Land Use Documentation for 2014 INSIGHT

Amy Wright, Integrated Water Management Analyst

December 31, 2013

Background

A geographic information system (GIS) was used to create a spatially distributed historical land use/cover dataset for INSIGHT. The land use datasets were originally developed individually, and in phases to correspond to NDNR integrated model development for annual Fully Appropriated Basin (FAB) analyses. This integrated modeling framework is a linked system of ground water models, surface water operation models (where applicable), and soil water balance (i.e. watershed) models. The land use datasets provided a basis for assumptions regarding water use for soil/water balance modeling within this framework. Table 1 provides more details about the modeling frameworks that the land use datasets were created for.

Model	Acronym	Model cell size	Model/land use time period	Year(s) land use was developed	Land Use Created by
Upper Niobrara White	UNW	1 mile x 1 mile	1950-2005	2010-2011	NDNR
Central Nebraska	CENEB	1 mile x 1 mile	1940-2011	2011-2012	NDNR
Blue Basin	Blue	½ mile x ½ mile	1950-2012	2012-2013	NDNR
Western Water Use	WWUM	¼ mile x ¼ mile	1953-2010	2012	Leonard Rice Engineers, Inc.
Lower Platte/Missouri Tributaries	LPMTribs	1 mile x 1 mile	1950-2012	2013	NDNR

Table 1: Details about land use development and associated models for 2014 INSIGHT sub-basins.

There were many overlaps between models, due to the fact that groundwater reservoirs do not follow surface water basin boundaries and overlying aquifers often overlap. To prevent double

counting of acres, model cells were assigned to particular INSIGHT sub-basins, and only the land use data associated with that model were used. Figure 1 shows the models from which land use data were acquired to create a seamless land use dataset for the 2014 INSIGHT sub-basins.

