Changes made

* Dd
  + Removed file column
  + Removed column long name
  + Updated column headers
  + Added dd and flmd column headers
  + Delete duplicates
  + Spelled out units
  + Capitalized beginning of definition and added periods
  + Added DOI to Abeshu reference
  + Defined RC
  + Ran spell check
* flmd
  + add flmd and read me
  + added files within TrainTest\_IndividualGrains folder and removed the row for this folder
  + deleted GIS row
  + added DOI to Abeshu reference
  + ran spellcheck

Action or response needed

* dd
  + ~~fill in data and term type for rows that do not have it~~
  + ~~add reference system to units of lat/long (eg decimal\_degrees\_WGS84)~~
  + ~~add unit to TOT\_STREAM\_SLOPE~~
  + ~~spell out unit for “logSlope“ using the same format as other units~~
  + ~~There are multiple duplicate columns with differing column headers or definitions. These columns should be renamed if they have two different definitions. If they have the same definition, the column names should match, and the duplicated rows should be removed~~
    - “d50\_mm\_nexss” and “D50\_mm\_nexss”
    - “label”
    - “lat”
    - “long”
    - “site\_id” and two “site\_ID”
    - “stream\_order” and “StreamOrde”
  + ~~Change the column names if possible (no special characters, no spaces); use underscores only~~
    - 221114\_usgs\_...
      * site\_no to site\_no\_usgs
      * d50\_mm to d50\_mm\_usgs
      * site\_rc to site\_no
    - d50\_Train1\_Prediction1\_40
      * photo width (pixel) to photo\_width\_pixel
      * photo height (pixel) to photo\_height\_pixel
      * resolution x(m/pixel) to resolution\_x\_m\_pixel
      * resolution y (m/pixel) to resolution\_y\_m\_pixel
      * width (m) to width\_m
      * heigth (m) to height\_m
      * long axis (m) to long\_axis\_m
      * YOLO box number to YOLO\_box\_number
      * YOLO accuracy (%) to YOLO\_accuracy\_percent
      * area covered (%) to area\_covered\_percent
      * YOLO threshold (%) to YOLO\_threshold\_percent
    - Nhd\_yrb\_D50\_variable\_updated
      * D50m\_YOLO to d50\_mm\_yolo
      * SiteID to site\_id
    - RC2 Spatial Study\_Responses\_Form
      * ID to site\_id
      * Latitude\_[dd] to lat
      * Longitude\_[dd] to long
    - RC2\_all\_D50
      * StreamOrde to stream\_order
      * site\_ID to site\_id
    - RC2\_all\_stream\_Attributes\_updated\_VGC
      * StreamOrde to stream\_order
      * site\_ID to site\_id
  + ~~Update definition for “YOLO threshold (%)”~~
  + ~~Can you please clarify the definition for the column “count”? What are the different scales and where does the number 50 come from~~
    - Thanks, updated here to reference the scaling factor and why it’s needed
  + It might be helpful to define catchment vs basin in the critical details section of the readme
* Flmd
  + If possible (and if it won’t break the scripts), the NA values should be changed to N/A for non-numeric columns and -9999 for numeric columns
  + ~~If “Table1.xlsx” is supposed to be included in the data package, it needs to be included as a csv (or pdf/jpg if formatting is important). It also needs to be added to the flmd~~
  + ~~Please check file descriptions for the .dat files. Please also add further description as to where the .dat file came from and how it connects to the photograph~~
  + ~~Please update description for “This file contains information about images used for YOLO d50 estimates, YOLO d50 estimates, and YOLO model diagnostics.“ YOLO d50 estimates is listed twice~~
  + ~~If you plan to publish anything outside of this folder (ie scripts), they need to be included in the flmd~~
* ~~Do not publish GIS data~~
* ~~Please include the flmd in the data package (it is not currently on github)~~
* ~~Consider renaming “abeyshu\_fig1\_digitzed.csv” to “abeshu\_fig1\_digitzed.csv”. If this change is made, update the flmd.~~