NASA/NCEAS/iPlant Update

Adam Wilson

June 22, 2012

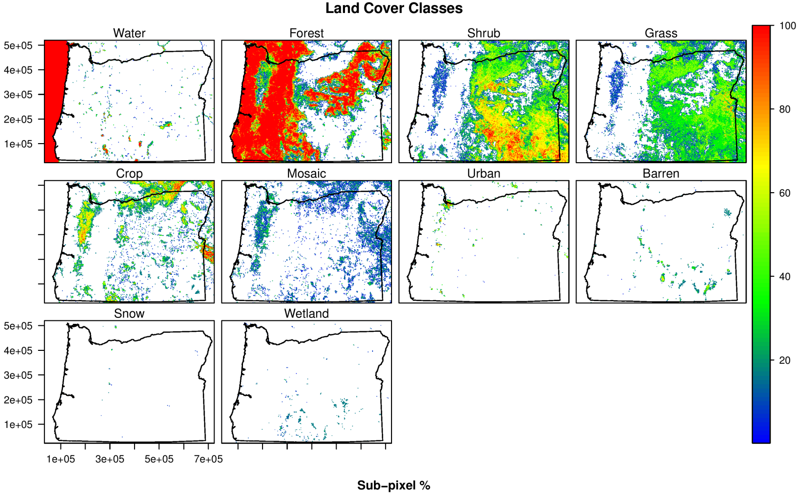
Exploring LST-LULC interactions

We know LST is affected by Land Cover (e.g. Teuling, 2010). Concerns for incorporation of LST into interpolation models:

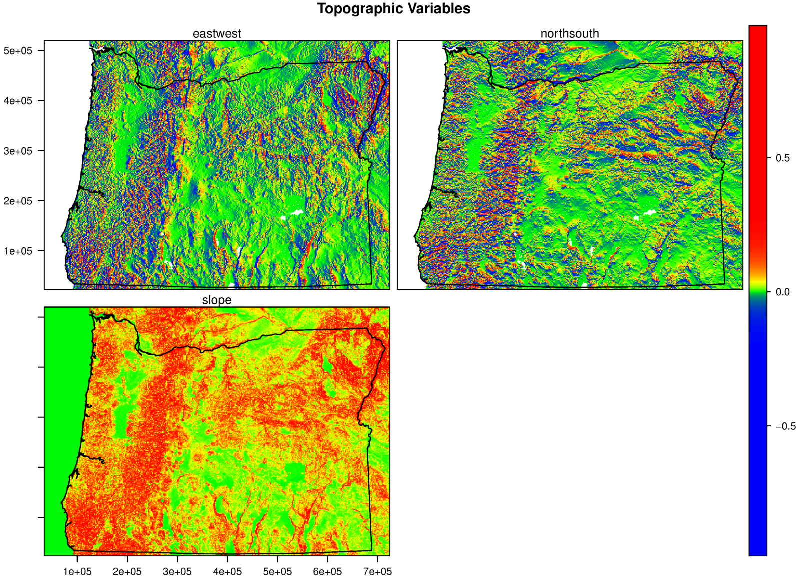
1. Land cover may introduce bias into LST-air temperature relationships
2. Biases may change throughout the year
3. Biases may be hard to estimate on a daily basis
4. Biases may be hard to estimate using only point (station) locations rather than the full grid

So I explored the LST-LULC relationships using the full 1km dataset (or a sample of 10k pixels for modeling, see below)

