ANOCOVA Method

**Part 1 (Corwin and Lesch 2014)**

* 6 reasons that it is hard to measure regional salinity
  + Between-field variation in field-average water content due to irrigation management
  + Between-field variation in soil texture
  + Condition of the soil surface (till v. no-till)
  + Surface geometry (furrows v. no furrows)
  + Temperature differences (surveys conducted at different times of the year)
    - Correction factor ft \* ECa
  + Between-field spatial variation in salinity
* 3 methods for regression
  + Field specific regression (FSR)
    - # of samples required: 3pM
      * P is # of soil depths
      * M is # of fields surveyed
    - Both slope and intercept need to change for field and depth changes
  + Common coefficient regression (CCR)
    - # of samples required: 3p
    - Accuracy tends to be poor
    - Neither slope or intercept change on field changes, but do for depth
  + Analysis of covariance model (ANOCOVA)
    - Calibration field samples required: pM + 2p
    - After calibration sample requirements: p
    - Best compromise between FSR and CCR
      * Intercept changes between fields, and slope remains constant
* ANOCOVA Process
  + Calibrate ANOCOVA regression with a few fields in the region
  + After calibration, as little as two sample locations at each field can be used for prediction
    - Step 1: compute shifted intercept residual
    - Step 2: estimate the intercept at a given depth
    - Step 3: estimate the associated depth-specific variance value
    - Step 4: repeat steps 1-3 for each sampling depth and compute model MSE estimate by averaging the Step 3 depth-specific variance estimates
* Questions/Comments
  + How are calibration fields selected?
    - ESAP uses 6, 12, 20 locations and so how is that implemented into the calibration. (The study had many sample locations and many fields with 6 or more sites for sampling)
    - How do I run an ANOCOVA model? In R maybe?
      * [See here for an example](https://www.r-bloggers.com/analysis-of-covariance-%E2%80%93-extending-simple-linear-regression/)
    - How well does this method work over time?
      * Maybe well because salts don’t move much
      * Maybe not well because moisture changes a lot
* Potential Implementation
  + Take calibration equations from muth2
    - Two sample locations per additional field
      * EM reading and Sat paste extract
    - Adjusts intercepts accordingly

**Part 2 (Corwin and Lesch 2017)**

* Introduction
  + Validation of ANOCOVA method against CCR and FSR method.
  + Two ways to look at regional salts
    - Satellite imagery
    - ANOCOVA
  + Of the 1.5 x 109ha of cultivated land, 23% saline and 37% are sodic (Massoud, 1981).
  + Red River Valley in CA
    - High water table due to management practices
    - Specifically Coachella Valley
* Materials and Methods
  + Study Area
    - Demographics and background of salinity issue
    - Dataset
      * Quality tested EM38 data for “internal consistency and reliability, that is, properly correlated and aligned EMI readings that were all positive and devoid of gross outliers, systematic instrument bias, and any obvious calibration error effects.”
    - EM surveys
      * Two depths measured for salinity distribution
      * Summary statistics provided (see paper table 1)
    - General characteristics of field sites
      * Where fields were, how they were sampled, **used ESAP to select sites in each field**
    - Compared ANOCOVA to FSR to CCR
      * See part 1 for description of methods
      * **IMPORTANT: ANOCOVA**
        + Many of the field specific effects on EM38 survey data, such as seasonal changes in soil temperature, bed-furrow geometry, surface roughness, and instrument placement height are multiplicative
        + *On a log transformed scale, these affects become additive constants, which in theory should only affect the intercept coefficient (WOW)*
        + This means that there are very good reasons to think that the ANOCOVA model could actually outperform the FSR model, particularly when only limited calibration data is available and the EM38 data was acquired over a short time span (Like us!)
    - Results
      * Some fields were influenced by high and low moisture based on the DPPC results
      * ANOCOVA out performed the FSR and CCR models in MPSE, but FSR was the best for just MSE.
        + This means that ANOCOVA is best at prediction, but FSR fit the sampled data best.
    - From Dennis Corwin at USDA ARS
      * How do you go about selecting the 1-3 additional sampling sites per field for ANOCOVA?
      * Does my survey and Sampling need to take place at the same time?
      * How do you convert an ESAP regression that uses z1 and z2 into ln(ECav) and ln(ECah)?
      * What does esap stand for?
        + Electromagnetic Sampling Analysis Design? Protocol?

