

30m CONUS Elevation Derivatives

STED30

Standardized Elevation Derivatives at 30m spatial resolution

Digital Soil Mapping Working Group

Colby Brungard, PhD
Assistant Professor of Soil Science
New Mexico State University
Las Cruces, NM



Rationale

1. Digital soil mapping requires spatially-exhaustive environmental covariates.
2. The most commonly used environmental covariates are derived from a digital elevation model (DEM).
3. A standardized dataset of terrain derivatives would be useful
 - a) Necessary for CONUS-scale DSM

Benefits of a standardized terrain dataset

1. A reduced need to calculate terrain derivatives for each project
2. A greater ability to compare results between studies
3. Better documentation and explanation!

Challenges

1. Most derivatives have 'optimization' parameters (Sorensen et al., 2006)
 - a) Slope = algorithm, neighborhood size
2. Different landscapes = different soil-landscape relationships = different parameters and scales (i.e., neighborhood sizes)
3. Problem at large (CONUS) extents, how to choose the right parameters?

R. Sørensen, U. Zinko, J. Seibert. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrology and Earth System Sciences Discussions, European Geosciences Union, 2006, 10 (1), pp.101-112. ffh1-00304825f
<https://hal.archives-ouvertes.fr/hal-00304825/document>

Solution

1. Use multiple parameters and neighborhood sizes for each covariate
 - a) Calculated for 40 terrain derivatives
2. This resulted in 199 terrain derivatives at CONUS extent

Methods 1

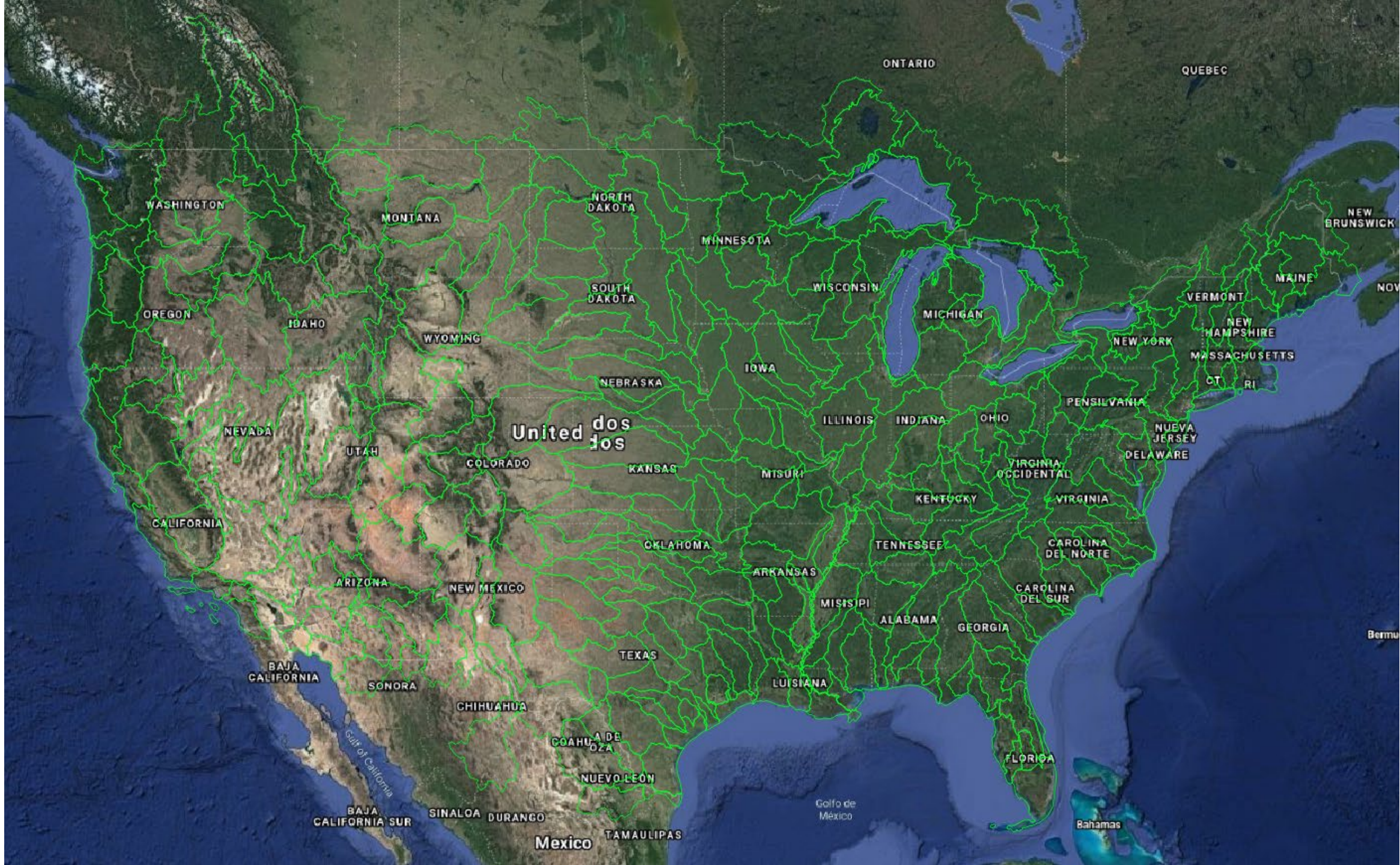
- Downloaded 1 arc sec (approx. 30m) DEMs from National Elevation Dataset.
 - Included parts of Canada and Mexico to avoid edge artifacts
 - 3370 DEM tiles
 - All NED elevation data is from the best available elevation data (i.e., LiDAR, resampled to 30m, where possible) – high vertical precision.
 - DEMs current as of 2/25/2021

Methods 1

- Downloaded 1 arc sec (approx. 30m) DEMs from National Elevation Dataset.
 - Included parts of Canada and Mexico to avoid edge artifacts
 - 3370 DEM tiles
 - All NED elevation data is from the best available elevation data (i.e., LiDAR, resampled to 30m, where possible) – high vertical precision.
 - DEMs current as of 2/25/2021
- DEM tiles came as both .tif and .img
 - Converted everything to .tif
- Mosaiced all DEM tiles into one CONUS-wide DEM with 30m resolution

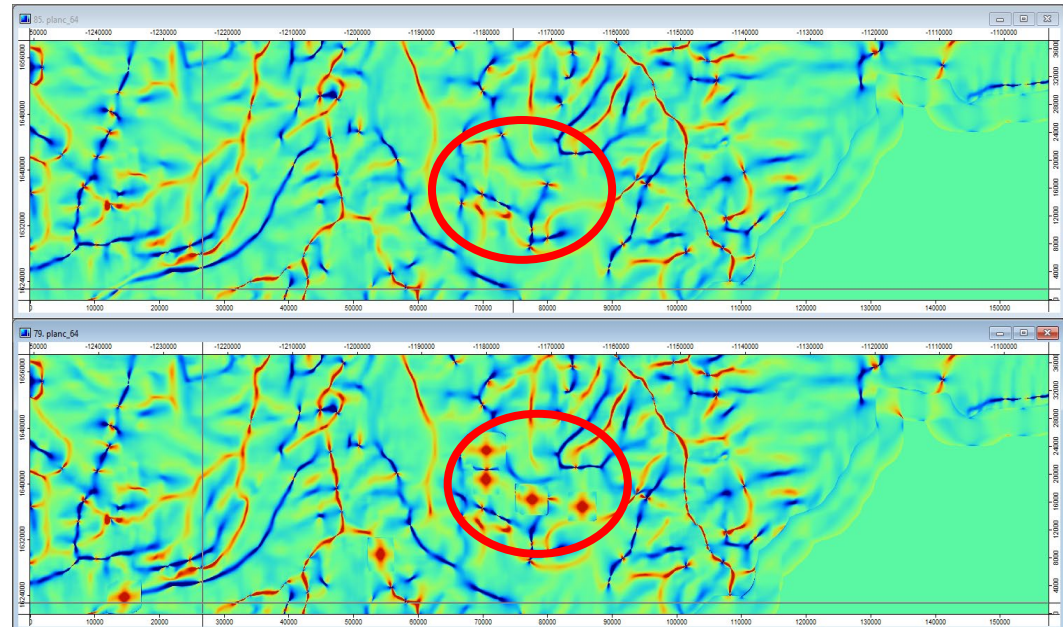
Methods 2a (Dave White, Las Cruces, NM field office did this)

- CONUS-DEM
 - Split CONUS-DEM into HUC-6 'watersheds'
 - Buffered 4 km
- Pitfilling (TauDEM)
 - Brought single cells (i.e. cells mostly likely data artifacts) to the level of surrounding cells so water would flow through. Very conservative approach.
 - Snapped to National Land Cover Dataset
 - Reprojected to Albers Equal Area (EPSG: 5070)



Methods 2b

- Waterbodies?
 - Tested with and without waterbodies > 4 acres removed
 - Removing waterbodies resulted in spurious artifacts in derivatives
 - Decided NOT to remove waterbodies (should probably mask waterbodies from final DSM products)
- Plan Curvature - 64 cell radius
 - Top waterbodies not masked
 - Bottom waterbodies masked



Methods 3

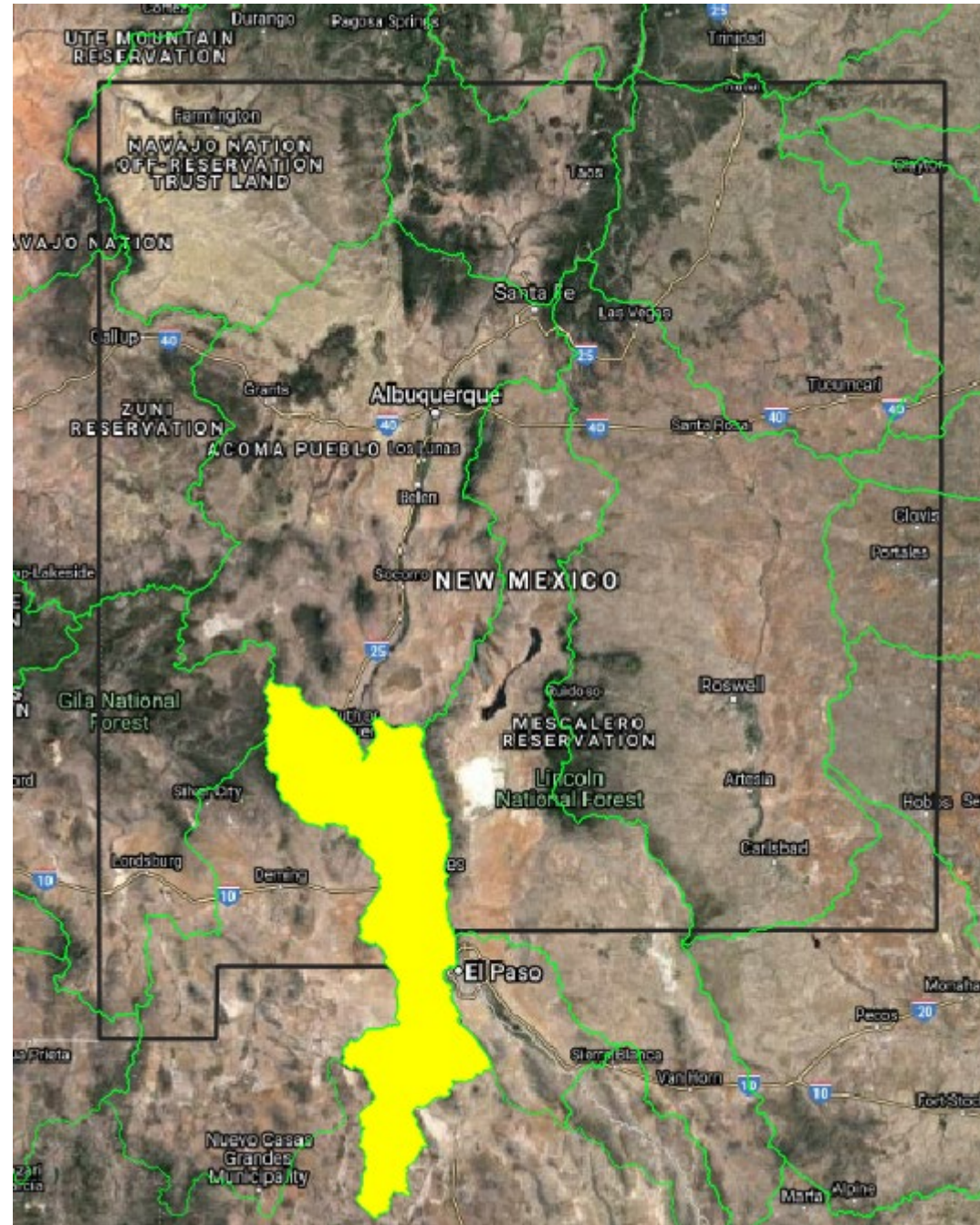
- On NMSU's High Performance Cluster
 - In parallel (on multiple nodes):
 - For each HUC-6 watershed {
 - Convert pitfilled DEM to saga grid format
 - In parallel (on multiple cores): calculated 40 elevation derivatives with multiple parameters
 - Took ~ 20 hours per HUC6 watershed
 - Trimmed off edges to 20 cell buffer
 - Converted derivatives to geotif, compressed with LZW compression or packbits
 - All no-data values set to -3.4e+38
 - Used SAGA GIS and gdal
 - All derivatives mosaiced into CONUS extent
 - gdal (overlapping cells, no feathering)
- *** still on-going ***

Explanation of Covariates

Example area

Example Area

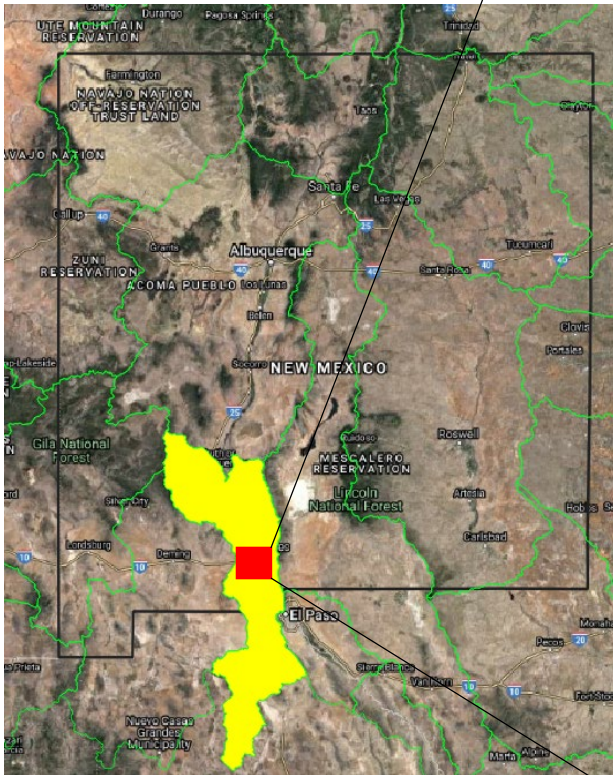
1. Southern New Mexico borderlands
2. HUC6 130301 (Rio Grande – Caballo)