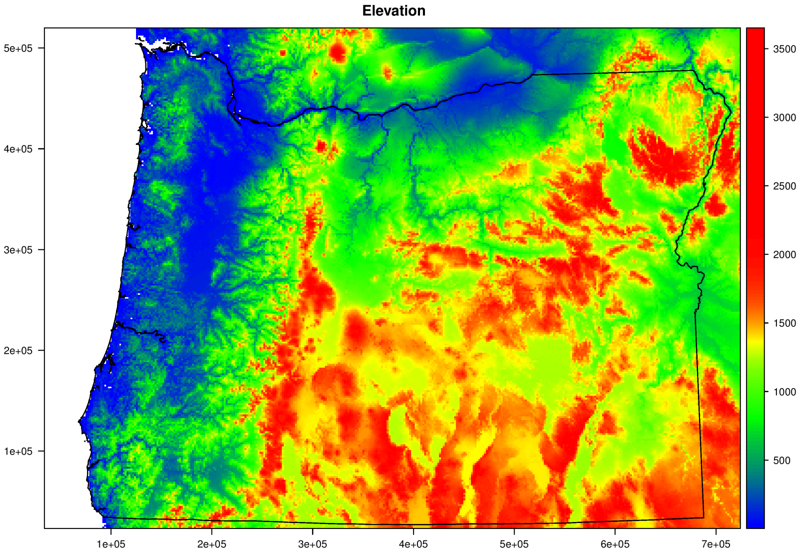
A review of the data



A review of the data

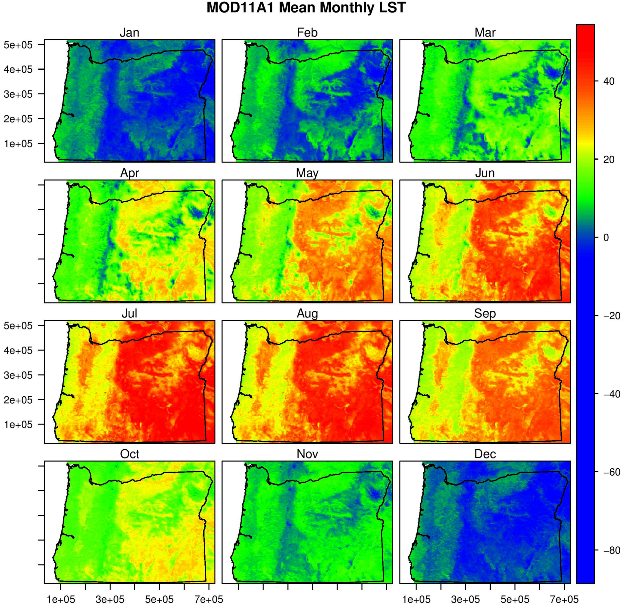


A review of the data



A review of the data

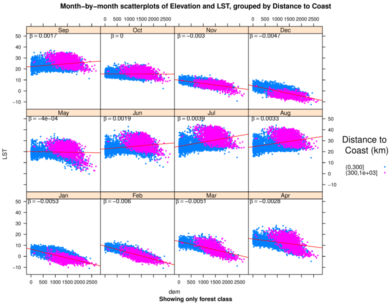
A review of the data



Sampling Bias

Stations are disproportionally in Urban and agricultural areas (surprise!).

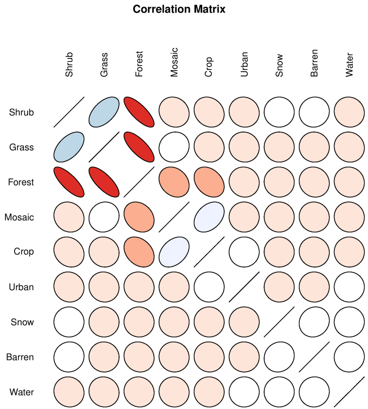
LULC Class	Sample %	Station %
Barren	0.00	0.00
Crop	0.08	0.17
Forest	0.51	0.55
Grass	0.09	0.05
Mosaic	0.00	0.01
Shrub	0.31	0.09
Urban	0.01	0.11
Water	0.01	0.02
Wetland	0.00	0.00



Sample % indicates the proportion of the 10k samples in each LULC class while Station % is proportion of stations in that class.