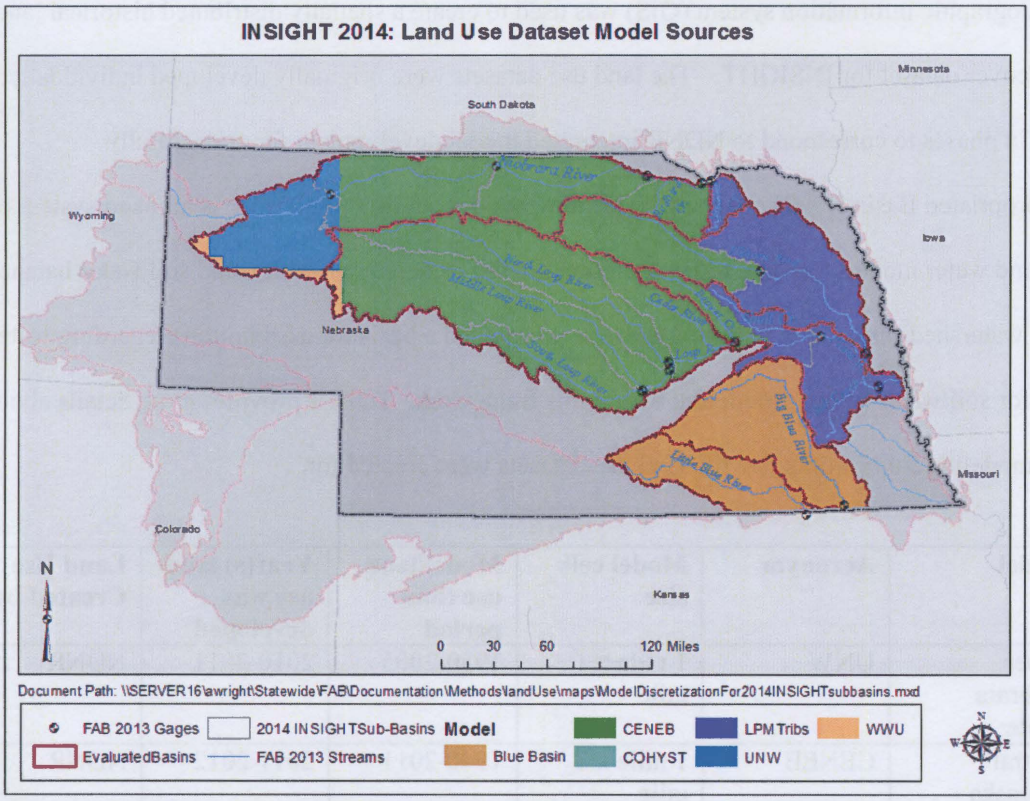


Figure 1: Model assignments within INSIGHT sub-basins, from which land use data was acquired.

Development of Irrigated Lands Datasets

Cropped lands were categorized with the source of irrigation such as groundwater, surface water or both (comingled), or dryland farmed. Residual lands were categorized as rangeland. The distribution of irrigated lands was developed using a combination of data sources, including the Natural Resources Districts (NRDs) digitized certified acres, county assessor digitized parcel boundaries, the Center for Advanced Land Management Information Technologies (CALMIT)

2005 land cover dataset, and NDNR surface water and groundwater data. NRD-digitized certified acres were considered the authoritative dataset to show the extent of groundwater irrigated lands and were used whenever available. In the absence of digitized certified acres, county assessor land use datasets were used if available. If these were not available, an updated version of the 2005 CALMIT irrigated parcels dataset was used. Updates were created by digitizing lands associated with post-2005 wells, identified by interpretation of the most recent Farm Service Agency (FSA) aerial imagery. NDNR surface water rights digitized fields were used to show the extent of surface water irrigated lands across the entire study area.

The historical component of groundwater irrigated lands was created by tying digitized fields to wells, with the assumption that the land could be potentially irrigated in the first year that a well was installed for that parcel. The historical component of surface water irrigated lands was created by tying digitized surface water parcels to points of diversion database, which contains information about the year the water right was granted. As such, all surface water lands were considered as potentially irrigated lands from the onset of the granting of a surface water right. Irrigation type was assigned as comingled in areas where surface water and groundwater irrigated lands overlapped. Figure 2 shows snapshots of potentially irrigated land through time.