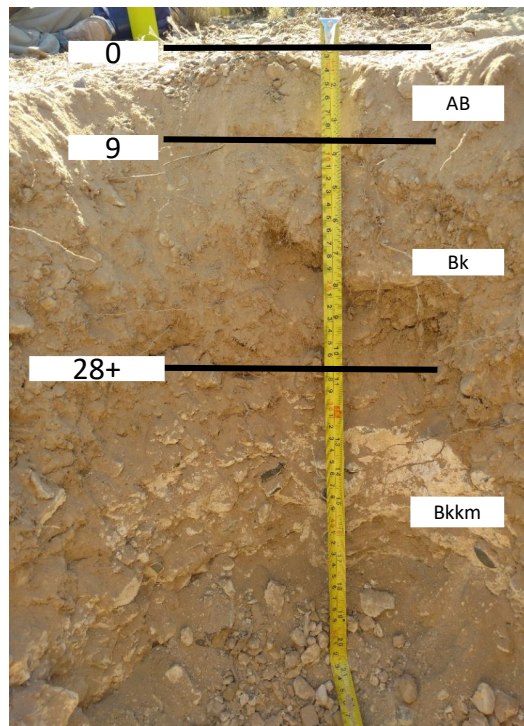


Example Area

Smaller area to view covariates

Approximate area of Desert Soil-
Geomorphology project





Explanation of Covariates

Lighting

Morphometry

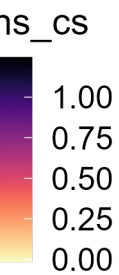
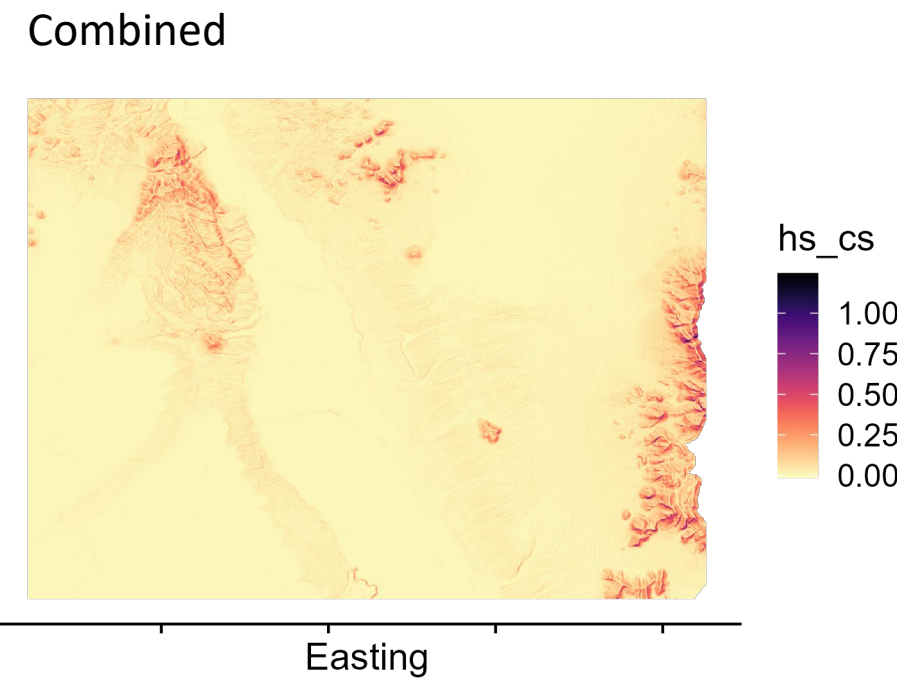
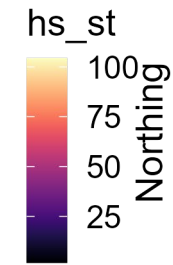
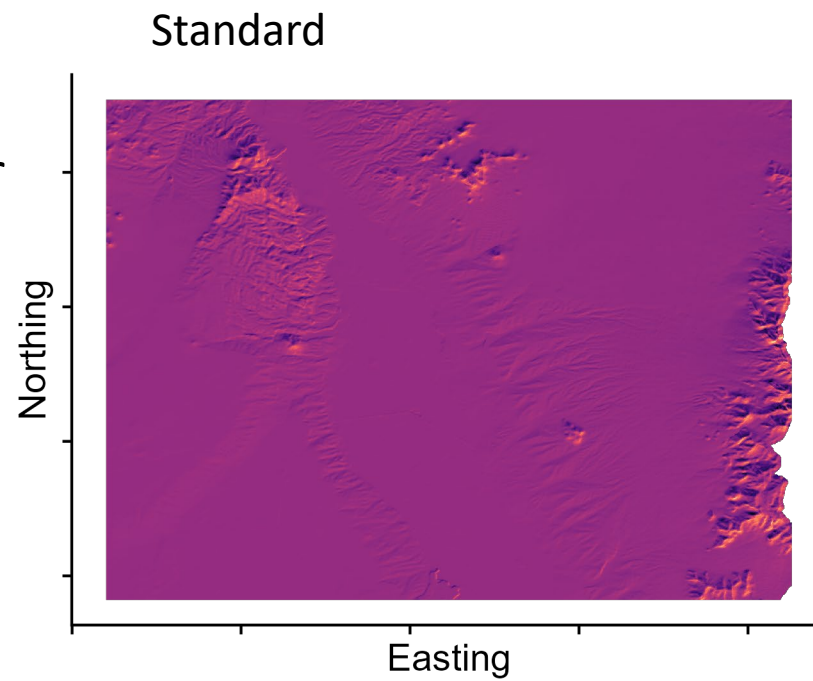
Hydrology

Lighting/Visibility

1. Hillshades
2. Potential Incoming Solar Radiation
3. Diurnal Anisotropic Heating
4. Topographic Openness

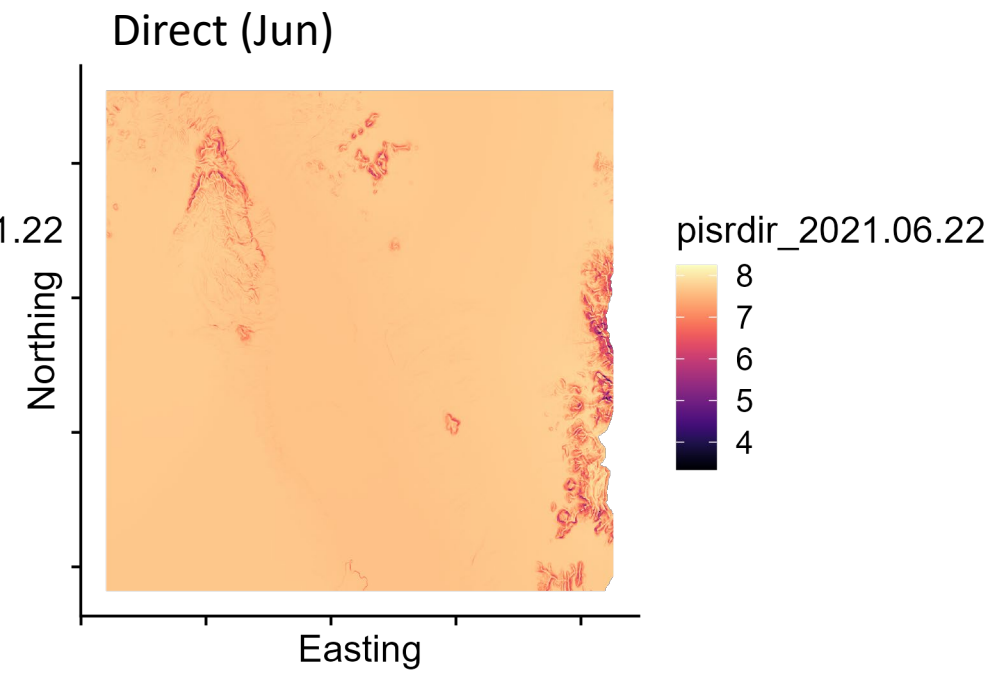
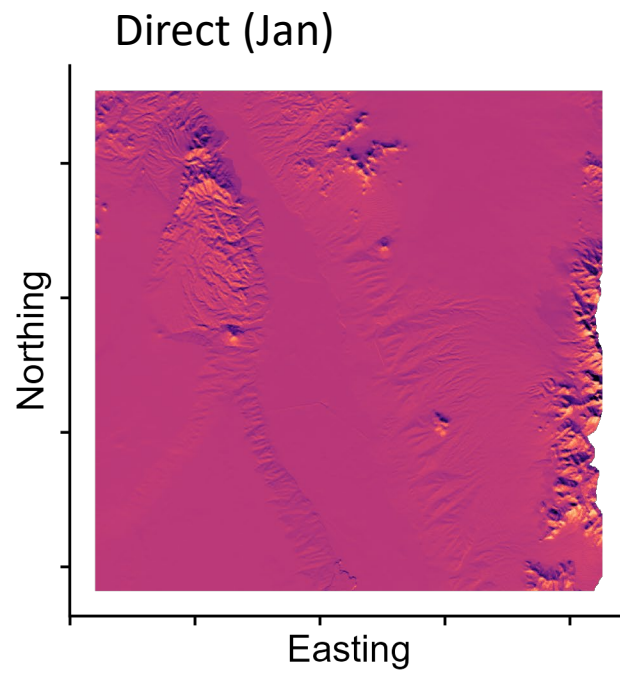
Hillshades

(mostly intended for visualization and not necessarily analysis)



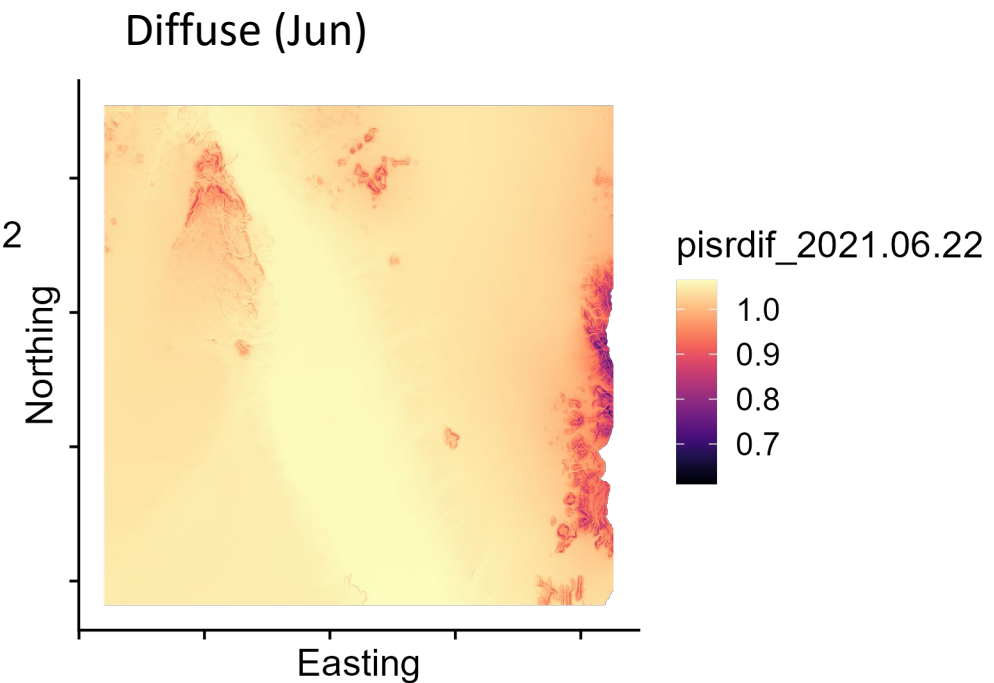
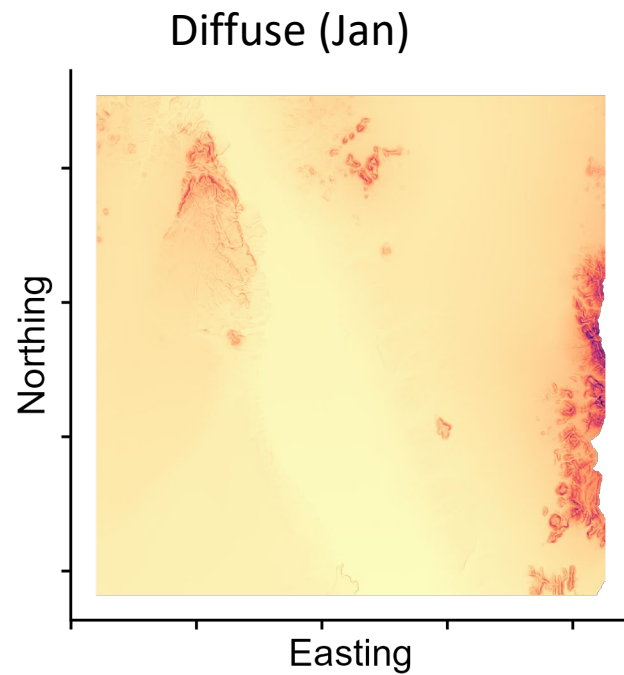
- Standard and combined hillshades
- Interpretation:
 - Low values = high incident radiation few shadows, high values = low incident radiation many shadows.

Potential Direct Incoming Solar Radiation



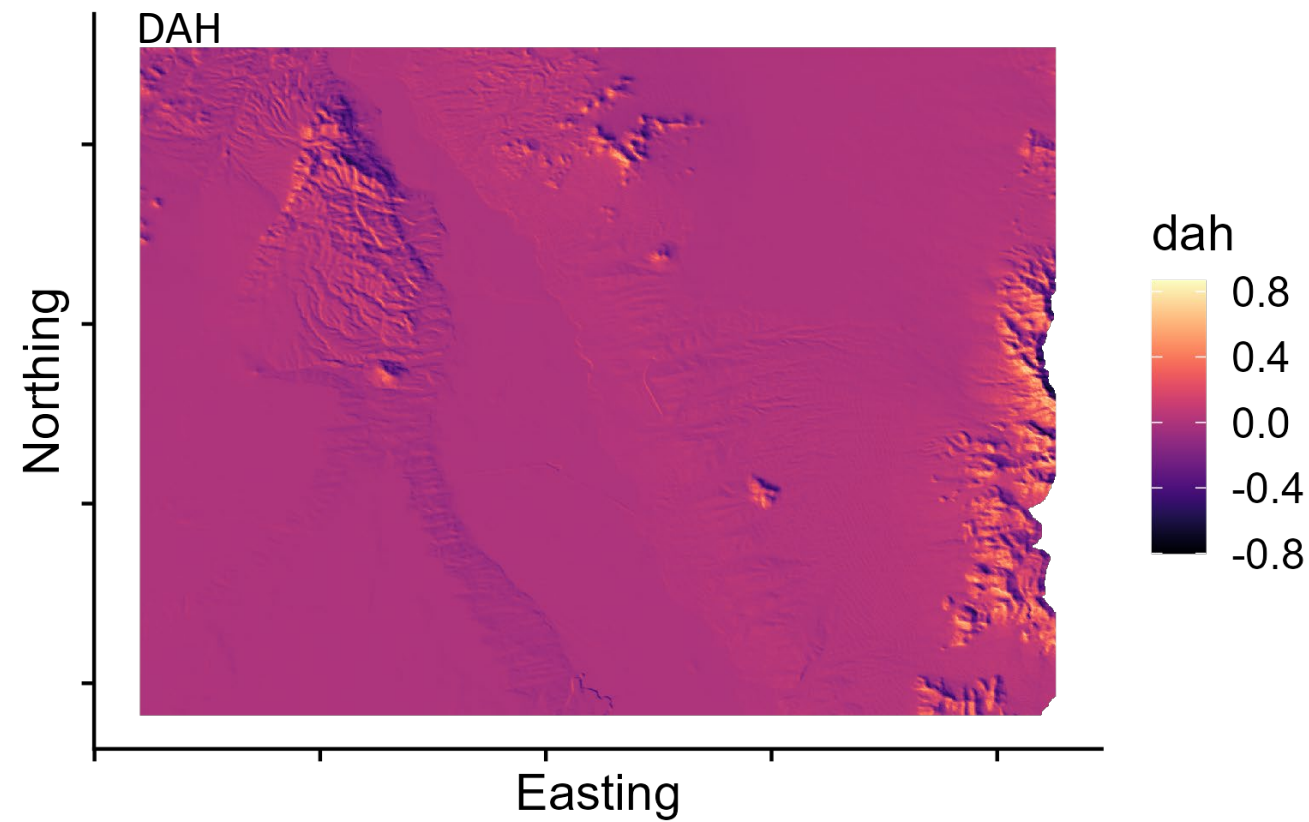
- Settings:
 - For the 22nd of each month (Jan – Dec) using a 4 hour time step and lumped atmosphere
- Interpretation:
 - Units are kWh m^{-2}
 - Direct = the energy directly from the sun that could potentially strike a cell for a single 24 hour day.
 - Intended as a better surrogate for aspect induced microclimates than aspect
 - Lower values indicate cells that should be cooler and moister than cells with higher values