LULC Correlations

Correlations between the sub-pixel proportions (%) of each LULC

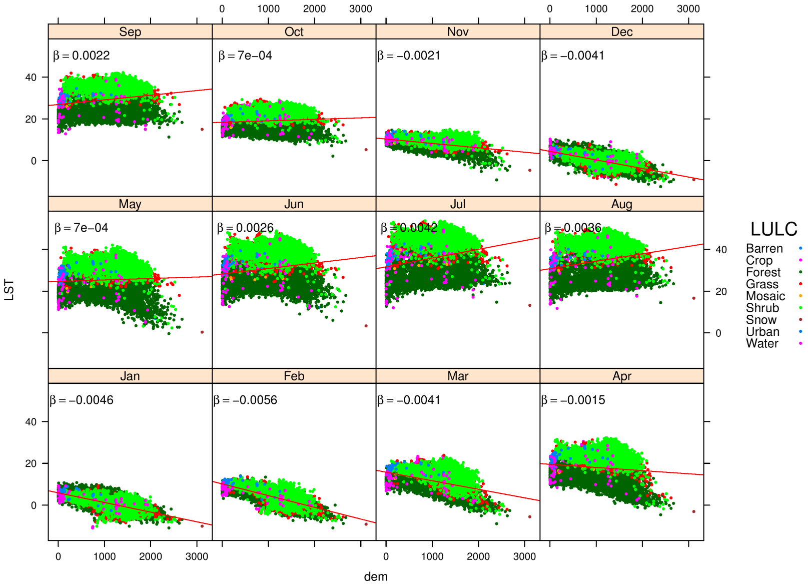


class.

Only Forest shows correlations > 0.6 with any others. In regression, this class can be removed (included in the intercept) and the remaining correlations are low enough to not worry about.