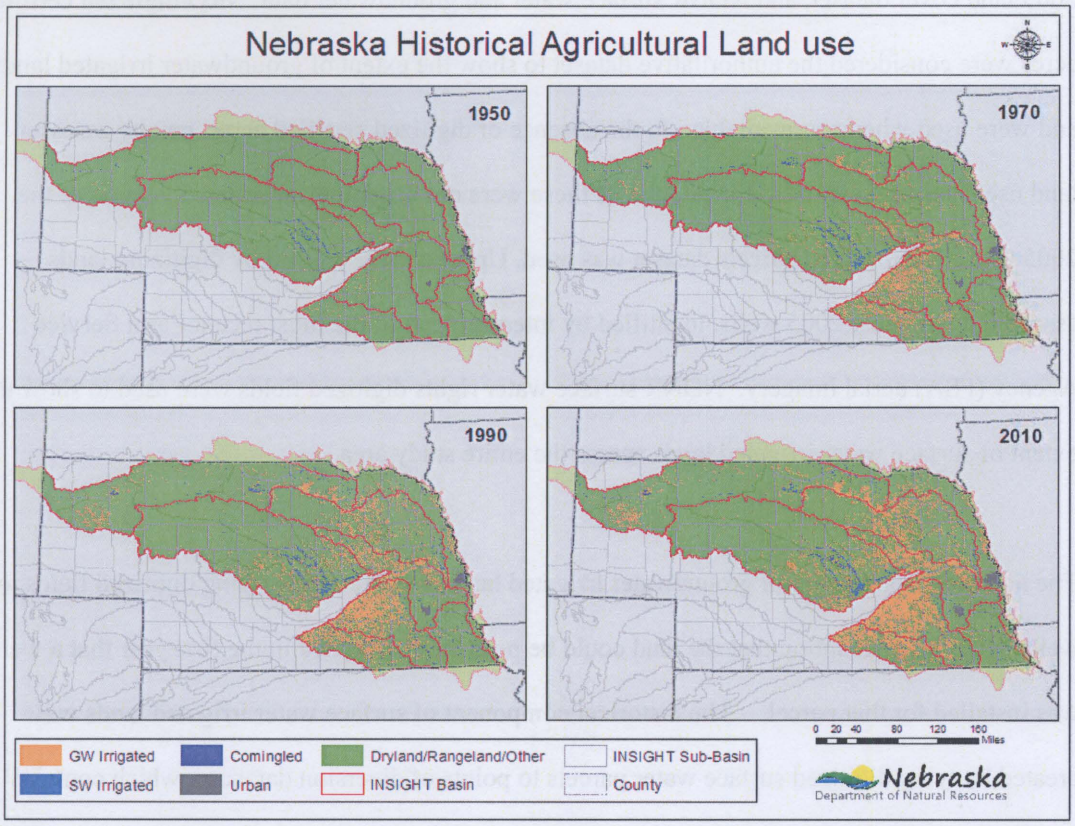


Figure 2: Snapshots of potentially irrigated lands through time.

Development of the dryland spatial dataset

There was limited data available to create a historical dryland spatial dataset. The 2005 CALMIT land use dataset was the only spatial dataset that showed the extent of dryland over all model areas. There were no historical time component data (e.g., well installation year, surface water right grant year) that could be attributed to dryland. Hence, the creation of historical dryland spatial dataset relied on two overarching assumptions:

- Lands that were dryland farmed in 2005 had potential to be dryland farmed historically
- Lands that are currently groundwater irrigated had potential to be dryland farmed prior to a well installation.

These assumptions are consistent with the general history of arable land cultivation in agricultural areas in the Midwest and Plains. By applying these assumptions, a "potential dryland" spatial dataset was created by combining the irrigation dataset with dryland areas extracted from the 2005 land use dataset. The resulting dataset represented all areas of land that could potentially be dryland farmed in any given year, provided the land was not being used for groundwater irrigation. A GIS program was implemented to partition the field-scale land use data into model cells to show the extent of potentially irrigated and dryland farmed lands at the model cell level.

Scaling to derive estimates of actual irrigated lands/dryland

It was recognized that the full extent of cropped lands defined as described may or may not actually be farmed in a given year. Therefore, the dryland and irrigated acreages were scaled according to county-level aggregate data from the National Agricultural Statistics Service (NASS), which provides estimates of irrigated and dryland acres for the model time periods, which are extrapolated from producer reports. The land use data derived from GIS were aggregated to the county level, and then scaled to NASS data using a 5-year moving average so that the estimates and trends in NASS data were captured. It should be noted that an exact match was not desired due to potential reporting discrepancies within the NASS data, and also because the NASS data that was used was based on harvested acres and did not account for acres lost to acts of nature. Because of inconsistencies in the pre-1960 NASS data, the scaling for the years prior to 1960 was based on 1960 NASS data values.

The 2005 CALMIT land use dataset estimates of irrigated and dryland acres were also plotted as a spot check for most areas. The 2005 CALMIT land use dataset was considered the best data source to show actual irrigated and dryland acres for this year only, because it was based on actual observations of the land via remote sensing.

Figure 3-5 shows county scaling as applied for the entire CENEB, LPMTribs and Blue Basin model areas. Here, county acreages were summed to portray the entire model area estimates of 1.) potential for irrigated and dryland acres, 2.)NASS reported acres, 3.) the adjusted acres (potential acres scaled to NASS data) and 4.) the CALMIT data estimates that were used as a spot check.

