Consulting Project of Soil Moisture: Final Report

Client: Department of Soil Science
Advisor: Dr. Ana-Maria Staicu
Zhou LI, Yingzi XU

Department of Statistics
North Carolina State University

April 24, 2013

Data description

For year=2008 and depth=10cm:

- M_{ij}: Soil moisture of subject i at time t_j (t_j is different from year to year and for 2008 t_j ranges from Jun 13 to Dec 19);
- g: Tillage group indicator for each subject i, $g = 1, 2, \ldots, 6$;
- t_j: Time point when soil moisture is measured;
- p_j: Precipitation (cm) at time t_j.

where $i = 1, 2, \ldots, 24, j = 1, 2, \ldots, 25$.

Data description

Main Goals

- How the trend profiles vary over time in each treatment group?
- Whether the trend is different across tillage treatments?

Generalized additive models (GAM)

$$g(\mu) = \beta_0 + f_1(x_1) + \cdots + f_m(x_m) \tag{1}$$

The functions $f_i(x_i)$ in (1) may be fit using parametric or non-parametric means, thus providing the potential for better fits to data than other methods.

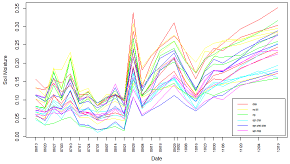
Remove Seasonality

Assume

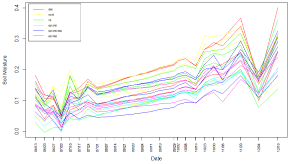
$$M_{ij} = \mu_{ij} + S_{ij} + \epsilon_{ij} \tag{2}$$

1. μ_{ij}: trend; ϵ_{ij}: noise; $S_{ij} = S_{i,j-d}$
2. $d = 12$ denotes the (approximate) number of weeks during the period
3. Treat fall and winter have the same seasonal component

Remove Seasonality



(a) Original data



(b) Deseasoned data

Figure: Soil moisture after removing seasonality for year 2008 and depth=10 cm

Determine covariance structure

Simple model

$$\boldsymbol{Y}_i^g = \beta_0 + \boldsymbol{\mu}_g + \epsilon_i \tag{3}$$

Model 1

$$\boldsymbol{Y}_i^g = \beta_0 + \boldsymbol{\mu}_1 + \sum_{g'=2}^{6} \alpha_{g'} \mathbb{1}\{g = g'\} + \epsilon_i \tag{4}$$

where $\epsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2)$.

Determine covariance structure

Hypothesis Testing

H_0 : No difference between any treatment is equivalent to $H_0 : \alpha_2 = \cdots = \alpha_6 = 0$.

Comments for Model 1

1. Assumption is too strong;
2. Still helpful for us to determine a better covariance structure.

GAM based on transformed data

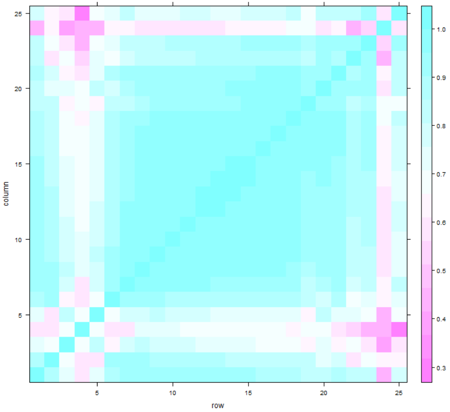


Figure: Level plot of correlation matrix for residuals of model 1

GAM based on transformed data

We assume that each element of ϵ_i is actually the sum of two random terms, i.e.

$$\epsilon_{ij} = b_i + e_{ij}$$

$$\boldsymbol{\Sigma}_2 = Var(\epsilon_i) = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

GAM based on transformed data

Model 2

$$\boldsymbol{Y}_i^{g*} = \beta_0^* + \boldsymbol{\mu}_1^* + \sum_{g'=2}^{6} \boldsymbol{\alpha}_{g'}^* \mathbb{1}\{g = g'\} + \boldsymbol{\epsilon}_i^* \tag{5}$$

where $\boldsymbol{Y}_i^{g*} = \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \boldsymbol{Y}_i^g, \beta_0^* = \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \beta_0, \boldsymbol{\mu}_1^* = \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \boldsymbol{\mu}_1, \boldsymbol{\alpha}_{g'}^* = \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \boldsymbol{\alpha}_{g'}, \boldsymbol{\epsilon}_i^* = \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \boldsymbol{\epsilon}_i$.