Potential Diffuse Incoming Solar Radiation



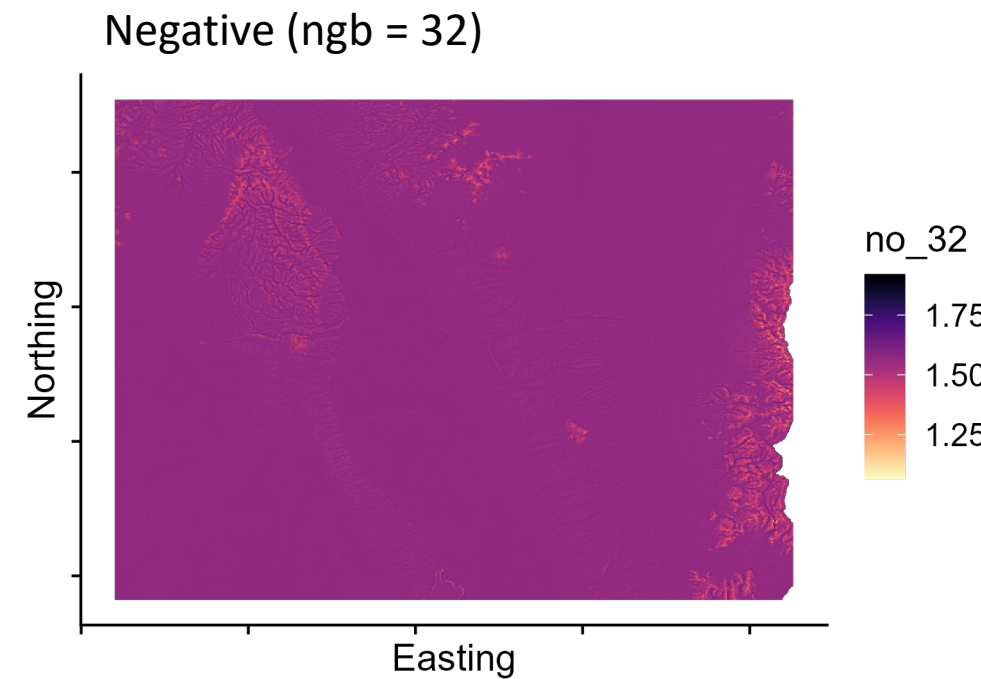
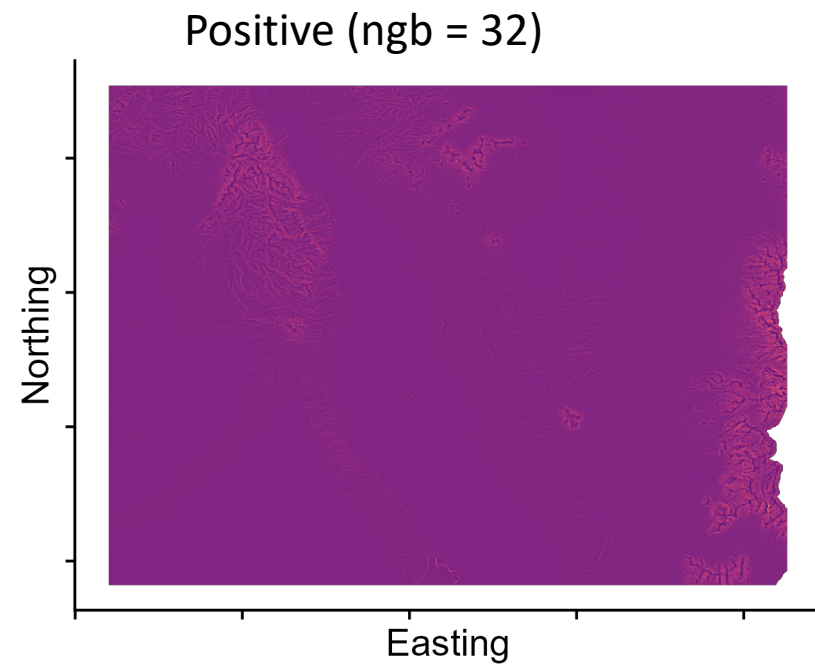
- Settings:
 - For the 22nd of each month (Jan – Dec) using a 4 hour time step and lumped atmosphere
- Interpretation:
 - Units are kWh m^{-2}
 - Diffuse = the amount of energy from the sky that could potentially strike a cell for a single 24 hour day.
 - Intended as a better surrogate for aspect induced microclimates than aspect
 - Lower values indicate cells that should be cooler and moister than cells with higher values

Diurnal Anisotropic Heating



- Settings:
 - Direction of sun set to 225 (southwest)
- Interpretation:
 - Dimensionless, ranges from -1 to 1, with 0 on flat slopes.
 - This is used to quantify differences in asymmetric heating between east and west slopes, where southwest slope are generally warmer and drier than are east facing slopes.
 - Can characterize topographically modified climate.

Topographic Openness

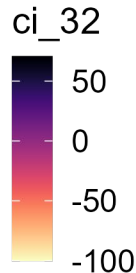
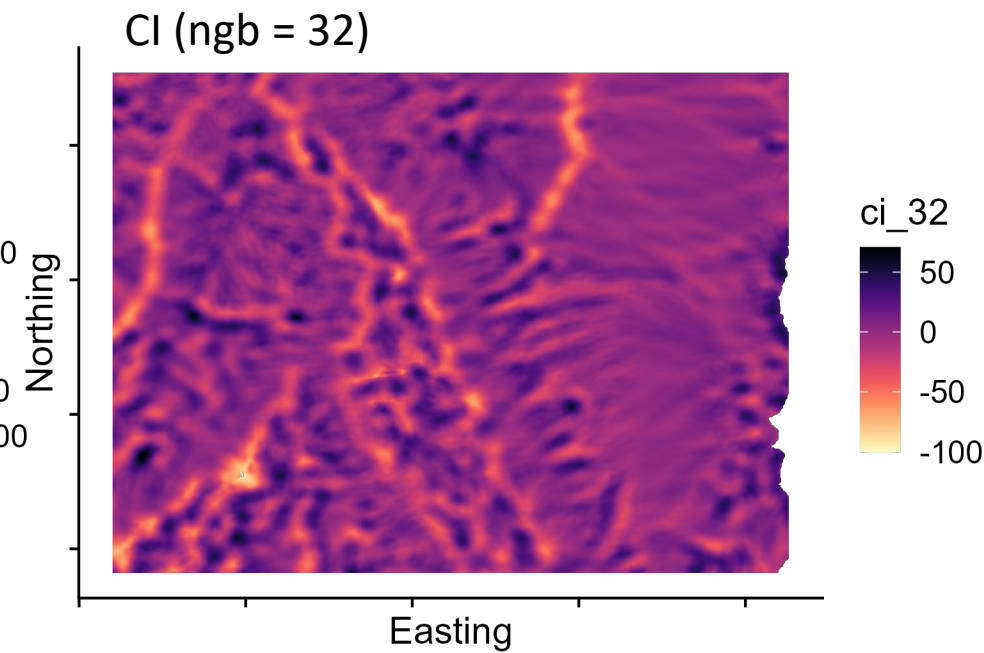
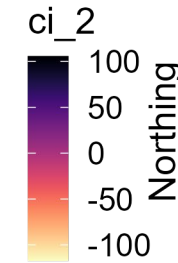
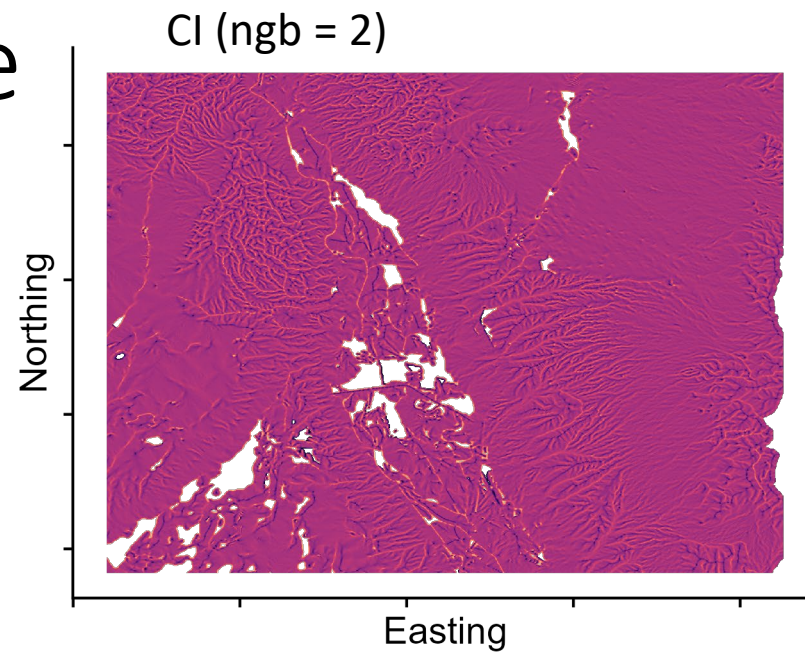


- Positive, negative, and differential openness
- Settings
 - Neighborhood sizes were 2, 32, and 256 cells
- Interpretation:
 - Units are degrees
 - Line-of-site openness; infer contextual landscape position related to convexities or concavities.
 - Observations: negative openness highlights stream channels, while positive openness highlights the surrounding stream valley.
 - Differences between positive and negative openness seem to be most noticeable at larger radii.
 - Negative openness is not the inverse of positive openness and can contain useful patterns.

Morphometry

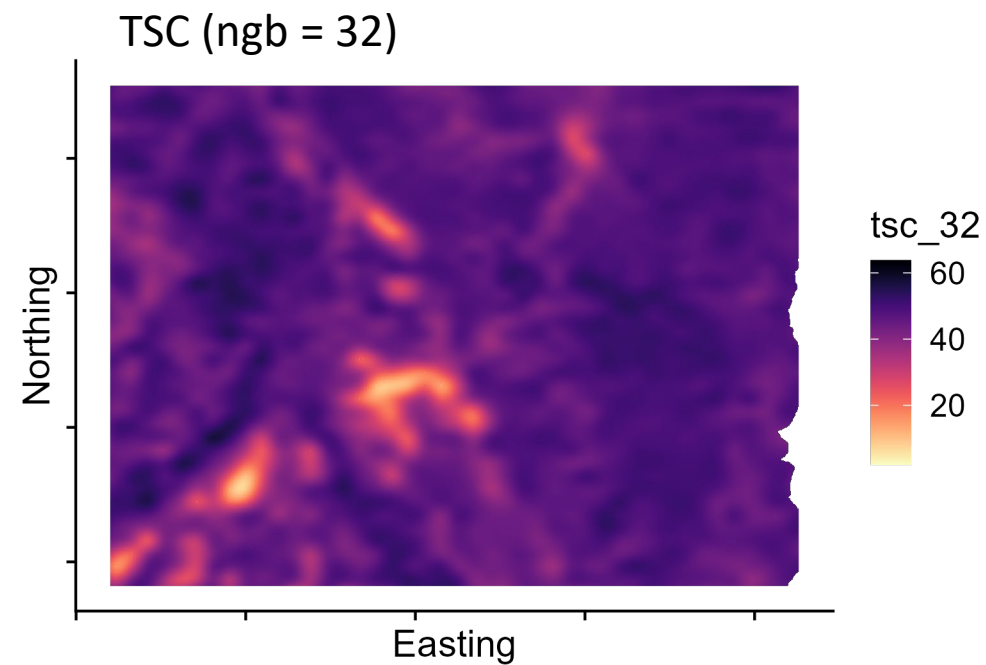
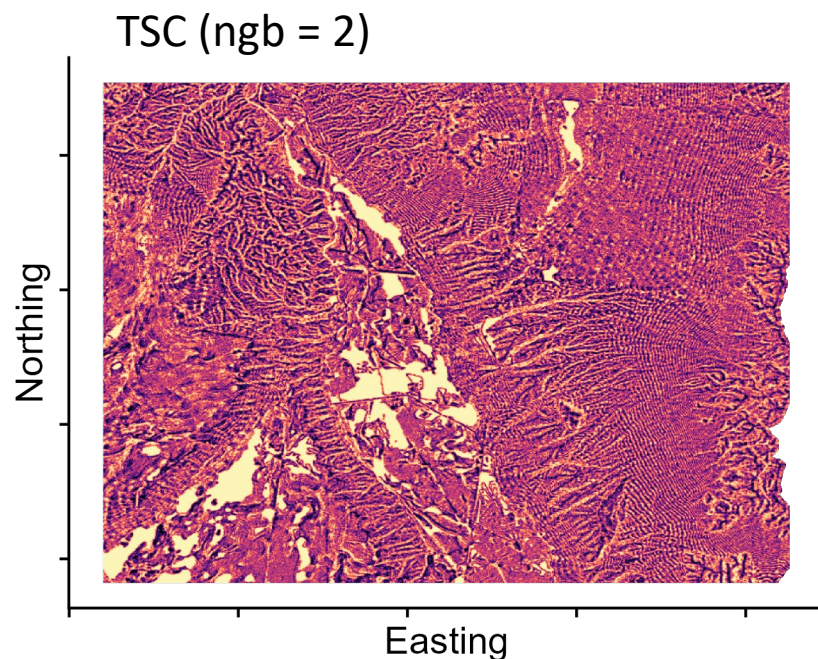
1. Convergence index
2. Terrain surface convexity
3. Multiscale topographic position index
4. Geomorphons
5. Morphometric Features
6. Terrain ruggedness index
7. Vector ruggedness index
8. Mass balance index
9. Slope, aspect, curvature
10. Focal statistics

Convergence Index



- Settings:
 - Calculated over radii of 2, 4, 8, 16, and 32 cells
- Interpretation:
 - Units = Unitless; all values scaled from -100 to 100.
 - Values > 0 are convergent, values < 0 are divergent areas.
 - A value of 100 indicates that all surrounding cells flow to this cell (i.e., a pit)
 - A value of -100 indicates that all surrounding cells flow away from this cell (i.e., a peak)
 - A value of zero indicates that all surrounding cells have the same aspect and thus water does not converge or diverge, but flows as a plane across the surface.
 - This is intended to approximate water flow across the surface and thus areas that are likely to be drier or wetter.