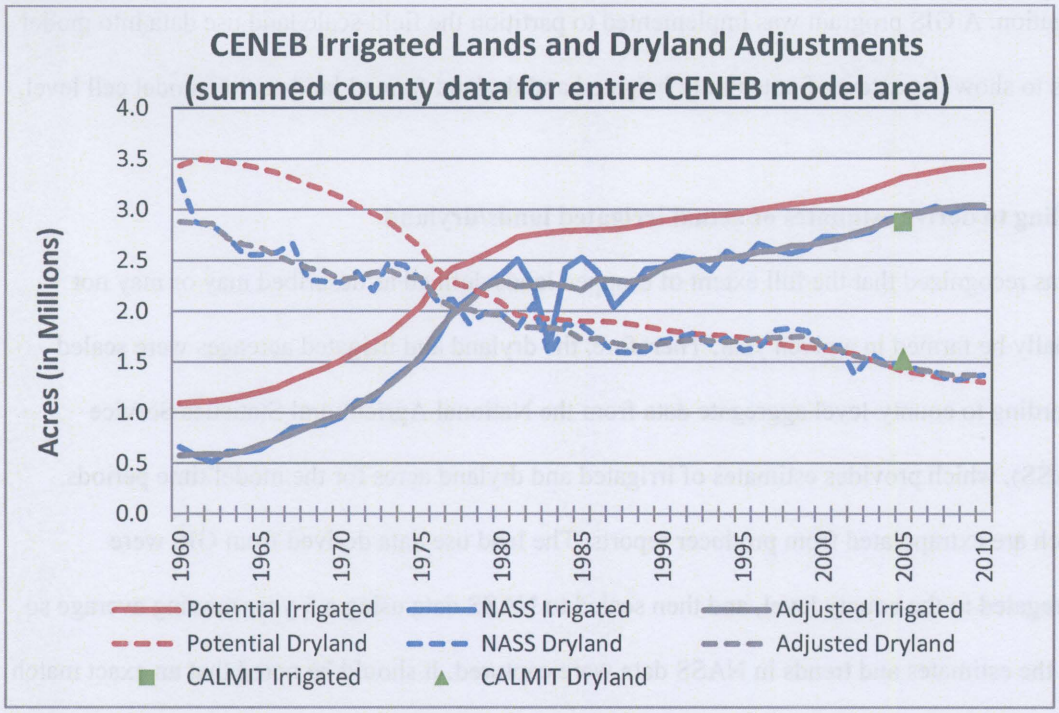


Figure 3: CENEB model aggregated irrigated and dryland acres representations.