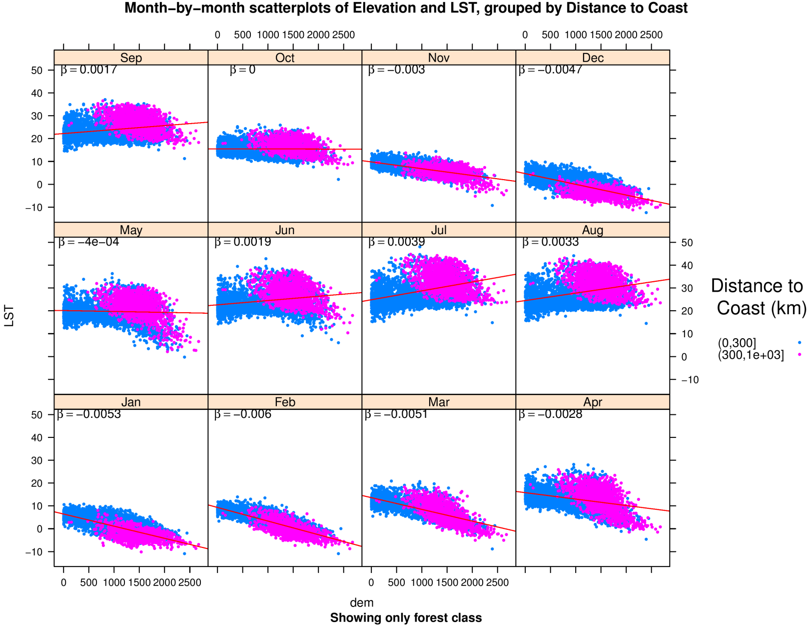
LST-LULC-Elevation Relationships

Month–by–month scatterplots of Elevation and LST, grouped by LULC

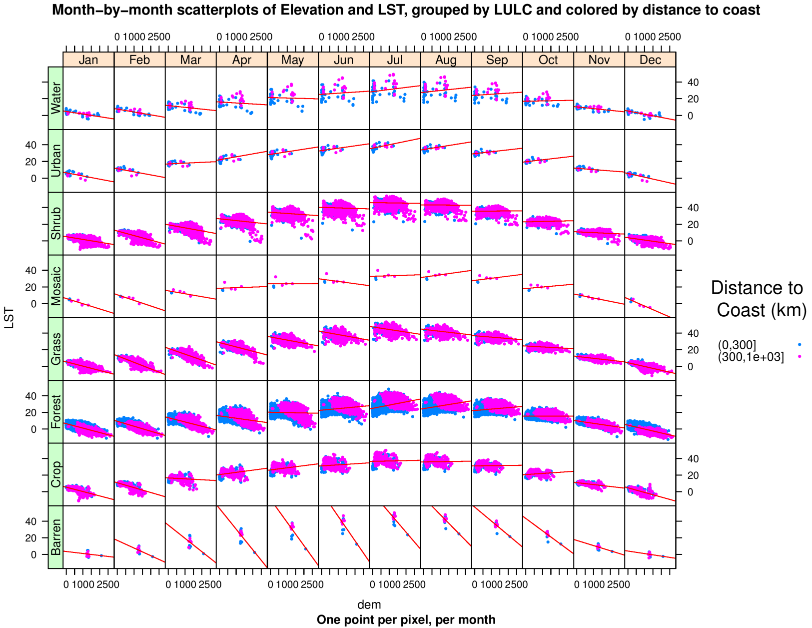


One point per pixel, per month

LST-LULC-Elevation Relationships



LST-LULC-Elevation Relationships



Effects of Distance to Coast - linear?

Seems unlikely that the effects of 'distance to coast' will be linear globally (especially for sites that are far inland). We need a transformation that has larger values by the coast and small values elsewhere, dropping to near zero at some point (~ 100km?) What are the options?

1. use linear anyways...
2. some kind of logistic curve that goes from 1 near the coast to 0 at some point inland. Hard to specify parameters (would be better to fit them, but this would require a more complex model and probably preclude use of existing packages)
3. log(distance)
4. 8-log(distance+1) (8≈log(farthest point on earth from coast)) This is what I tried (but this certainly needs more exploration).

Spatial Regression

$$Y(s) = X^T(s)\beta + W(s) + \epsilon(s) \quad (1)$$

$$\epsilon(s) \sim \mathcal{N}(0, \sigma^2) \text{ : Measurement error} \quad (2)$$

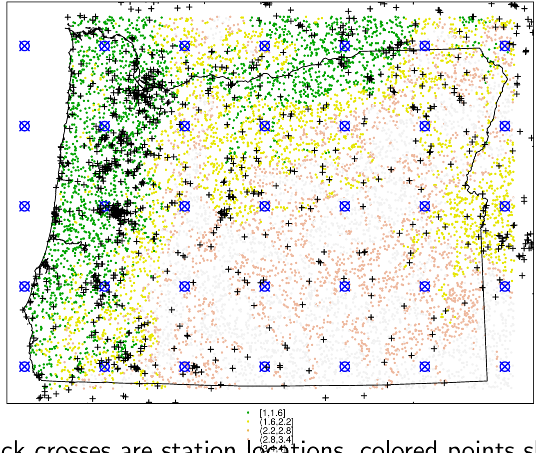
$$W(s) \text{ : Zero-centered Gaussian Process} \quad (3)$$

$$(s) \text{ : indexes space} \quad (4)$$

Objective: Isolate the effects of LULC on LST after accounting for topography (eastwest, northsouth,distance to coast, elevation) and spatial autocorrelation. Fit using spBayes.