Terrain Surface Convexity



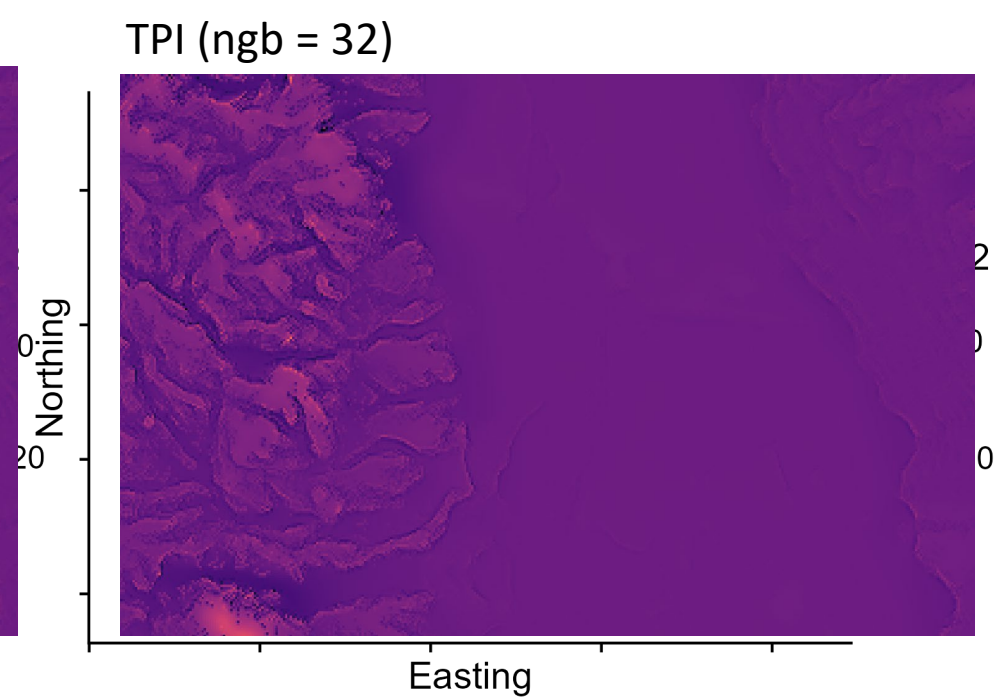
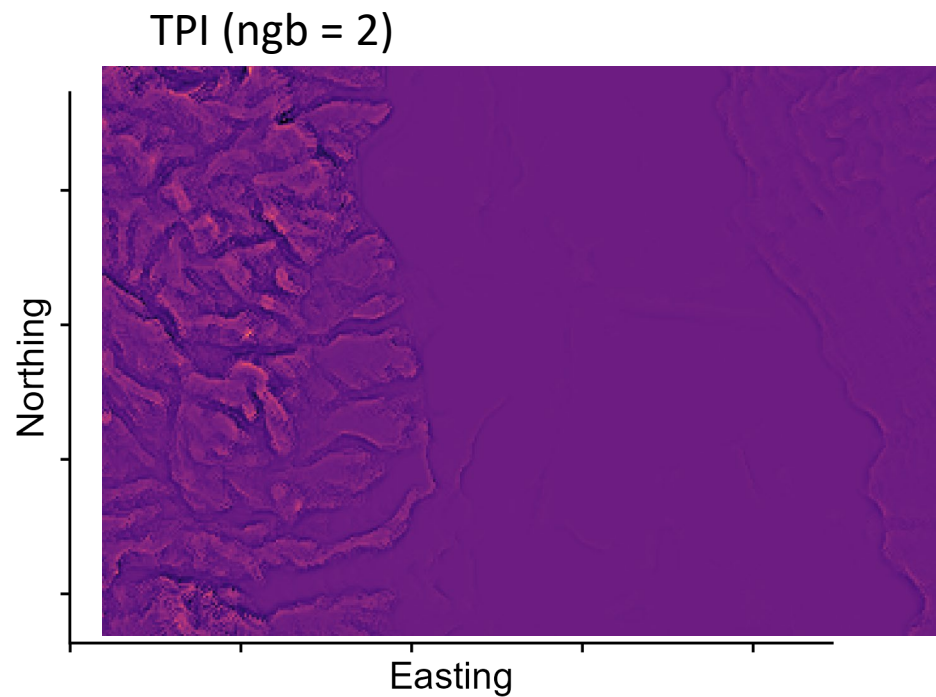
• Settings:

- Calculated over radii of 2, 4, 8, 16, and 32 cells
- A cell was considered 'flat' if the convexity was < 0.01

• Interpretation:

- Units are % (of the neighborhood that is convex)
- This is a measure of how convex the entire search radius
- Higher values will have more 'convex' areas.
- This could also be calculated as concavity, but the results would be the reciprocal of convexity

Multiscale Topographic Position Index



- **Settings:**

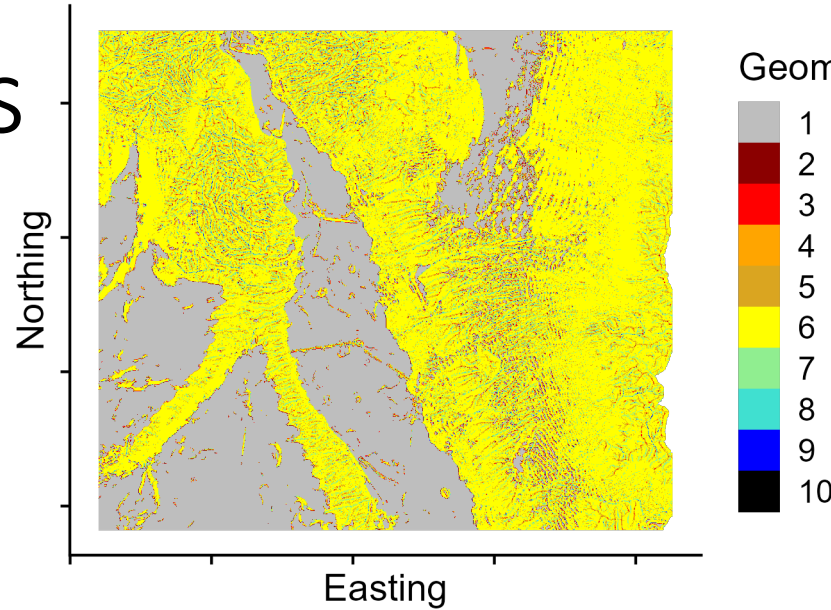
- Calculated over radii of 2, 32
- Minimum and maximum (i.e., inner and outer) distance of annulus set to 0 and x, where x is the radii.
- Number of scales set to two.

- **Interpretation:**

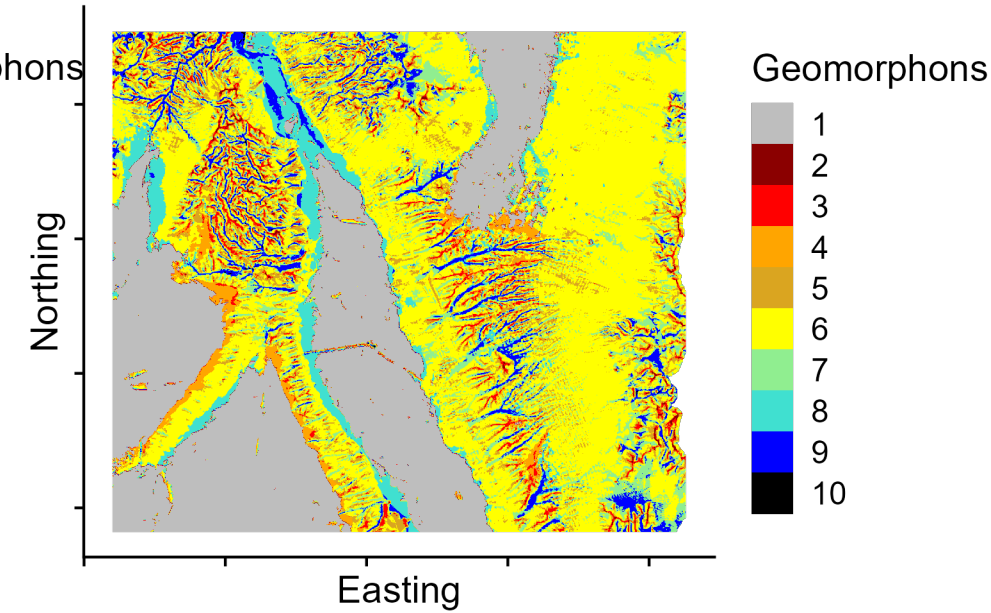
- Units are meters (difference from the mean elevation in the neighborhood)
- Positive values are locations that are higher than the average elevation of the surrounding area
- Negative values are locations that are lower than the average elevation of the surrounding area
- Values near zero are either flat areas (where the slope is near zero) or areas of constant slope (where the slope of the point is significantly greater than zero).
- This can approximate landscape position/landform/geomorphic environment

Geomorphons

Geomprh (radius = 30)

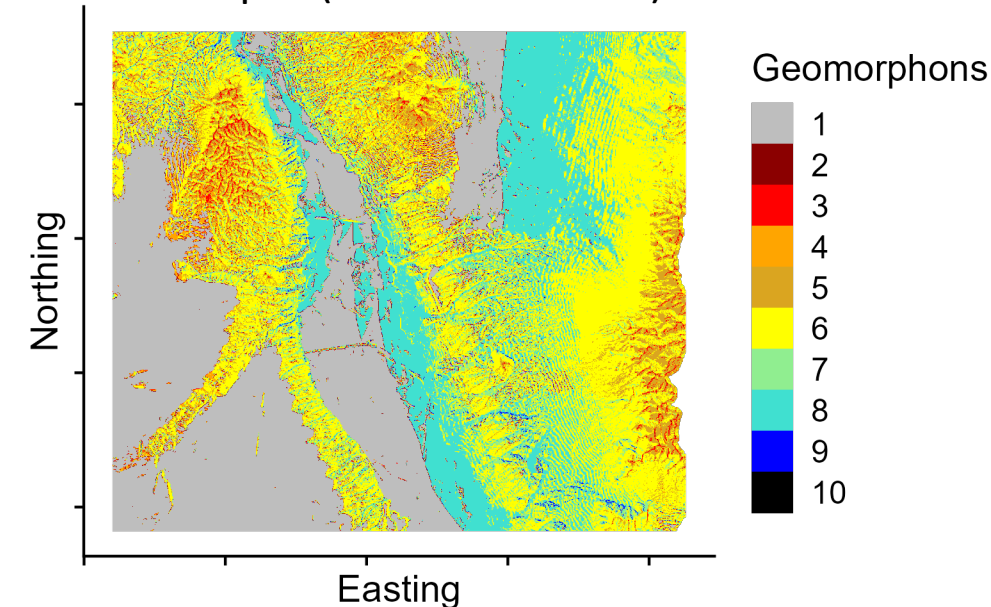


Geomprh (radius = 3000)

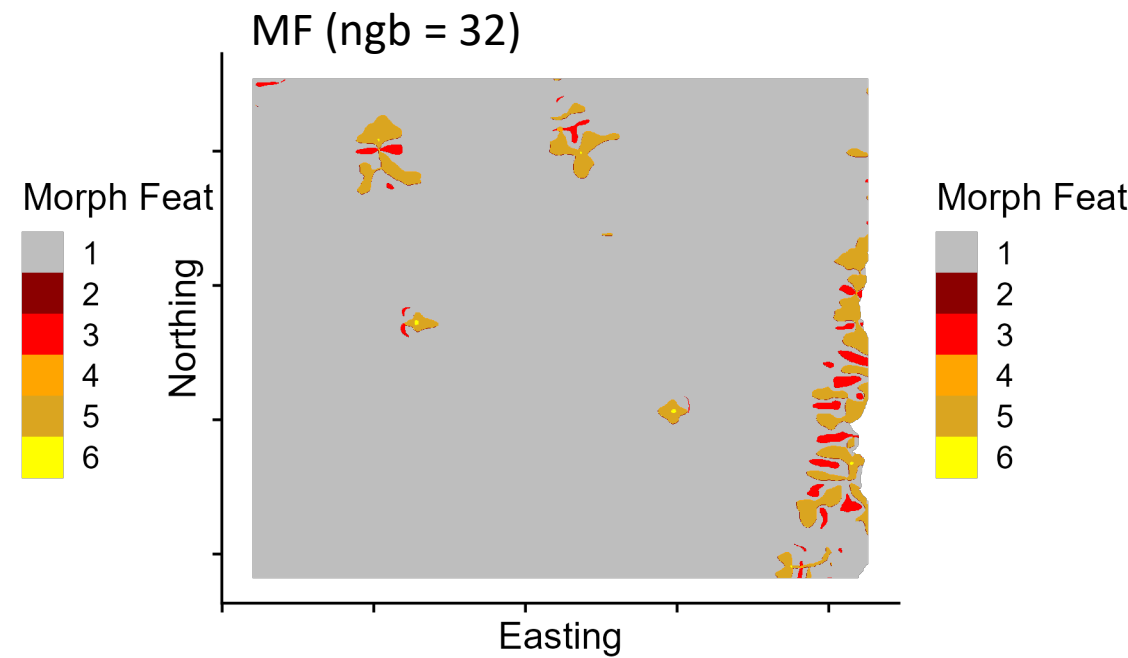
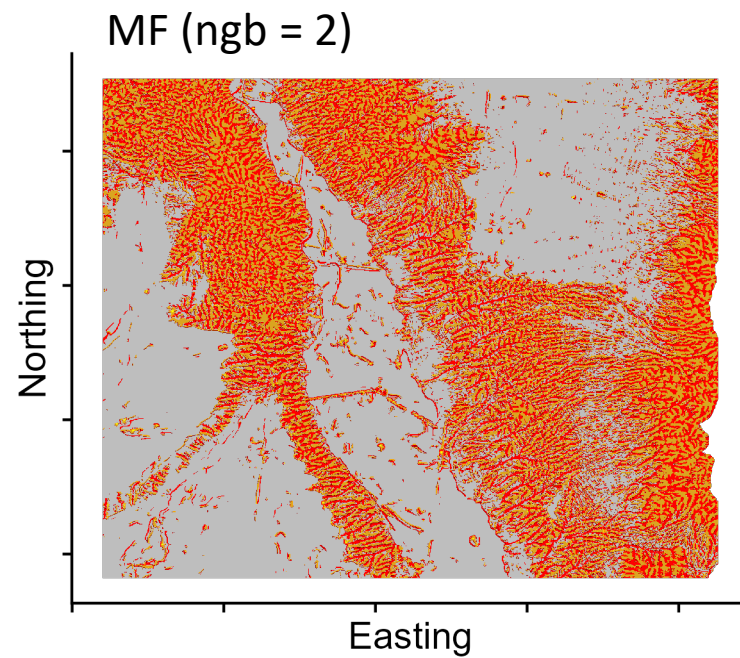


- Both 'set radius' and multiscale
- Settings:
 - Only changes > 0.5 m elevation were considered non-flat
 - Radii values = 30m, 300m, 3000m
 - Multiscale = 30m, 300m
- Interpretation:
 - Geomorphons are a landform classification:
 - 1 = flat, 2 = summit, 3 = ridge, 4 = shoulder, 5 = spur, 6 = slope, 7 = hollow, 8 = footslope, 9 = valley, 10 = depression, 0 = nodata

Geomprh (multiscale = 300)

