AG RES BIOMASS RASTER BUILD PHASE

Comprehensive summary (Raster Construction and Validation)

1. Purpose and scope

This phase converted the harmonized, RM-corrected, ground-truth-scaled biomass dataset from tabular form into geospatial rasters compatible with GIS platforms.

The objective was to spatially allocate per-pixel biomass values while preserving RM- and municipality-level variation and ensuring consistency with MASC ground-truth data.

The phase also addressed issues such as invalid values, border leakage, and visual discontinuities.

2. Data inputs

Primary inputs included:

\* `aci\_biomass\_per\_pixel\_<year>.csv` – biomass data by RM, municipality, and crop Label.

\* Annual Crop Inventory (ACI) raster – `aci\_<year>\_mb\_v\*.tif`.

\* `municipalities.geojson` – boundaries of all Manitoba municipalities.

\* `Manitoba\_Provincial\_Boundary\_2830875....geojson` – official provincial boundary.

\* Crop classification reference (`aci\_crop\_classifications\_iac\_classifications\_des\_cultures.csv`) mapping crop Labels to integer ACI Codes.

Each row in the CSV defines a unique RM × MUNI\_NAME × Label combination with associated acreage, yield, biomass, and pixel counts.

Each pixel represents 30 by 30 metres, corresponding to approximately 0.22 acres.

The “tonnes per pixel” metric therefore represents absolute biomass mass at the pixel level.

3. Raster outputs

Two rasters were created for each year:

\* `biomass\_codes\_<year>.tif` – integer crop codes (uint16) with nodata value 0.

\* `biomass\_values\_<year>.tif` – biomass in tonnes per pixel (float32) with nodata value -9999.0.

Both rasters retain the projection, extent, and resolution of the ACI base raster and use LZW compression.

This structure allows categorical and continuous data to be handled efficiently and displayed separately in GIS.

4. Assignment logic

Processing was conducted at the municipality level.

Each municipal polygon was converted to the raster’s coordinate reference system, and a bounding window was extracted using the ACI raster.

A mask limited pixel operations to the area inside the polygon.

Biomass assignment followed two steps:

1. Pixels whose existing crop code matched the crop’s ACI code were directly assigned the municipality’s biomass value.

2. If the number of existing pixels for a crop was below the expected `aci\_pixels` value, additional pixels were randomly selected and assigned that crop and biomass value.

Only unassigned pixels within the polygon and outside of protected classes (e.g., water, forest, urban) were eligible.

All assignments were recorded in a diagnostic log (`raster\_assignment\_summary\_<year>.csv`) for verification.

5. Early technical issues

\*\*Invalid biomass values\*\*

Rows with zero acreage in both ACI and MASC data caused division-by-zero errors, producing infinite values.

These rows were removed during preprocessing, ensuring all subsequent biomass values were finite.

\*\*Performance\*\*

Initial implementations processed the full raster globally.

This approach was replaced with per-municipality window processing and direct writing to disk, which significantly reduced computation time and memory usage.

\*\*Border artifacts\*\*

Early raster versions contained biomass “bleeding” into Saskatchewan, North Dakota, and major lakes.

This occurred because rectangular geometry windows extended beyond the Manitoba boundary.

Introducing a per-window geometry mask confined assignments to the actual polygon shape, resolving the problem.

\*\*Overlapping boundaries\*\*

A set of twenty-four municipalities extended slightly outside the provincial boundary due to source data geometry.

All municipal polygons were clipped to the provincial boundary before rasterization, ensuring all pixels remained inside Manitoba.

\*\*Projection consistency\*\*

All vector layers were reprojected to the raster’s Albers Equal Area CRS, preventing alignment discrepancies between the raster and vector data.

6. Biomass scaling and normalization

Earlier versions used mean biomass per crop label, collapsing spatial variation.

The logic was revised so that each RM × MUNI\_NAME × Label record retained its specific biomass-per-pixel value.

This restored full spatial differentiation and produced a raster with biomass values identical in range to the tabular source data.

Validation confirmed:

\* The raster’s biomass range matched the CSV range of approximately 0.004 to 8.26 tonnes per pixel.

\* The total number of valid biomass pixels corresponded to the expected cultivated area of Manitoba.

\* The summed biomass total was within a few percent of the tabular provincial total, a discrepancy attributable to rounding, clipping, and exclusion of protected land classes.

7. Visual discontinuities along boundaries

Because each RM’s biomass was normalized independently to MASC ground-truth data, visible transitions appeared at administrative boundaries where adjacent RMs have differing per-pixel biomass values.

This pattern is common in datasets calibrated at the administrative level and is also visible in federal and international yield mapping products.

Three approaches were evaluated:

\* Full Gaussian smoothing across all pixels, which removed boundaries but diluted real spatial variation.

\* Weighted Gaussian smoothing using a valid-pixel mask, which preserved nodata handling but still blurred interior patterns.

\* Boundary-only smoothing applied within a narrow buffer around RM boundaries, which retained interior values and visually softened seams.

The smoothing steps were left optional and excluded from the final script.

Cartographic smoothing will instead be applied in ArcGIS Pro using Focal Statistics and Raster Calculator.

8. Nodata and visualization

\* Code raster: nodata = 0, displayed as transparent.

\* Biomass raster: nodata = -9999.0, displayed as transparent.

GIS applications sometimes show truncated value ranges due to sampling; recalculating full statistics displays the true range.

9. Verification results

Tests confirmed that:

\* Raster projection and alignment match the ACI base raster.

\* No biomass is written beyond Manitoba borders.

\* Municipal and RM differentiation is preserved.

\* Nodata areas display correctly in GIS.

\* Total biomass differs from the tabular total by only a small percentage, within acceptable tolerance for rasterization and clipping.

10. File management and repository guidance

GitHub restricts files larger than 100 MB and recommends repository sizes under one gigabyte.

Because the raster files exceed these limits, they are excluded from version control.

Typical management practice includes:

\* Ignoring `/data/` and `/outputs/rasters/` directories in `.gitignore`.

\* Storing large files in external or institutional storage (e.g., ArcGIS Online, shared drives).

\* Keeping all scripts, metadata, and diagnostic CSVs under version control.

\* Using Git LFS only when smaller large files must be included.

11. Final workflow summary

The final raster build process performs the following operations:

12. Load processed biomass CSV, crop code reference, municipality, and provincial boundaries.

13. Select the corresponding ACI raster for the year.

14. Reproject and clip geometries to the raster CRS.

15. Initialize output rasters for codes and biomass values.

16. For each municipality:

\* Extract raster window and geometry mask.

\* Assign biomass based on municipal crop information and fill deficits where required.

\* Write the updated window back to disk.

17. Record pixel assignments and biomass statistics in a diagnostic file.

18. Produce the final `biomass\_codes\_<year>.tif` and `biomass\_values\_<year>.tif`.

12. Outcomes

The final rasters:

\* Accurately represent ground-truth-scaled biomass distributions across Manitoba.

\* Retain all RM-level differentiation.

\* Contain no invalid or cross-border pixels.

\* Display correctly with transparent nodata.

\* Provide a defensible, reproducible foundation for analytical and cartographic applications.

The outputs constitute the authoritative biomass surface for each modeled year, suitable for integration into ArcGIS Pro and ArcGIS Online environments.