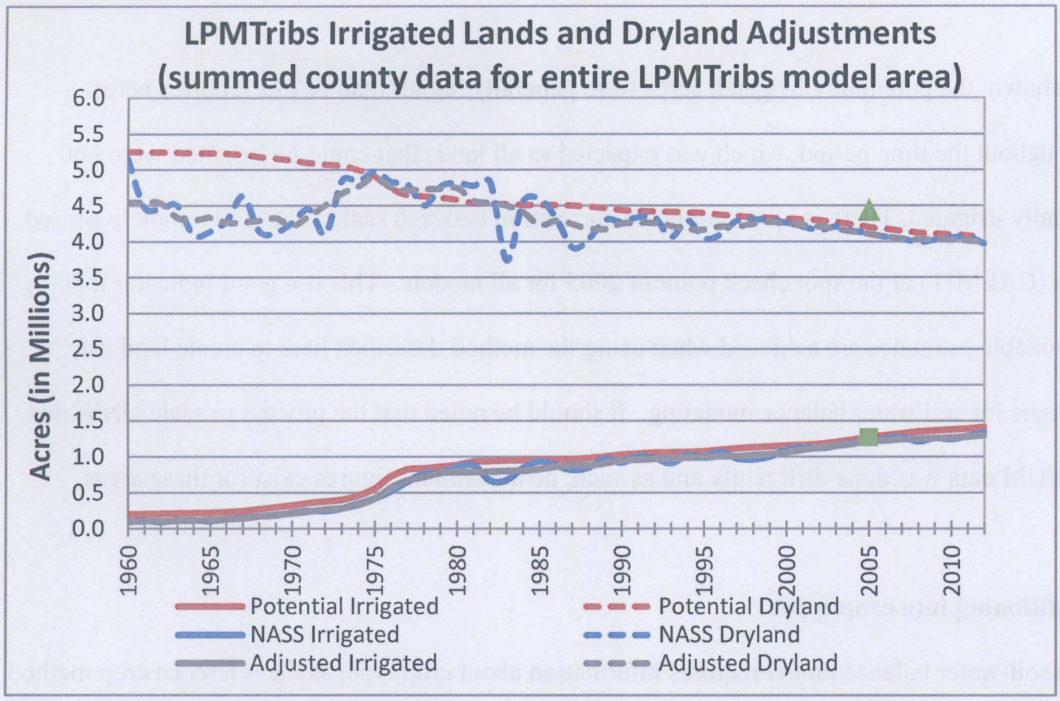


Figure 4: Lower Platte Missouri Tributaries model aggregated irrigated and dryland acres representations.

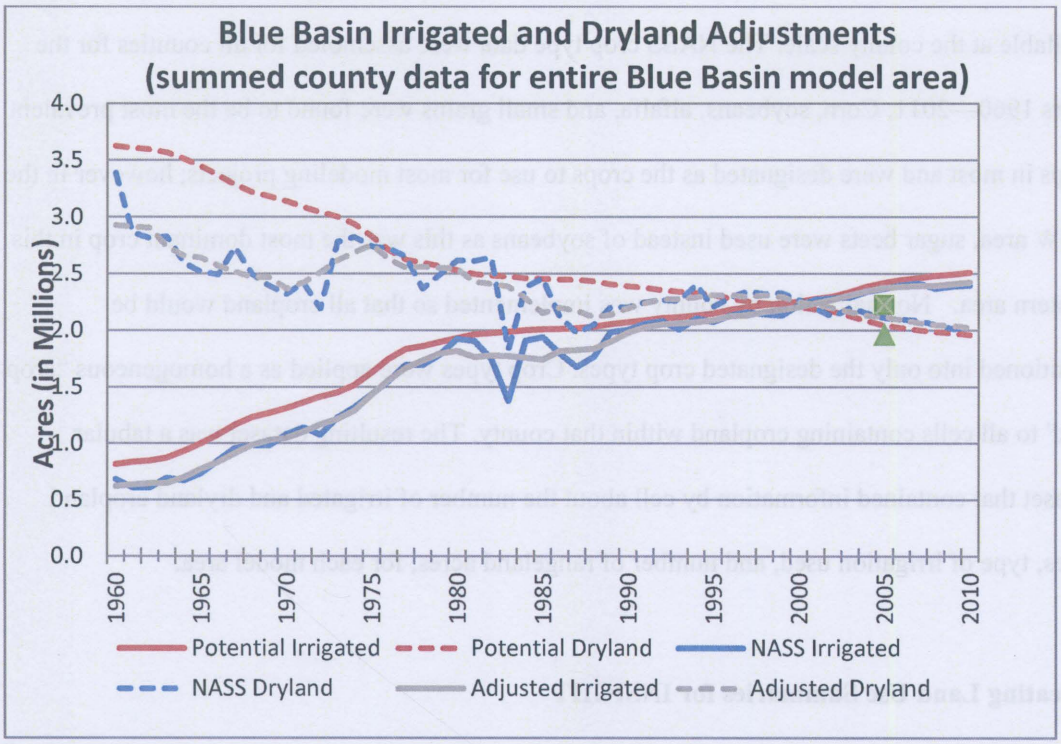


Figure 5: Blue Basin model aggregated irrigated and dryland acres representations.

As shown, the potentially irrigated acres were generally higher than NASS reported acres throughout the time period, which was expected as all lands that could be irrigated were not actually irrigated. There is also reasonable correlation between scaled data and remotely sensed data (CALMIT) at the spot check point of 2005 for all models. This is a good indicator that reasonable estimates are achieved when using the method described here to create land use datasets for soil water balance modeling. It should be noted that the process to scale UNW and WWUM data was done differently and as such, no adjustment figures exist for these areas.