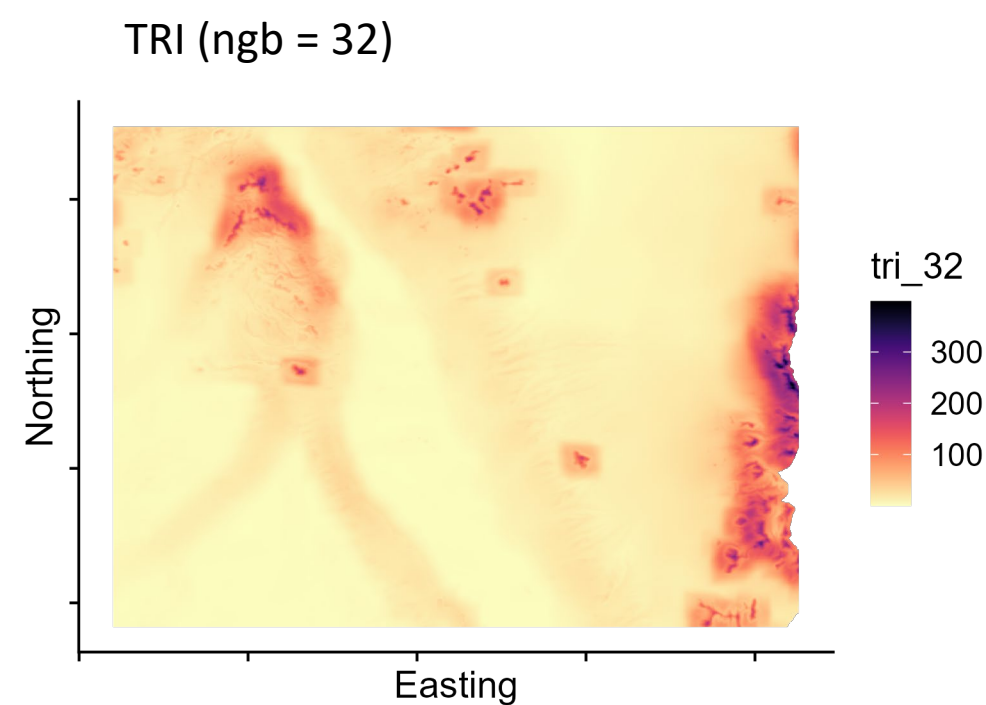
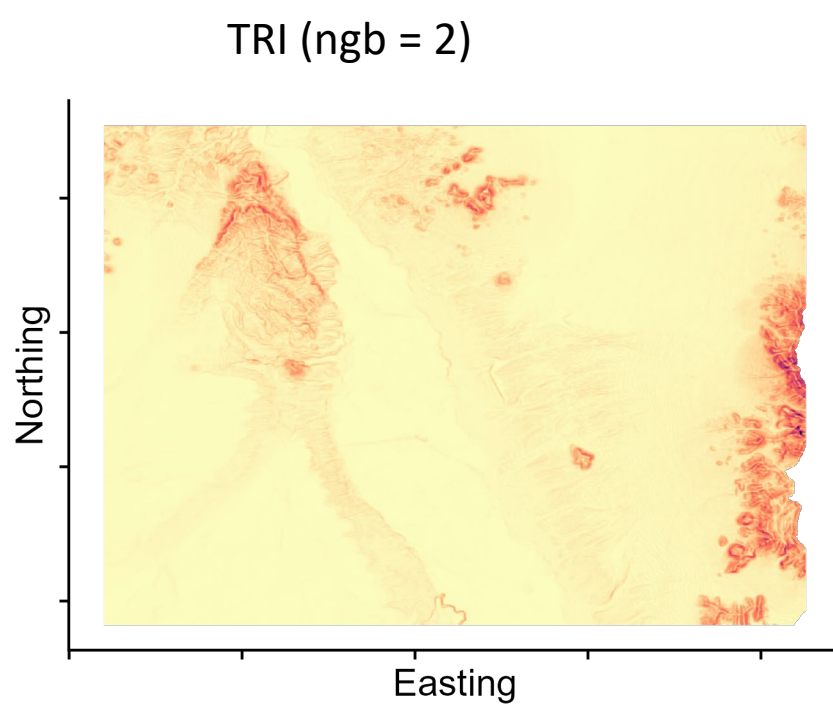


Morphometric Features (terrain classification)



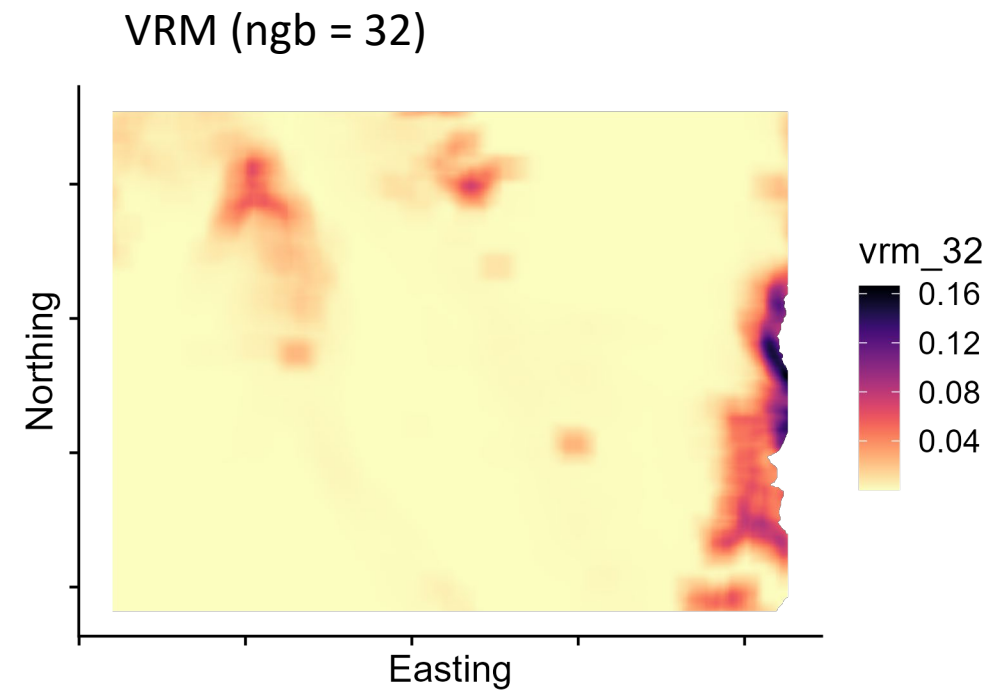
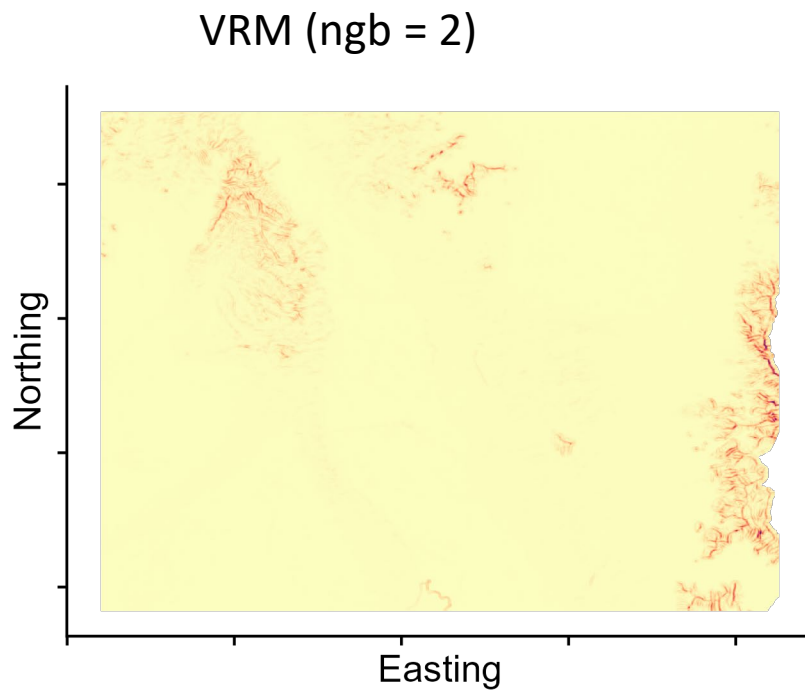
- Settings:
 - Radius set to: 2, 4, 8, 16, and 32
 - Slope tolerance set to 1 (slopes less than this are considered flat)
 - Curvature tolerance set to 0.0001 (curvatures less than this are considered flat)
- Interpretation:
 - 1 = plane, 2 = pit, 3 = channel, 4 = pass(saddle), 5 = ridge, 6 = peak

Terrain Ruggedness Index



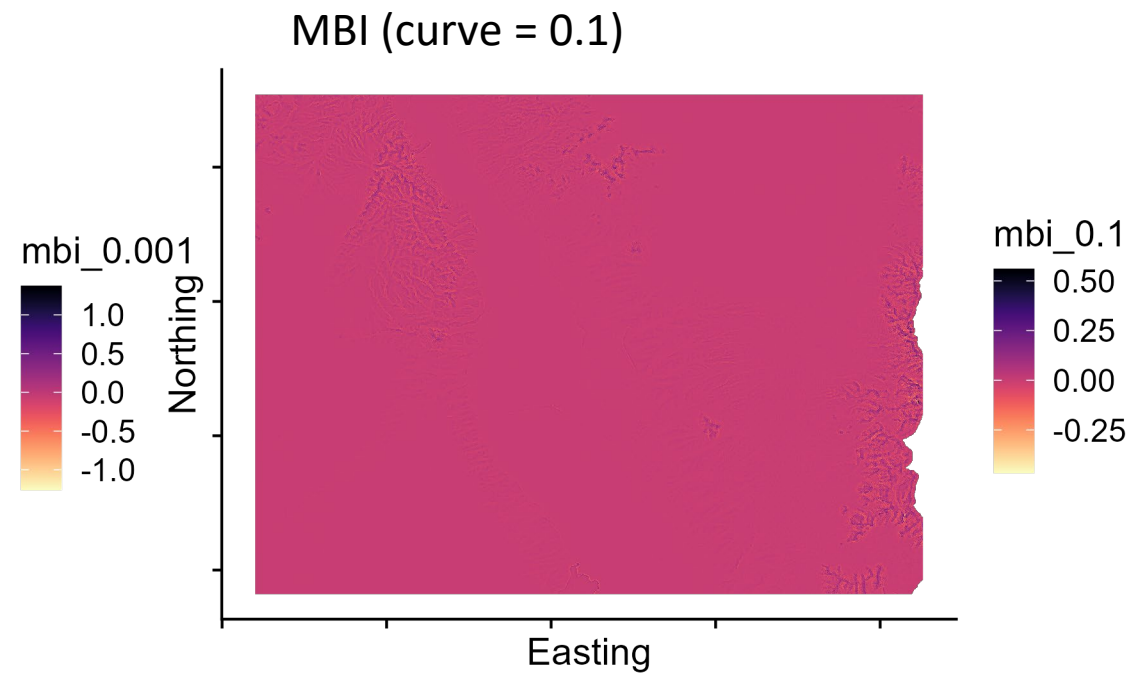
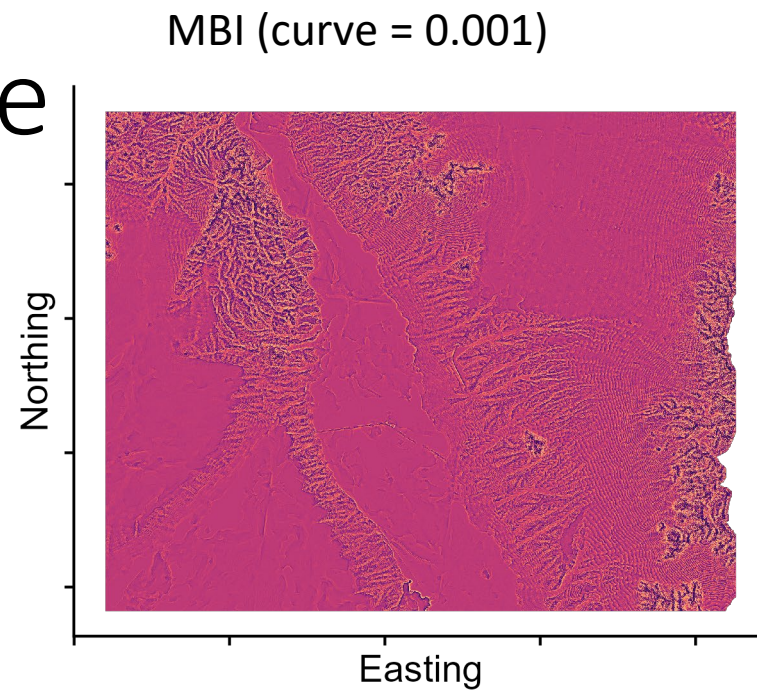
- Settings:
 - Radius in cells: 2, 4, 8, 16, and 32
 - Square neighborhood
- Interpretation:
 - Units are in meters
 - The average elevation change between any point on a grid and it's surrounding area
 - Possibly highlight areas of active erosion: edges of cuestas, badlands, or other breaks

Vector Ruggedness Measure



- Settings:
 - Radius in cells: 2, 4, 8, 16, and 32
 - Square neighborhood
- Interpretation:
 - Unitless; ranges from 0 (flat) to 1 (most rugged)
 - Values are low in both flat and steep areas, but are high in areas that are both steep and rugged.
 - May be useful for identifying areas where active erosion is occurring.
 - Could also be related to geology with differential erosion and thus highlight parent material differences

Mass Balance Index

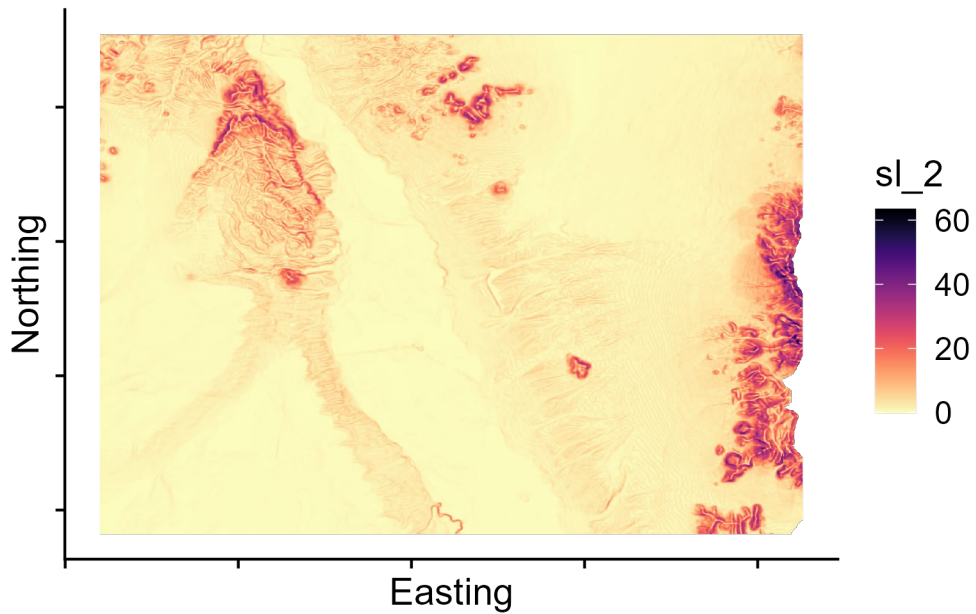


- Settings:
 - Curvature thresholds used: 0.001, 0.01, 0.1.
- Interpretation:
 - Unitless
 - Negative values = areas of net deposition such as depressions and floodplains
 - Positive values = areas of net erosion such as hillslopes
 - values close to zero = areas of balance between erosion and deposition such as low slopes and plain areas
 - Intended to approximate geomorphological processes of sediment redistribution