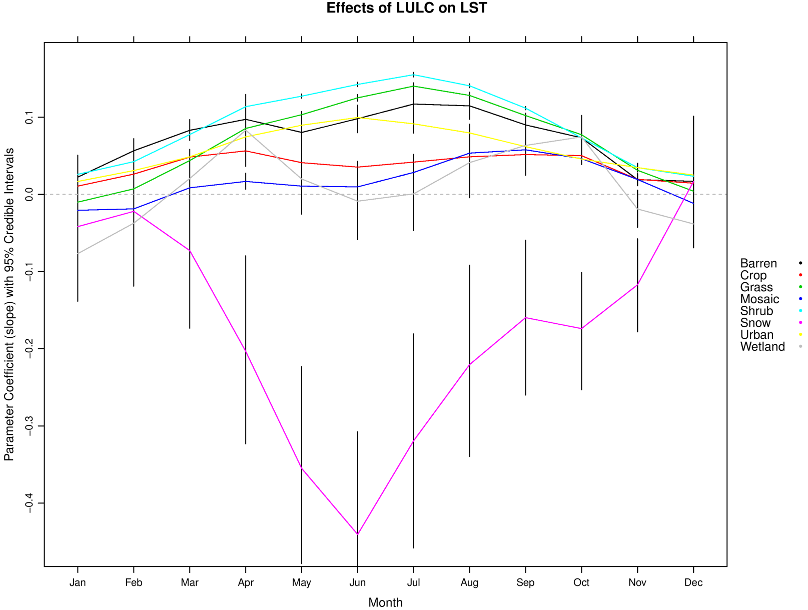
Sampling Bias

10,000 sample locations, stratified by elevation



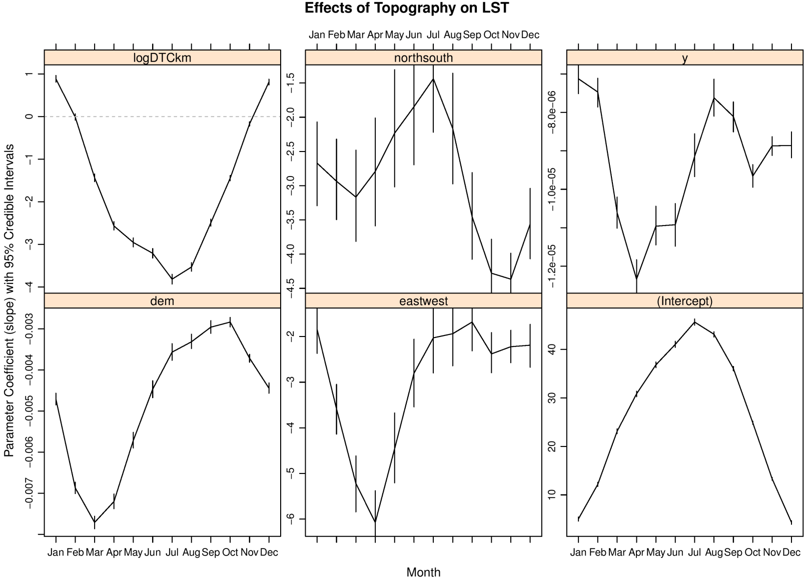
Black crosses are station locations, colored points show the 10k random samples (stratified by elevation - highest class nearly white in color). Blue crosses show points used in spBayes 'predictive process' to estimate spatial effects.

Estimated Parameters by month



Coefficients are unstandardized and represent the change in LST expected with a 1% increase in that class from 100% Forest

Estimated Parameters by month



Coefficients are unstandardized. Intercept is degrees C, eastwest and northsouth range from –1 (90 degree slope) to 0 (flat) to 1 (90 degree slope), and dem is m logDTCkm is in log(km), and y (lat) is m

Summary

- LULC clearly affects LST measurements differently throughout the year (forests are cooler, shrub is warmest).
- Strongest biases are in summer months
- Lapse rate changes significantly throughout the year ($<$-7^{o}C/km in spring, $>$-3^{o}C/km in fall).
- Should compare these results to parameters estimated using only station locations to assess whether spatial station bias (and low n) is biasing model fitting.

Other thoughts

A few other things that caught me attention while working on this...

- Values in monthly LST climatologies range from -88^{o}C to 54^{o}C. Need to do additional quality control?
- Need to estimate effects for all (or at least most common across region) LULC classes to predict.
- Will interaction terms in interpolation models (i.e. lst*lulc) be sufficient to capture these patterns?
- Possible solution: use informed priors from full dataset to inform relationships at the stations (Bayesian only).