Partitioning into crop types

The soil-water balance model requires information about crop type, as the reference crop method takes into account that different ET levels occur for different crop types. The NASS dataset was the only available resource to show crop distribution through time. Again, these data were only available at the county scale. The NASS crop type data were assembled for all counties for the years 1960—2011. Corn, soybeans, alfalfa, and small grains were found to be the most prevalent crops in most and were designated as the crops to use for most modeling projects; however in the UNW area, sugar beets were used instead of soybeans as this was the most dominant crop in this western area. Normalization by county was implemented so that all cropland would be partitioned into only the designated crop types. Crop types were applied as a homogeneous "crop mix" to all cells containing cropland within that county. The resulting dataset was a tabular dataset that contained information by cell about the number of irrigated and dryland cropland acres, type of irrigation used, and number of rangeland acres, for each model area.

Creating Land Use Summaries for INSIGHT

Information was added to each tabular dataset, to denote which cells would be pulled for INSIGHT sub-basin analyses. As noted earlier and shown in Figure 1, model assignments were

made to prevent double counting of acres. These assignments were tied to individual model land use datasets using the unique identifier for cells within that model. The data that were extracted from each individual land use dataset were then merged to create a seamless dataset for the entire area to be evaluated as a part of INSIGHT 2014. From here, sub-basin information was added to the combined datasets, and the acres were summed according to crop type and irrigation practice for each sub-basin, for each year between 1985 and 2012. If later years were not available for a particular dataset (e.g. CENEB only went to 2011, and UNW only went to 2005) the data for the latest year was copied for subsequent years. For more information on the INSIGHT sub-basin aggregation process, please see appendix A.

APPENDIX A
Creation of INSIGHT Subbasin Land Use Aggregates

Amy Wright

August 12, 2013

Input Dataset Locations

- Land Use Datasets from specific models
  - LPMTribs
    Server16\\awright\LPMTribs\LandUse\Processing\LPMTribsFinalLUdataset.gdb\LPMTribsLandUse1950_2012
  - Blue Basin:
    Server16\\awright\BlueBasin\allLandUse\FinalDataset\BlueLandUseAug2013.gdb\BlueLandUse1940_2011; BlueLandUse_2011as2012only
  - CENEB:
    Server16:\\awright\CENEB\LandUseAll\Processing.gdb\LUgrid1940_2011_wCropTypes_July2013
  - UNW
    Server16:\\awright\UNW\UNWlandUseForINSIGHT.gdb\UNWlanduse_1960_2005; UNWluse_1960as1950to1959; UNWluse_2005as2006to2012
- Model grids, INSIGHT cells only, with attributes to link to the NeGrid and INSIGHT attributes:
  - Server16\\awright\FAB\INSIGHT\StudyArea\INSIGHTgridAssignments.gdb\INSIGHTcellsByModel\BlueBasinINSIGHTcells and CenebINSIGHTcells and LPMTribsINSIGHTcells and UnwINSIGHTcells.

Working Directory: Server16\\awright\FAB\INSIGHT\SubbasinAggregates\LandUseBySubbasin.gdb

Process Steps

1. Ensure each land use dataset has the following populated attributes: Year, INDEX, AdjAcres_GW, AdjAcres_SW, AdjAcresComg, AdjAcresDry, Rangeland (or RangelandUrban), overage.
2. The attributes and information shown in Table 1 should be added to existing land use datasets, with the resulting table appearing similar to Figure 1.