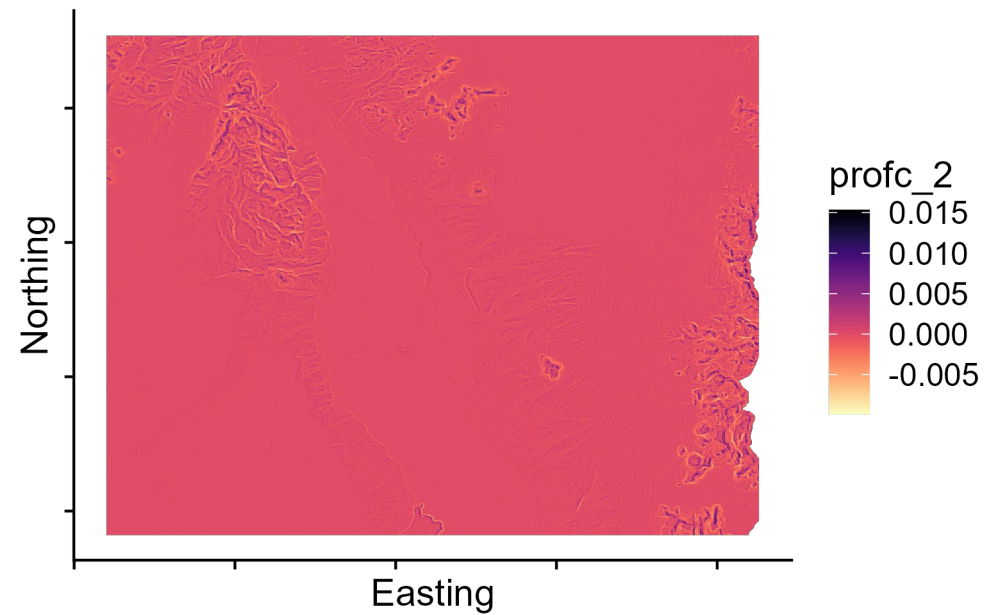
Slope, Aspect, Curvature

- Slope; aspect;
- Curvatures: profile, plan, longitudinal, cross-sectional, maximum, minimum
- Settings:
 - Radius set to 2, 4, 8, 16, and 32
- Interpretation:
 - Slope units are degrees
 - Aspect units are degrees
 - Curvature units are in m^{-1}
 - Positive curvature values describe convexity
 - Negative curvature values describe concavity
 - Slope and curvature approximate geomorphological processes related to soil redistribution by water and gravity

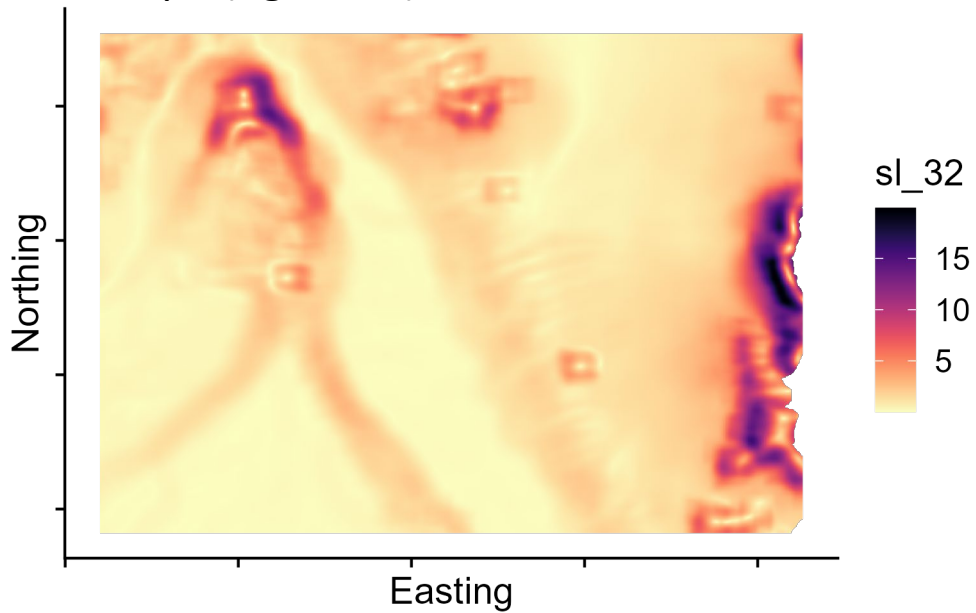
Slope (ngb = 2)



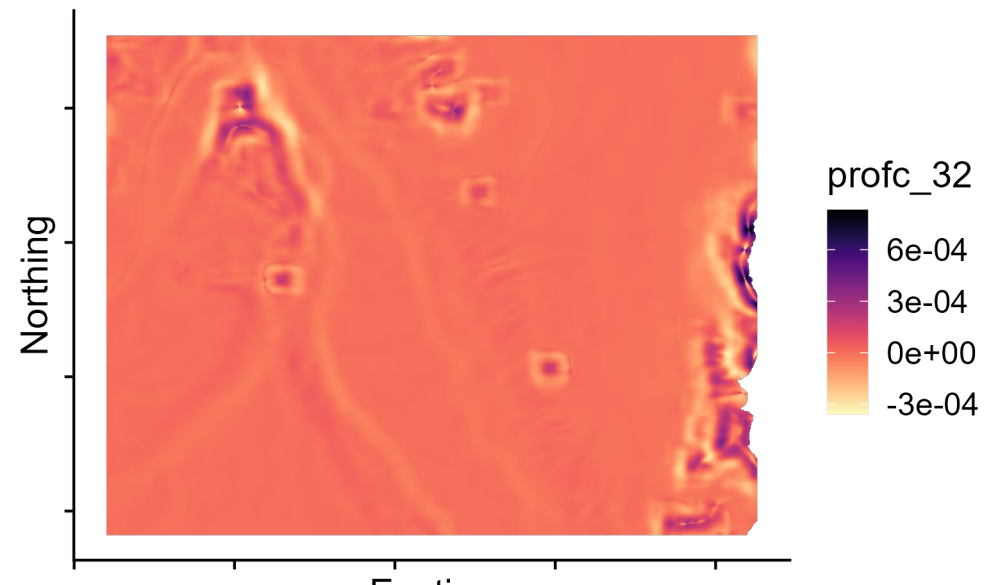
Profile Curv (ngb = 2)



Slope (ngb = 32)



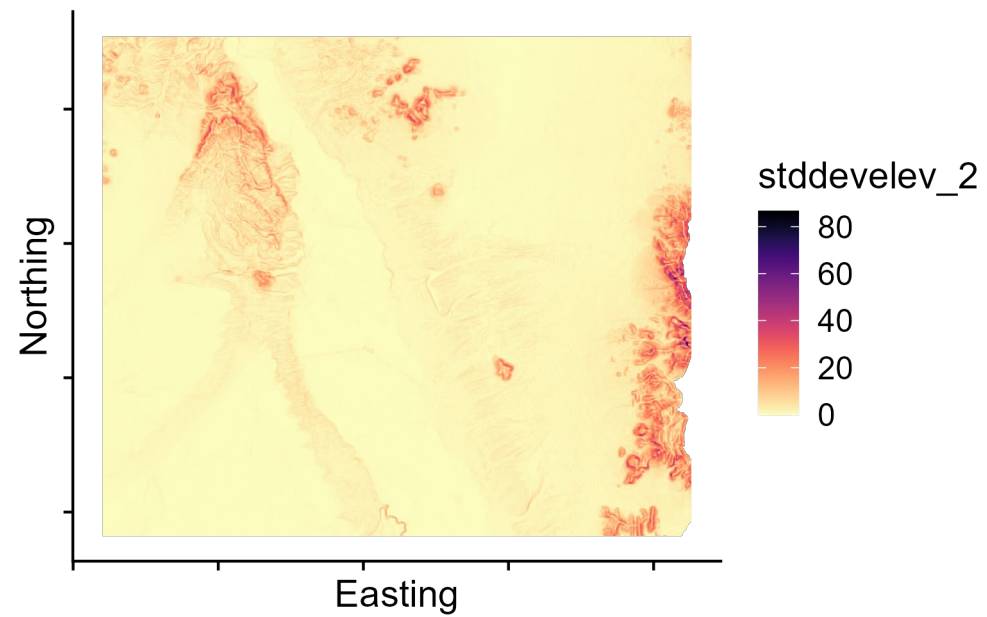
Profile Curv (ngb = 32)



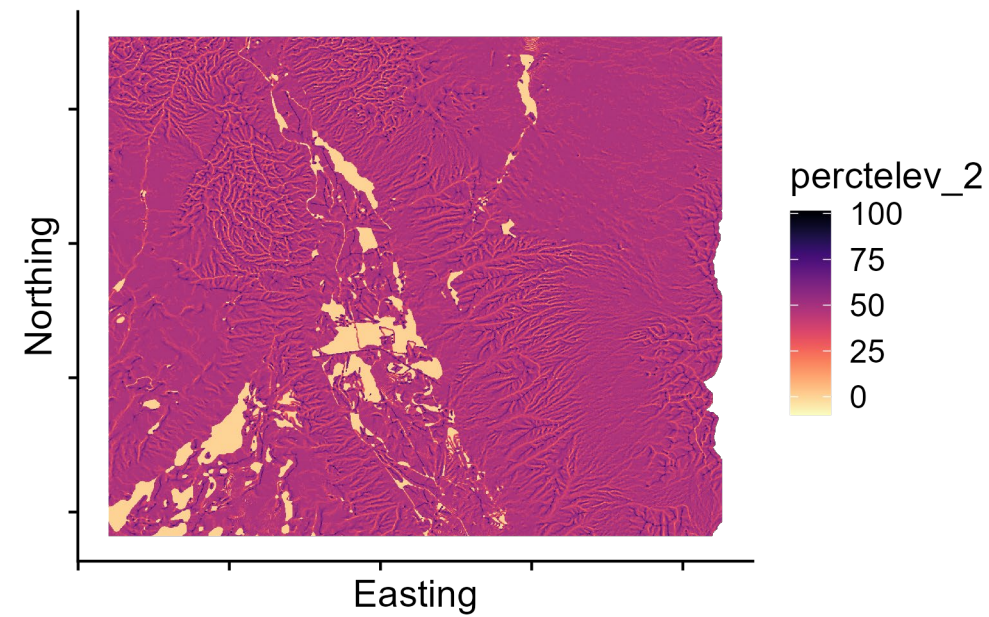
Focal Statistics

- Mean, minimum, difference from mean elevation, deviance from mean, relative elevation, relative mean elevation, standard deviation, percentile
- Settings:
 - Calculated over a radius = 2, 4, 8, 16, and 32 cells
- Interpretation:
 - Units are meters (except percentile)
 - Percentile (a value of 90 means that the cell is higher than 90% of all other cells in the neighborhood)
 - These are intended to contextualize landscape position which can often provide information about water accumulation/distribution or geology/landform.

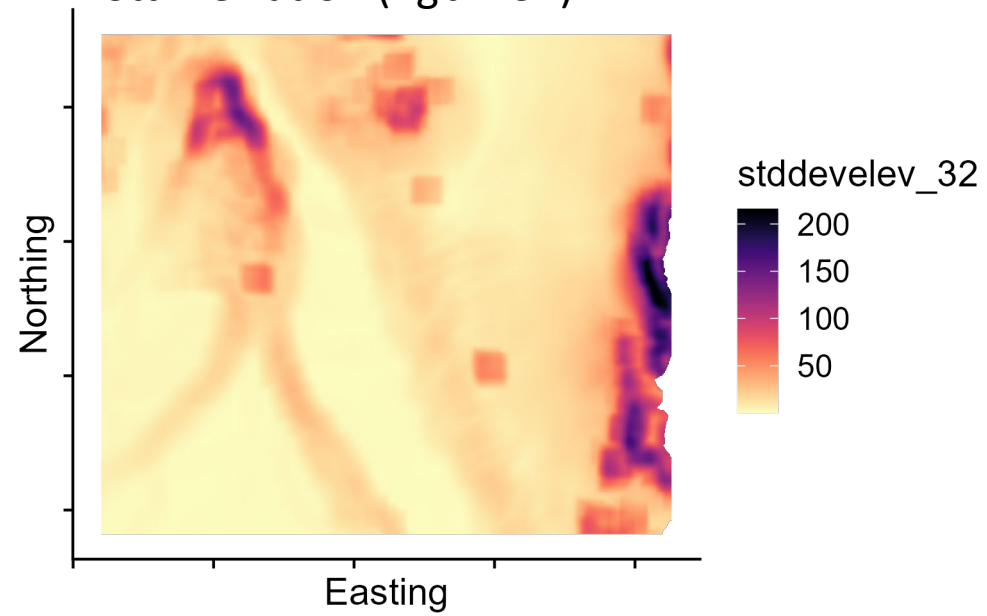
Std Deviation (ngb = 2)



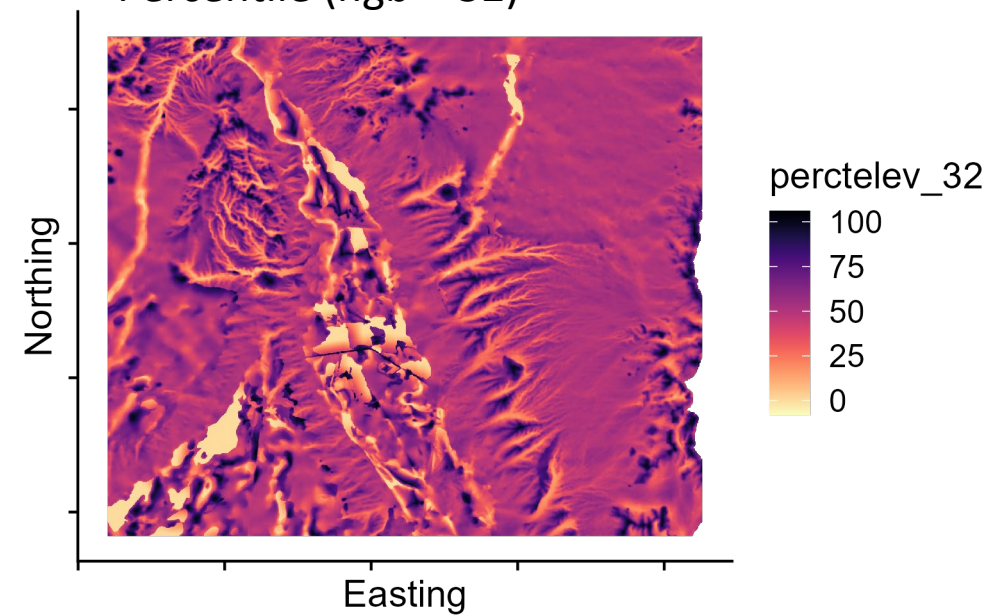
Percentile (ngb = 2)



Std Deviation (ngb = 32)



