David Leifer

11/03/2016

Notes

Reading 5

Spatio-temporal interpolation of soil water, temperature, and electrical conductivity in 3D + T: The Cook Agronomy Farm data set

**Abstract**

-two approaches to spatio-temporal geostat models for 3D and time soil data:

-1. Random forest algorithms

-2. Kriging model post-de-trending the observations of depth-dependent seasonal effects

-uses Cook Agronomy Farm

-includes hourly measurements of soil water content, temp, electrical conductivity

-taken at 42 stations at five depth

-leaves one station out for cross-validation

-kriging was more accurate than trees- temp was most, water was second, electrical was least

**1. Introduction**

-regression-kriging- combines multiple regression model and variogram produces unbiased continuous prediction surfaces

-can summarize 3D + time direction in 2D + time by approximating space/time as horizontal distance to distance in depth and temporal separation

-2D and time is good at predicting soil water when incorporated with daily evapotranspiration and net precipitation covariate estimations

-combine terrain attributes and RS land surface temps to explain 70% of variation in weather station observations

**2. Materials and methods**

*2.1 The Cook Agronomy Farm dataset*

-UW (Washington) State farm

-deep silt loams formed on loess hills-fertile

-covariates: DEM, SAGA wetness (much lawls), NDRE (m), NDRE (sd), Bt horizon, BLD, PHI, Precip\_cum (I’m LOLing), MaxT, Min\_T, Crop

-these are a combination of 2D, 3D, 2D+time

-response variables: VW, C, EC (ded)

-only 3D+time

-NDRE > NDVI and Red Edge > Red

*2.2 Conceptual foundation for 3D +time modeling*

-Z(x, y, d, t) = m(x, y, d, t) + ε′(x, y, d, t) + ε′′(x, y, d, t)

-spacetime-

coordinates

-*d* depth

-*m* is the trend

-ST stochastic

-uncorrelated noise

*2.3 3D + Time random forests model*

-linear regression Generalized Linear Model

-trend model using random forests algorithms positives:

-1) doesn’t need to assume specific distributions for linear relationships

-2) fits predictive model for dataset with high dimensionality

-negatives: computationally intensive, over-fit noisy datasets

*2.4 3D + T kriging model*

-uses seasonal de-trending instead of meteorological covariates

-a lot of R code fitting 2D+time to pseudo 3D variables

-…..

*2.5 Cross-validation*

-cross validations versions:

-Random Forests Prediction:

RF-loc- strict cross validation, leave-one-station-out

RF-rnd- simple cross-validation, randomly subsets points using 5 fold sets of model fitting/validation

-kriging- [kriging-loc] leave-one-station-out using fitted variogram model then validation

*2.6 Software implementation*

-buncha rPackages

**3. Results**

*3.1 3D + T random forests model*

-(%IncMSE) temp is most important covariate for all three observations

-(IncNodePurity):

-temp for water content

-max temp for soil temp

-PHI for EC

*3.2* *3D + T kriging model*

­-yo no say un idea

-more temporal neighbors than spatial neighbors

*3.3 Model accuracy*

-goodness of fit == 90%

-water content and conductivity == not lit

-r2 for soil temp remained high in cross validation

**4. Discussion**

*4.1 Model performance*

-expected weather to be important- only for MSE

-expected air temp to be an important predictor- NOT

-soil properties and covariates water content and electrical conductivity were important

-