Attribute	Properties	Explanation	Values	Notes
INSIGHTcell?	Text, 5	Tells if this should be pulled for INSIGHT analysis	"Yes" or "No"	Records attributed with "Yes" are those that successfully joined to the associated model INSIGHT cells, AND have a year of 1950-2012.
Model	Text, 15	When the data are combined, this field will tell which model it came from	LPMTribs CENEB Blue Basin UNW	All records have the same value in specific land use datasets.
NEgridRCID	Text, 10	This will contain the NE grid index values	e.g. '64-430', etc. for those that are	Join the records of the model INSIGHT cells to

		from the INSIGHT model cells layer.	INSIGHT cells; otherwise "None".	the land use datasets, and populate with values from the NEgridRCID field. For remaining NULL values, AND years <1950, populate with "None".
INSIGHT_ Subbasin	Text, 50	This will contain the values from the INSIGHT model cells layer for the INSIGHT subbasin name.	"	With join in place, calculate field = Sub_basin from model INSIGHT cells layer.

Table 1: Attributes/information to add to existing land use datasets for INSIGHT aggregations.

Table

BlueLandUse1940_2011

INDEX_	Year_	CountyYear	RnglandUrban	AdjGWirr	AdjSWirr	AdjComg	AdjDry	Overage	INSIGHTcell?	NeGrid_RCID	INSIGHTsubbasin	Model
10-228	2010	Butler_2010	160	0	0	0	0	0	No	None	None	Blue Basin
10-228	2011	Butler_2011	160	0	0	0	0	0	No	None	None	Blue Basin
10-229	1940	Butler_1940	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1941	Butler_1941	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1942	Butler_1942	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1943	Butler_1943	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1944	Butler_1944	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1945	Butler_1945	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1946	Butler_1946	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1947	Butler_1947	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1948	Butler_1948	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1949	Butler_1949	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1950	Butler_1950	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1951	Butler_1951	159.949676	0	0	0	0.050326	0.000002	No	None	None	Blue Basin
10-229	1952	Butler_1952	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1953	Butler_1953	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1954	Butler_1954	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin
10-229	1955	Butler_1955	159.949676	0	0	0	0.050327	0.000003	No	None	None	Blue Basin

1 (0 out of 3514320 Selected)

BlueLandUse_2011as2012only | BlueLandUse1940_2011

Figure 1. Individual model land use dataset with added information for INSIGHT aggregates (RnglandUrban was created to combine Rangeland and Urban Acres).

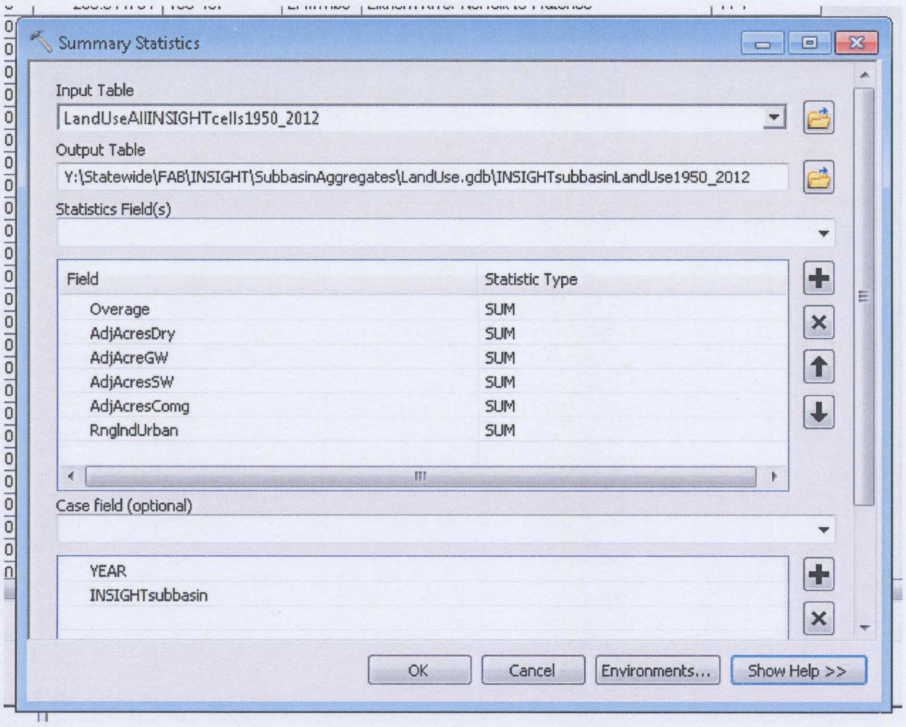
3. ***QA:*** *QA each land use dataset to ensure there is the same number of features for every year.* To do this, select by attributes INSIGHTcell? = "yes", and perform sum stats on the year column for selected features, then sort the "count field" of the output to ensure all have the same value. Also ensure no years are missing (1950 – 2012 = 63 years/records in output).[1]
4. Ensure all model datasets have the years 1950 to 2012 available. If necessary, create an additional year(s) of data, for those models that only go to 2010 or 2011, i.e. Blue Basin, CENEB and UNW. To do this, select the most recent year of data that corresponds to INSIGHT cells (YEAR = 2011 AND INSIGHTcell? = "Yes"). Export selected features to create a new feature class called e.g. BlueBasin2011as2012 (if 2010 isn't included, copy and paste an additional year of most recent data into table). Store it in the same geodatabase as the original data. Within that dataset change all fields that contain e.g. "2011" as the year, to 2012.