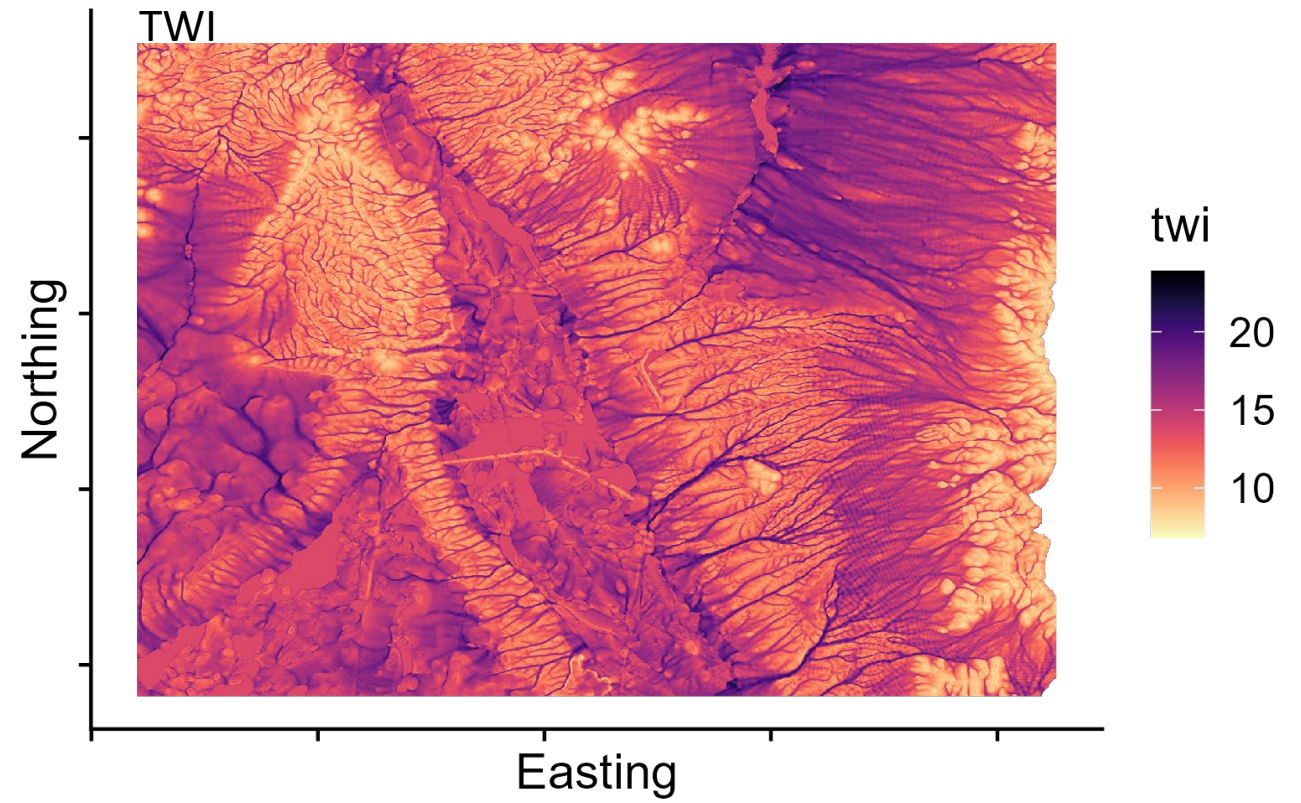
Percentile (ngb = 32)



Hydrology

1. Topographic wetness index
2. Saga wetness index
3. Stream power index
4. Vertical distance to channel network
5. Valley depth

Topographic wetness index

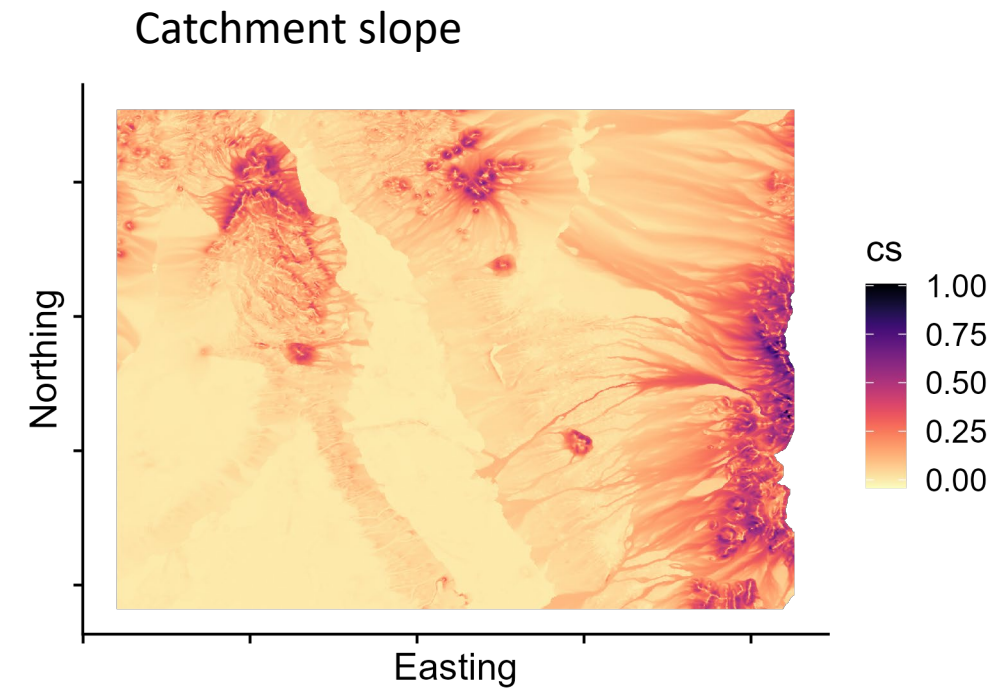
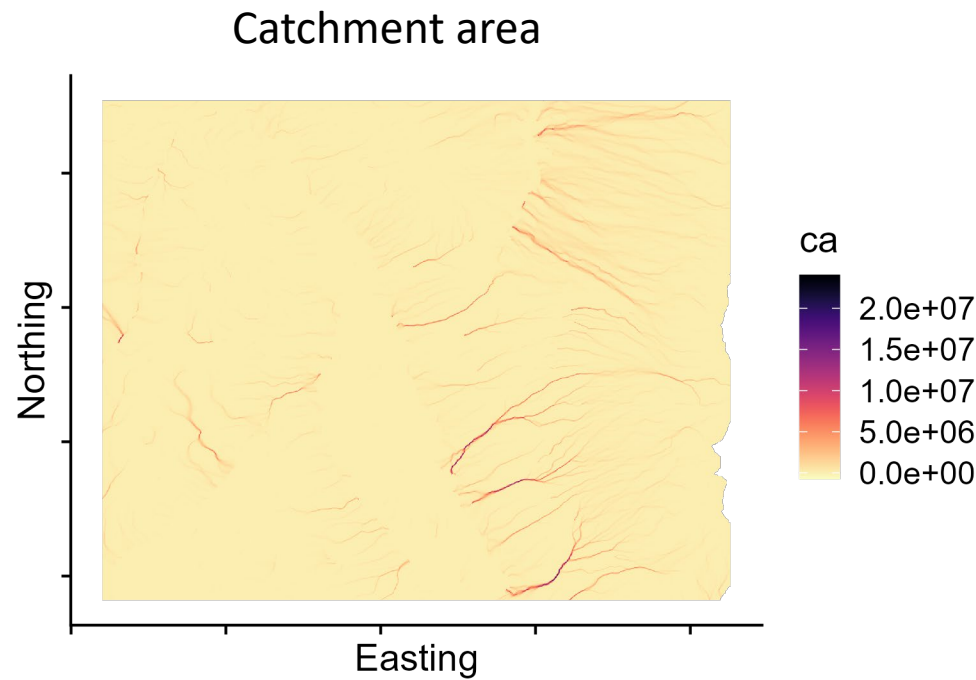


- Settings:
 - Input specific catchment area and slope
 - Specific catchment area calculated using top-Down flow accumulation
 - Slope calculated with a 9 parameter 2nd order polynomial (Zevenbergen and Thorne, 1987).
- Interpretation:
 - Unitless
 - The tendency of a cell to accumulate water based on the idea of mass balance (water in vs water out).
 - Cells with lower slope (so water will not run off as fast) and a larger contributing area (larger area = more area to 'catch' water) should have a higher potential wetness
 - Compared to grid cells with either a larger slope (so water will run off faster) or smaller contributing area (less potentially available water).
 - Higher values = potentially wetter areas

Saga wetness index

- Catchment area, Catchment slope, Modified catchment area, Saga wetness index
- Settings:
 - Suction = 10 and 10,000 (only affects modified catchment area)
 - Never really explained
 - Smaller values means higher suction and result in a higher wetness index values.
 - Area type = specific catchment area
 - Slope type = local slope
 - Minimum slope (all values smaller than this value will be set to this value) = 0.
 - Slope offset = 0.01.

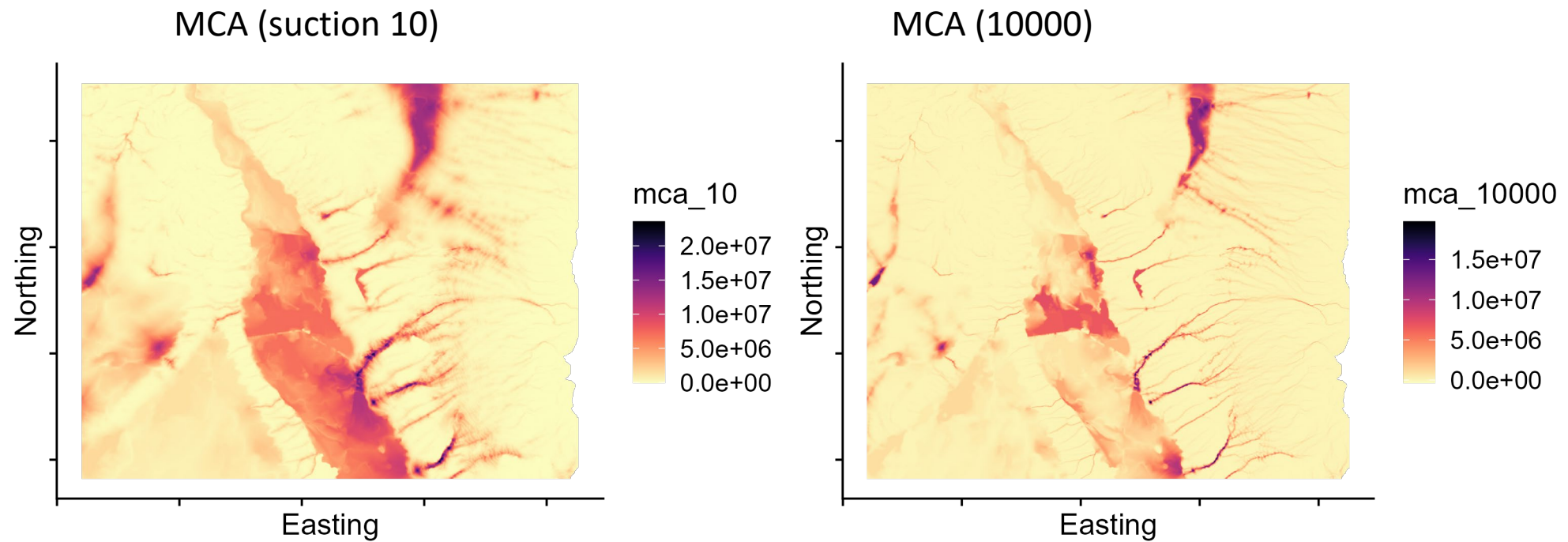
Catchment area and slope



- Interpretation:

- Catchment area (aka upslope area).
 - Units are m^2
- Catchment slope - the mean slope of all the cells upslope from a cell
 - Unitless (?)
 - Indicates the potential energy as cells with steeper contributing areas are likely to have higher velocity flows

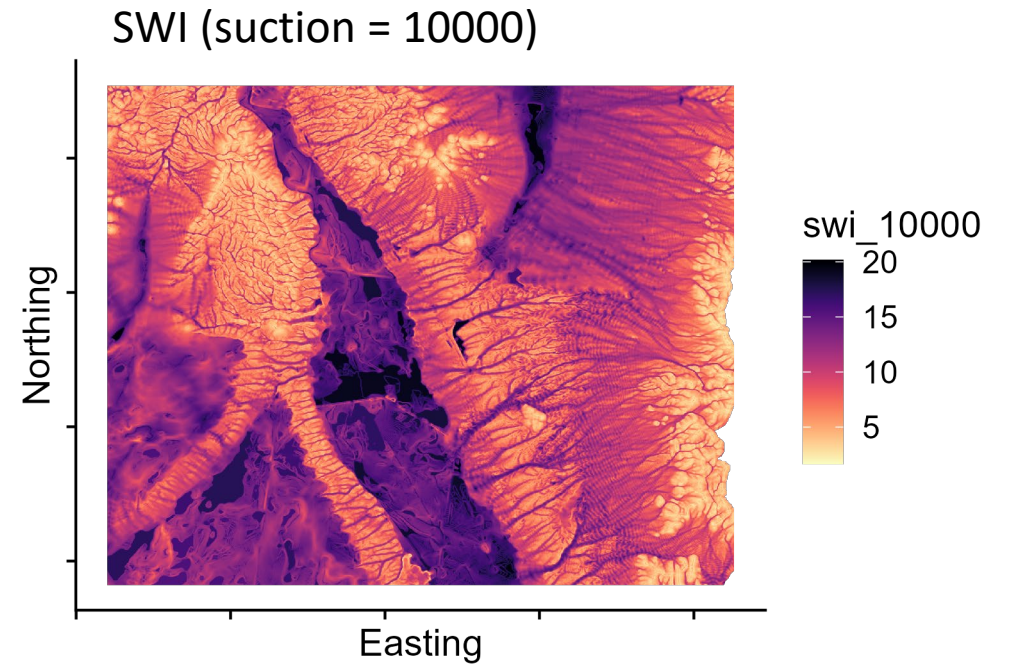
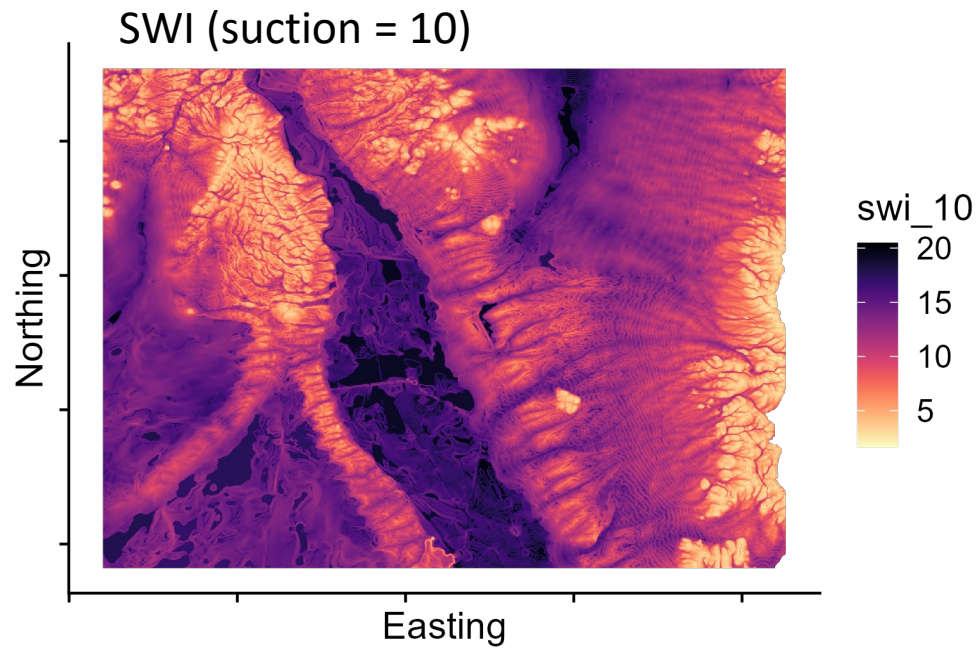
Modified Catchment Area



