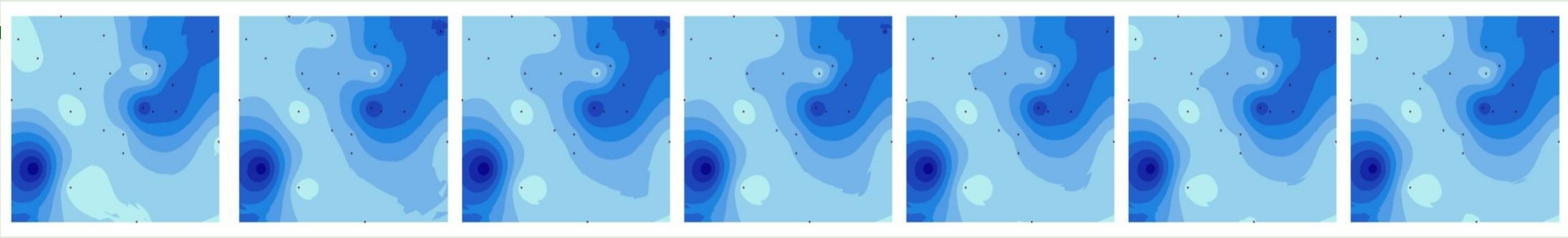
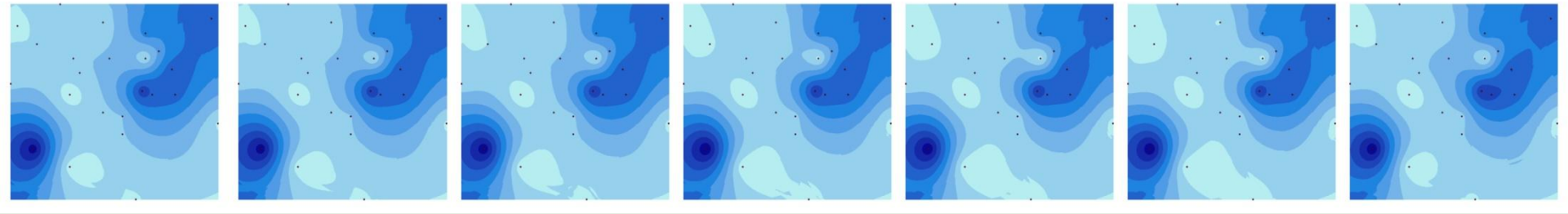


# Analysing spatial and temporal patterns of soil water in small forest areas by using GAMs



**Yvonne Morgenstern**
**Forest Research Institute of Baden-Württemberg, Germany**

Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg

# Content

1. Aim of project
2. Dataset – RMP concept
3. First steps with GAMs
4. Many questions

# Main Aim of Project

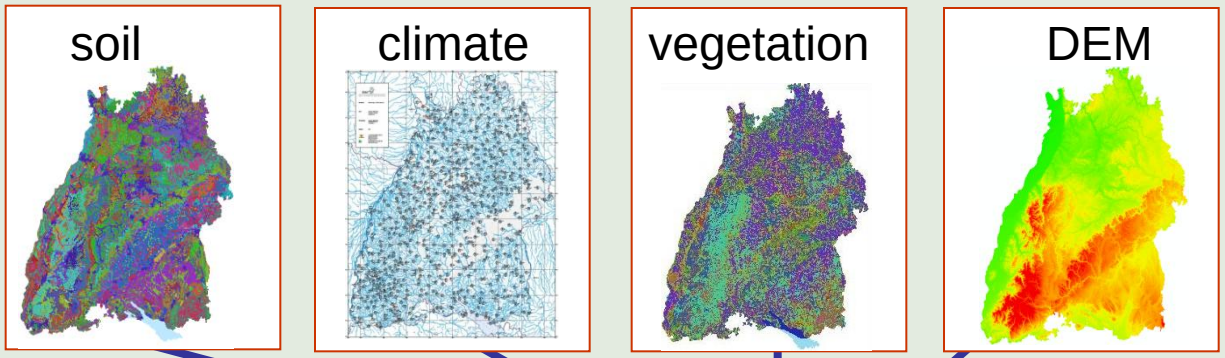
Development of a regional concept to quantify the probability of soil water deficit in small forest areas

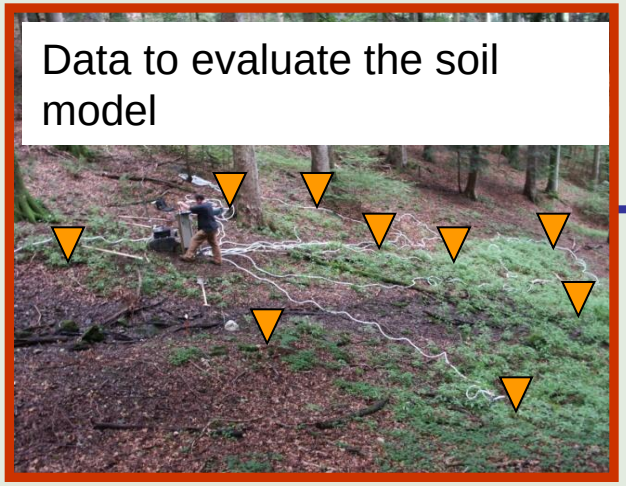
→ spatial distribution of water deficit

→ temporal dynamics of water deficit
(frequently and duration)

→ change of drought stress risk due to climate change

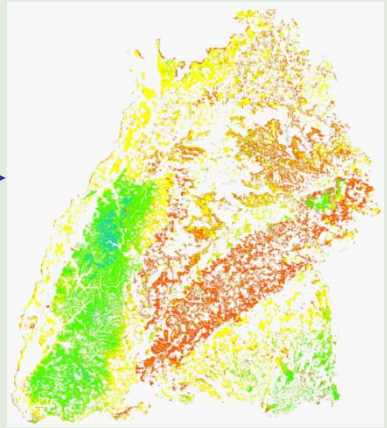
# Modeling of soil water with a water balance model




?

