stagg:: spatiotemporal aggregation of climate data in R

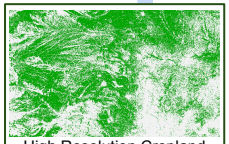
Cheat Sheet

The R Package stagg enables simple and efficient pairing of climate and economic or political data for use in nonlinear regression analyses. This is accomplished by aggregating gridded ERA5 data to the level of administrative regions in a 3 step process.

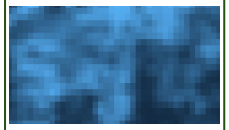
1. secondary_weights()

Resample a raster layer to a different spatial resolution

Argument	Description	Format
secondary_raster	Data on a seperate variable to weight climate data by during aggregation	Raster layer, raster brick, or raster stack
grid	Grid with the same resolution as climate data to resample the secondary raster to, defaults to the ERA5 grid	Raster layer, raster brick, or raster stack
extent	Longitude and latitude boundaries to crop the secondary raster for greater efficiency, defaults to reading in entire raster	Numeric vector of length 4, in the order c(xmin, xmax, ymin, ymax)



High Resolution Cropland



ERA5 Resolution Cropland

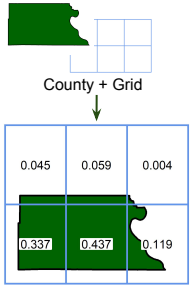
x	y	weight
-98.75	37.25	2.127619e-01
-98.50	37.25	4.731410e-01
-98.25	37.25	5.053359e-01

Example Usage

```
secondary_weights_kansas <- secondary_weights(secondary_raster =
  cropland_cropland_world_2011, grid = era5_grid, extent =
  c(-103, -94, 37, 41))
```

2. overlay_weights()

Calculate the portion of each polygon that falls in each grid cell



(Optional)

Argument	Description	Format
polygons	Borders of administrative regions	Simple features object
polygon_id_col	The name of the column with unique identifiers for each polygon	String
grid	A grid with the same resolution as climate data to overlay the polygons onto, defaults to the ERA5 grid	Raster layer, raster brick, or raster stack
secondary_weights	Optional table of weights determined by a separate variable to be normalized by area, created using previous function	Data.table with 3 columns: 'x', 'y', and 'weight'

Example Usage

```
overlay_weights_Kansas <- overlay_weights(polygons =
  kansas_counties, polygon_id_col = "GEOID", grid =
  era5_grid, secondary_weights =
  secondary_weights_kansas)
```

x	y	poly_id	w_area	weight
258.25	37.75	075	0.1303709778	0.1390821970
258.25	37.75	187	0.1558011301	0.1459069310
258.25	38.00	075	0.2358176588	0.2181691016

3. staggregate_*

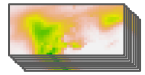
Transform climate data and aggregate gridded values to the polygon level

staggregate_* is a family of functions which take mostly the same arguments and perform the same role. The difference between each is the transformation performed, and arguments specific to that transformation.

Argument	Description	Format
Common to all staggregate_* functions		
data	Climate data to aggregate	A raster brick or raster stack containing a multiple of 24 layers
overlay_weights	Table of area weights (and possibly area-normalized secondary weights) to use in aggregating to the polygon level, created using previous function	Data.table with 4 or 5 columns: 'x', 'y', 'poly_id', 'w_area', and, if desired, 'weight'
daily_agg	How to convert hourly values into daily values	One of two strings: "sum", or "average"
time_agg	The temporal scale to aggregate transformed values to	One of three strings: "year", "month", or "day"
Unique to staggregate_polynomial() - [Polynomial Transformation]		
degree	The highest order to raise the daily values to	Whole number greater than 0
Unique to staggregate_spline() - [Restricted Cubic Spline Transformation]		
knot_locs	Knot locations	Numeric vector
Unique to staggregate_bin() - [Binning Transformation]		
num_bins	Number of non-edge bins to draw, defaults to 10	Whole number greater than 0
binwidth	Width of non-edge bins, overrides num_bins, defaults to minimum in data	Positive number
min	Minimum value that non-edge bins must capture	Number
max	Maximum value that non-edge bins must capture	Number
start_on	Where to draw the left edge of a bin. Only one placement (start_on, center_on, and end_on) may be specified. If none of these are specified, start_on is set to min	Number
center_on	Where to center a bin on. Only one placement may be specified	Number
end_on	Where to draw the right edge of a bin. Only one placement may be specified	Number

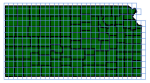
staggregate_* process

- Climate values are converted from hourly to daily
- Daily values are transformed
- Transformed values are aggregated to polygon level
- **Output ready for use in regression analysis**

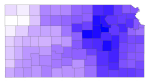


Hourly, Gridded Precipitation

+



Area Weights



Precipitation at Monthly, County Level

Example Usage

```
polynomial_output <- staggregate_polynomial(
  data = prcp_kansas_dec2011_era5,
  overlay_weights = overlay_weights_kansas,
  daily_agg = "sum", degree = 3)
```

year	month	poly_id	order_1	order_2	order_3
2011	12	181	0.03156041	0.0003409671	2.237109e-06
2011	12	023	0.02187841	0.0001325466	1.037352e-06

```
spline_output <- staggregate_spline(data =
  prcp_kansas_dec2011_era5, overlay_weights =
  overlay_weights_kansas, daily_agg = "sum",
  knot_locs = c(-1.7e-16, 1.1e-6, 1.6e-2))
```

year	month	poly_id	value	term_1
2011	12	071	0.04700092	1.297768e-09
2011	12	199	0.04279329	1.067354e-09

```
bin_output <- staggregate_bin(data =
  prcp_kansas_dec2011_era5,
  overlay_weights =
  overlay_weights_kansas,
  daily_agg = "sum", binwidth = .02,
  min = 0, max = .03)
```

year	month	poly_id	bin_min_to_0	bin_0_to_0.02	bin_0.02_to_0.04	bin_0.04_to_inf
2011	12	187	17.50392	12.40965	0.08112779	0
2011	12	071	17.81776	12.18224	0.00000000	0

daily_quants()

Values are converted from hourly to daily before transformation, so daily_quants() calculates desired quantiles of daily values for guidance in placing bins or knots. The function takes data, overlay_weights, and daily_agg, as well as a vector of quantiles to calculate (0 to 1). Note that while stagg can be used in parallel, climate data statistics like daily_quants() or default min and max values will not work as intended if data is broken into chunks for parallel processing.

Install stagg by running `devtools::install_github("tcarleton/stagg")` in R.

For further information on how to use stagg, please view the readme and documentation available at https://github.com/tcarleton/stagg