**Phone Call with Dennis Corwin (2/5/18)**

* In both papers, you selected enough sites to run a Field Specific Regression, and an ANOCOVA model. If I were to start with only the ANOCOVA method in mind, how would I select my 1-2 sample sites for each additional field? ESAP chooses a minimum of 6 using RSSD. I thought perhaps you could just choose an area of high conductivity and an area of low conductivity (with a "decent" distance between them). I was curious as to your thoughts on this.
  + 2 approaches
    - Take EM survey, and just take high and low survey value (preferably 3 points though)
    - Take EM survey, and use ESAP to generate points for collection
      * Pick high and low from ESAP (3 points if possible)
      * If texture is an issue, probably take all 6 samples
    - Make and ANCOVA for each
* I have lots of survey data from an EM38 survey performed 2 years ago during the growing season (aprox. June 2016). Would you feel comfortable using this conductivity data to select sampling sites for this year? Or would you recommend performing another series of surveys before choosing sample sites?
  + DEEP ploughing ruins correlation (ask Leon about it)
* Lastly, in the ANOCOVA papers, your models consistently used ln(EMv) and ln(EMh) as the regessors for the ln(ECe), but in ESAP, transformed signals (z1 and z2) are used as the regessors for ln(ECe). How would I go about handling this in developing an ANOCOVA model, when my ESAP model only uses my transformed z1 and z2 inputs?
  + First you wanna see if ECa data is log normally distributed or normally distributed
    - Create model from log normal readings of EMv and EMh
    - Create model with EMh if water table too shallow
      * Trend factor (slope)
      * Validation set of data
* Course in Salinity at Riverside!
  + UCDavis in March, half day program

**Phone Call with Dennis Corwin (4/4/18)**

* Did you obtain your gravimetric water contents from a separate soil sample taken at the same location of the saturated paste samples? (i.e. two holes at the same sampling location, one for ECe and one for GWC) and how did you establish saturation percentage?
  + Will only really affect me if there’s abrupt change in texture
  + E.g. Kaess 1 is sandly, muth2 is clay
    - You want to separate surveys for each field because the water contents will vary so much
    - Texture is so important!
    - 450g
    - Take sample of saturated paste for saturation percentage. Texture is almost perfectly correlated with SP%, so much is can be determined very reliably
    - When defending, DO NOT assume saturated paste SP% is the same as field SP%. Soil physicists will be angered.
* Where did you account for changes in EM38 readings due to water content in the ANOCOVA equation? The equation shown shows an intercept, EMv, and EMh. Don't adjustments need to be made to the EMv and EMh values because they fluctuate based on water contents?
* How important is it to perform the whole survey at once? Ideally I'd like to get my surveys and sampling done within a two day period, but with my limited labor available, I was thinking about breaking up the survey in to 2 or 3 sections. Does this prevent fields from being grouped into the same ANOCOVA analysis?
  + It is OK to split up the survey so long as the moisture in the field is good, and the readings and samples are taken simultaneously
* Travel Plans to California
* What Dennis Reccommends
  + Stay away from DA1, the high water table will mess with calibration
  + In all cases, when performing the survey, try to get to field capacity (below 70% is okay, but 50% is really not good)
  + Recommend 3 spots per field, but Muth6 should have 6 (can potentially)
  + Deep ripping
    - Sample between rippers, and survey perpendicular to the ripping
  + Water table
    - DO NOT trust readings from an area with a shallow water table
    - Base everything on horizontal readings
  + Any reading below .1 dS/m is questionable –could vary, this is for san juaquin and coachella valley. Colorado may be different.
    - Some areas don’t go above that though b/c of low salinity, so it’s only detecting other factors (OM, texture, ETc…)

Works Cited

Corwin, D. and S. Lesch (2017). "Validation of the ANOCOVA model for regional‐scale ECa to ECe calibration." Soil Use and Management **33**(2): 178-190.

Corwin, D. L. and S. M. Lesch (2014). "A simplified regional-scale electromagnetic induction—Salinity calibration model using ANOCOVA modeling techniques." Geoderma **230**: 288-295.