DECISION: If a model does not have all years available, use the nearest year and copy the data for the missing year(s), changing only the year in the attributes.

5. ***QA:*** *Ensure Rangeland (or RangelandUrban) and Overage fields are calculated properly.* There should be no negative values, and the total of land uses should equal 640 or 160 acres, depending on scale. If there are overages, the total will be more, but the overage column should specify the correct amount over 640 or 160 acres.
6. Export data from one land use dataset to (have all fields hidden except those shown in Figure 1, and select only the features that have INSIGHTcell? = Yes) to create a new geodatabase table called "FAB2013_allLandUse_1950_2012.

QA's:

1. **Ensure there is the same number of cells for all years.** This is done in both in the individual model land use tabled (choose only INSIGHT cells = yes) and then again after all the models' land use have been combined for INSIGHT (i.e. LandUseAllINSIGHTcells1950_2012). Use summary statistics on the year field to perform the QAs. In the output, ensure that all years have the same number of cells for a given model, or for the INSIGHT aggregate, by sorting both ascending and descending. Also look at the min, max and mean values to make sure there are no missing years. In this way, if there is model that is missing cells or an entire year of data, it will be obvious as the number of features will be different for that year, or the mean value will not make sense for the number of years represented. If these are not caught, those missing data would wrongly be attributed as rangeland.
2. **Compare the summarized model cells' residual rangeland to subbasin residual rangeland.** In this case, the summary of the summarized cells' rangeland, urban and the holes (in certain subbasins), minus the summary of the overages, is compared to the subbasin residual rangeland.
   When compared (cell vs. subbasin residual), differences in estimates of rangland were within 1% of total subbasin area. Looking at this another way, the differences ranged from -13 (subbasin rangeland is smaller than cell based aggregate), to +29 cells (subbasin rangeland is larger than cell based aggregate) cells of rangeland for specific subbasins in specific years, with an average difference of 3.6 cells for all subbasins by year. The differences are probably due to vector (subbasin) and raster (model cells) aggregations of rangeland The spreadsheet with rangeland comparisons is located at: Server16:\\awright\Statewide\FAB\INSIGHT\SubbasinAggregates\QAs\comparisonOfrangland.xlsx3.

3. **Compare the number of cells pulled from individual models to the number of cells for each model in the INSIGHT varied scale discretization.** Be sure to just pick the INSIGHT cells that will be used for the FAB2013 discretization.

Model	**From indiv. Models, the # of cells per year that are in INSIGHT area**	**# of total features**	**From INSIGHT model discretization, the number of cells attributed to each model**
CENEB	25,465	1,604,295	25,465
Blue Basin	28,921	1,822,023	28,921
LPMTribs	13,082	824,166	13,082
UNW	4147	261,261	4147
Total	71,615	4,511,745	71,615