- Interpretation:

- Units are m^2
- Imagine the situation where a single cell is slightly higher (a few cm) than all surrounding cells.
- Most likely from imprecision in the DEM rather than actual elevation differences
- However since this is the highest grid cell (like a peak) then it's contributing area is only itself which would result in random patterns related to the DEM noise
- So the already calculated specific catchment area is iteratively modified until results remain unchanged

Saga Wetness Index



- Interpretation:
 - Unitless
 - Same as TWI, higher values = potentially wetter areas

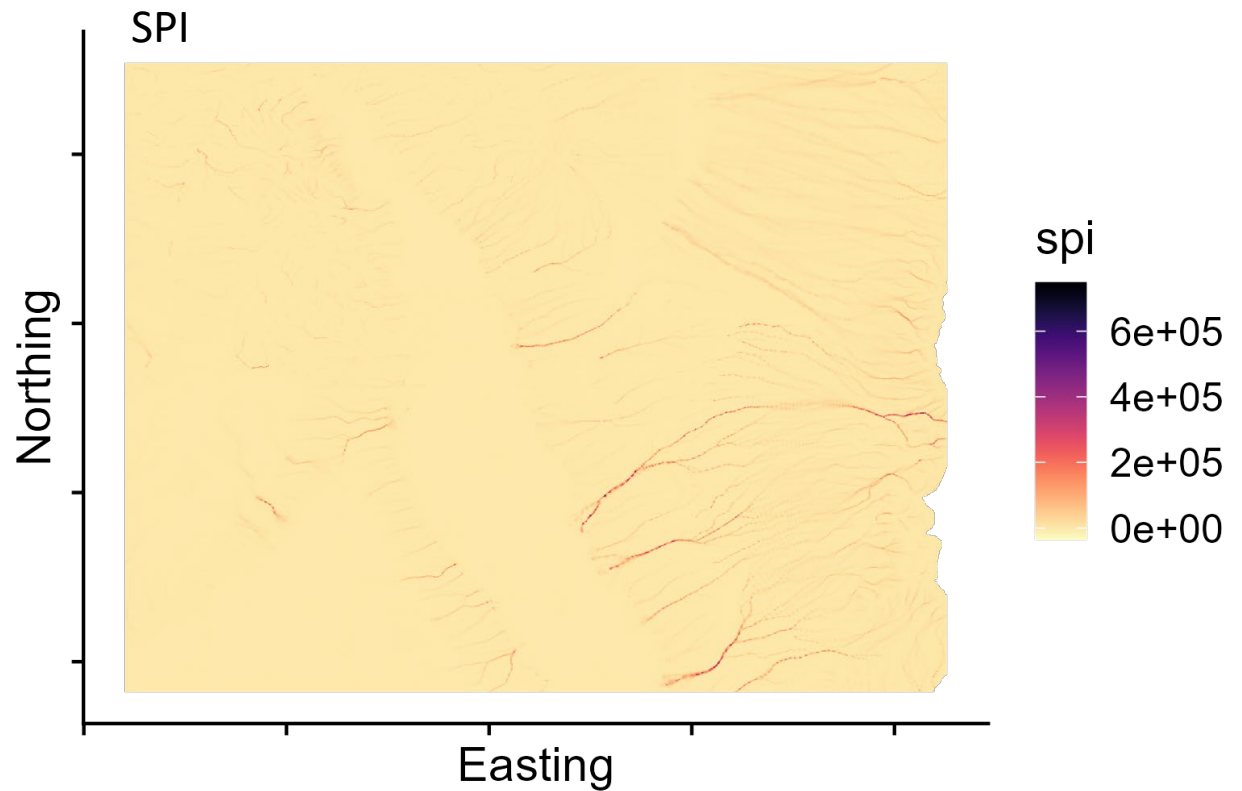
Stream Power Index

- Settings:

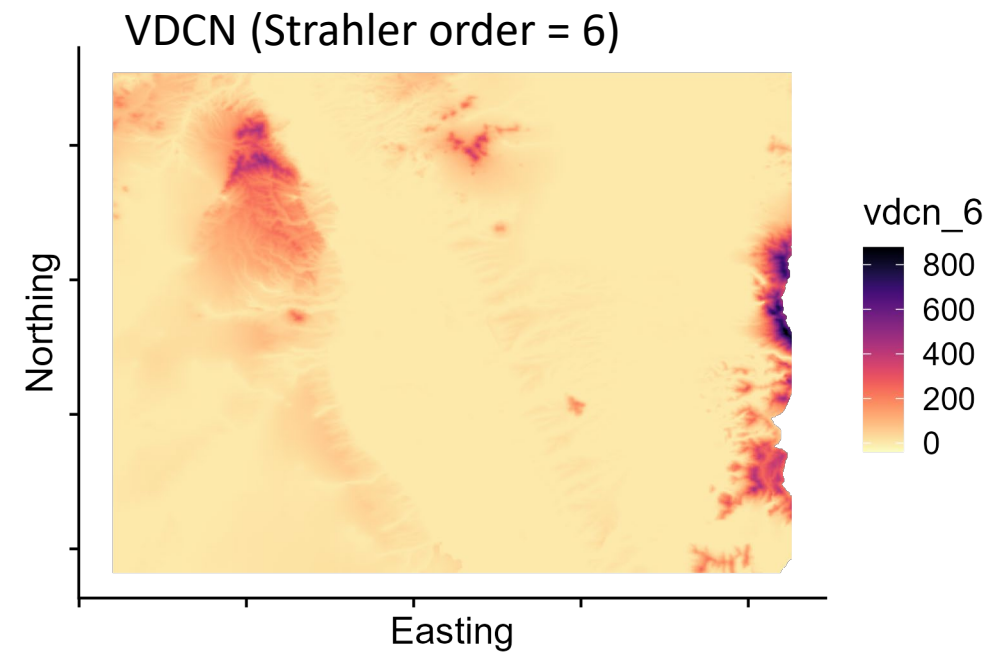
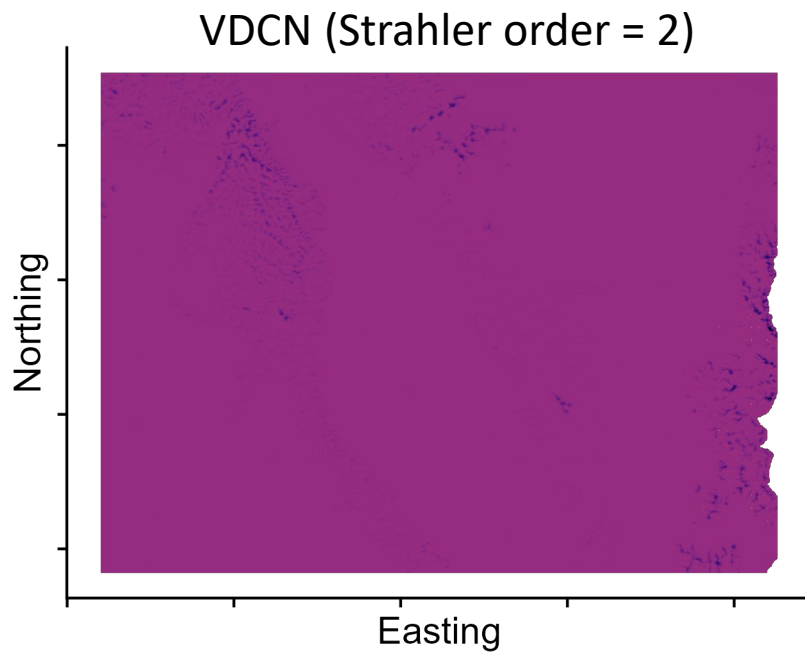
- Input specific catchment area and slope
 - Specific catchment area calculated using the Top-Down flow accumulation
 - Slope calculated with a 9 parameter 2nd order polynomial (Zevenbergen and Thorne, 1987).

- Interpretation:

- Unitless
- Approximates potential soil erosion from water flow.
- As the specific catchment area and slope increase the amount of water from upslope areas as well as the velocity of water flow increase.
- Higher values indicate that erosion is more likely

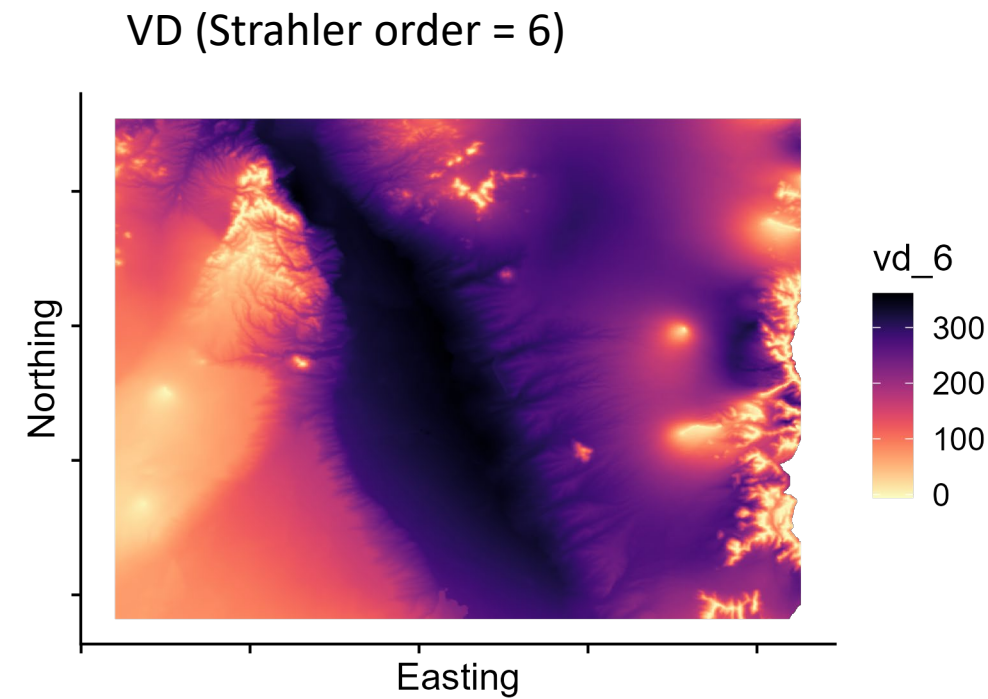
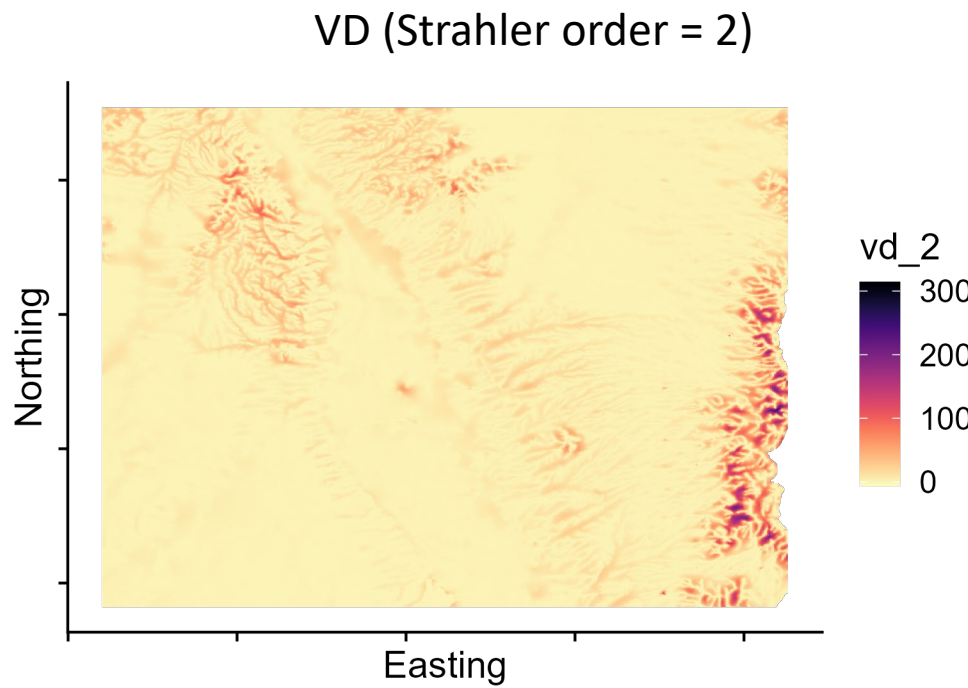


Vertical Distance to Channel Network



- Vertical distance to channel network (base level)
- Settings:
 - Maximum change = 100m
 - Calculated for Strahler order 'streams' = 2, 3, 4, 5, 6
- Interpretation:
 - Base level = theoretical elevation surface between stream lines
 - Units are meters
 - Vertical distance to channel network = difference between actual elevation and base level
 - Units are in meters above the interpolated base level
 - This could capture surface age in a stream or fan-terrace environment or estimate relative erosion (differences in geology/parent material)
 - Has been useful for mapping depth to groundwater and maybe used for mapping wet soils or depth to redox

Valley Depth



- Valley depth (ridge height)
- Settings:
 - Maximum change in meters = 100
 - Calculated for Strahler order 'streams' = 2, 3, 4, 5, 6
- Interpretation (inverse of base level and vertical distance to channel network):
 - Ridge height
 - units are meters.
 - Theoretical elevation surface between ridge lines
 - Valley depth is the difference between an interpolated ridge level and the actual elevation.
 - Units are meters below the interpolated ridge level.
 - Conceptually useful for relative position within a valley, identifying landscapes in a terrace environment, or location along a hillslope

Data availability

- Data provided by HUC6 watershed and CONUS mosaic
 - Complete HUC 6 dataset is ~ 6 TB
 - Conus mosaic is ~ 6 TB
 - Individual HUC 6 directories are < 94.3 GB (most are MUCH smaller ~ 5 – 10 GB)
- FTP @ NMSU (set up, currently working with IT to make it public facing)

Practical Applications

- Resolution (30 m)
 - You may think that this is too coarse for your landscape
 - Encourage you to try this before you invest effort in calculating these for higher spatial resolution data.
 - Base DEM is often from LiDAR
 - Multi-scale + different parameters
 - Can improve accuracy
 - Could help you identify relevant variables

Thank you!



Loamy Skeletal Typic
Haplocalcid
Dona Ana County, New Mexico

Questions to clarify in publication

- How did I decide on what parameters
 - Iterated over each covariate
 - Generated for a small NM HUC12, chose those parameters that produced the largest difference and or/least correlation between outputs. Assumption that these parameters would be useful on a CONUS scale. I think that this is a relatively robust assumption because the neighborhood size (for those with specific radii) was always smaller than the HUC12 (and will be << smaller than the HUC6 units) and because most soil properties are probably driven by processes occurring on the landscape/hillslope scale which will almost always be smaller than the HUC12. Need to explain this more.
- Actual computational steps + processing specs (cpu, ram, etc.)
- How to deal with multi-collinearity when modeling? (this is really a question about how to choose variables)
 - Iterative process, probably combined
- Each variable needs to be discussed more and explained to a soil scientist interested in using it (I wonder if I could farm this out to specific individuals so they could add a paragraph)
 - Underlying equation
 - Underlying assumptions/backgrounds
 - How it has been used