In model 2, $\epsilon_{ij}^* \overset{\text{iid}}{\sim} N(0, 1)$

GAM based on transformed data

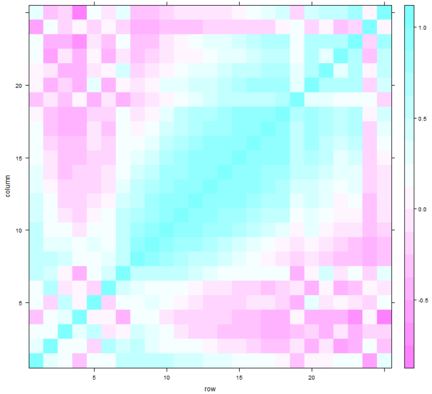


Figure: Level plot of correlation matrix for residuals of model 2

GAM based on transformed data

	edf	Ref.df	F	p-value
μ_1^*	7.95	8.51	61.33	0.00
α_2^*	7.53	8.60	1.52	0.14
α_3^*	7.53	8.60	0.62	0.78
α_4^*	7.53	8.60	2.65	0.01
α_5^*	7.53	8.60	1.42	0.18
α_6^*	7.53	8.60	2.93	0.00

Table: Approximate significance of smooth terms

GAM based on transformed data

We can also restore the original trend estimate by setting

$$\begin{aligned} \hat{\beta}_0 &= \hat{\boldsymbol{\Sigma}}^{\frac{1}{2}} \hat{\beta}_0^* \\ \hat{\boldsymbol{\mu}}_1 &= \hat{\boldsymbol{\Sigma}}^{\frac{1}{2}} \hat{\boldsymbol{\mu}}_1^* \\ \hat{\boldsymbol{\alpha}}_{g'} &= \hat{\boldsymbol{\Sigma}}^{\frac{1}{2}} \hat{\boldsymbol{\alpha}}_{g'}^* \end{aligned}$$

GAM based on transformed data

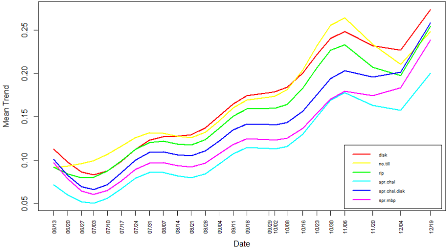


Figure: Estimated mean trend for different tillage groups

GAM based on transformed data

Model 3

$$\boldsymbol{Y}_i^g = \beta_0 + \boldsymbol{\mu}_1 + \sum_{g'=2}^{6} \boldsymbol{\alpha}_{g'} \mathbb{1}\{g = g'\} + \boldsymbol{B}_{b(i)} + \epsilon_i \qquad (6)$$

where $B_1, B_2, B_3, B_4 \overset{\text{iid}}{\sim} N(0, \sigma_B^2)$, $Var(\epsilon_i)$ follows Markov Structure.

Sad results.

Conclusions

- Treatment 1 (disk) does not differ significantly on soil moisture from treatment 2 (no.till), treatment 3 (rip) and treatment 5 (spr.chsl.disk);
- Treatment 1 is different from treatment 4 (spr.chsl) and treatment 6 (spr.mbp), which lead to lower soil moisture;
- Similarly we can test the difference between any tillage groups by changing the baseline trend in (4).

Conclusions

Trt	vs	Trt	Trt	vs	Trt	Trt	vs	Trt
2	$\approx$	1	4	$<$	1	6	$<$	1
2	$=$	2	4	$<$	2	6	$<$	2
2	$\approx$	3	4	$\approx$	3	6	$\approx$	3
2	$>$	4	4	$=$	4	6	$\approx$	4
2	$>$	5	4	$\approx$	5	6	$\approx$	5
2	$>$	6	4	$\approx$	6	6	$=$	6

Table: Pairwise comparison between different treatments

Thank you!