Water balance model LWF-Brook90



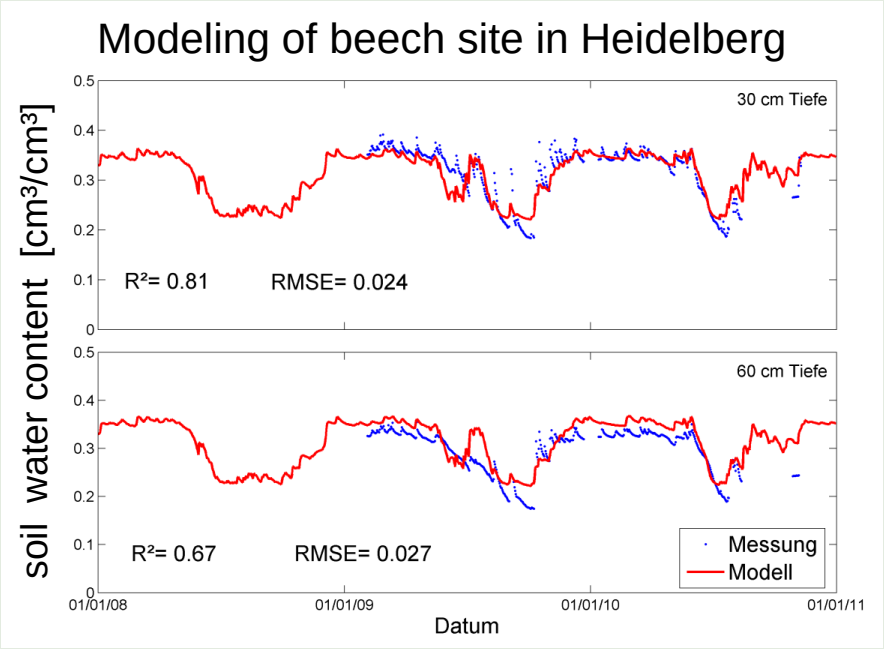
Result:

Maps of values of available water content per day

# Data for validation a water balance model

Location of intensive measurement stations in forest areas in Baden-Württemberg


Only temporal patterns of soil water in different depths

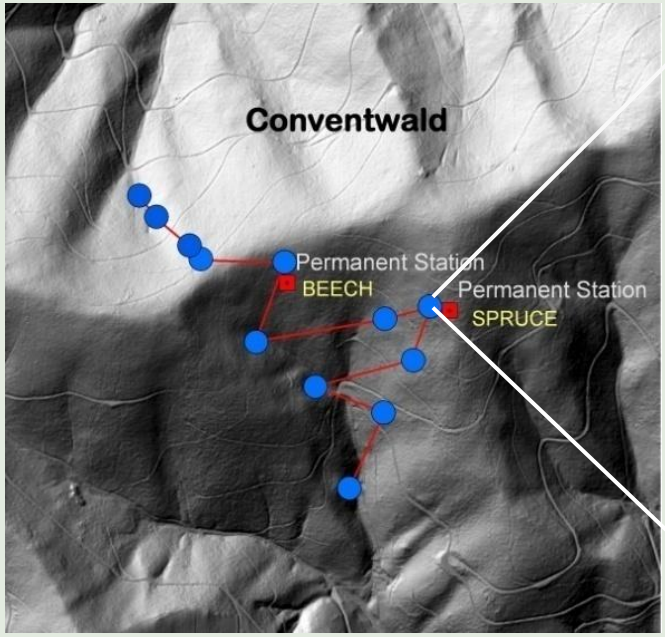



No measurements of spatial variability → Apply RMP-concept

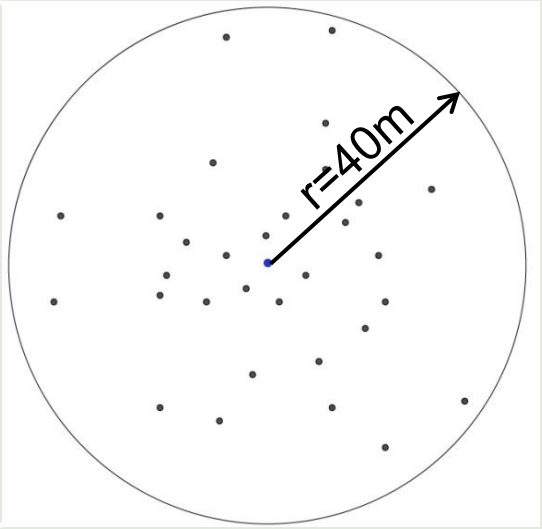
Which site factors influence the variability of soil moisture ?

# Application of Randomised Moving Plots- concept

## 4 x 12 plots in forest areas in southwest Germany




**Transect 800 – 1000 m, 12 plots**




**1 plot with**
**31 TDR- position**

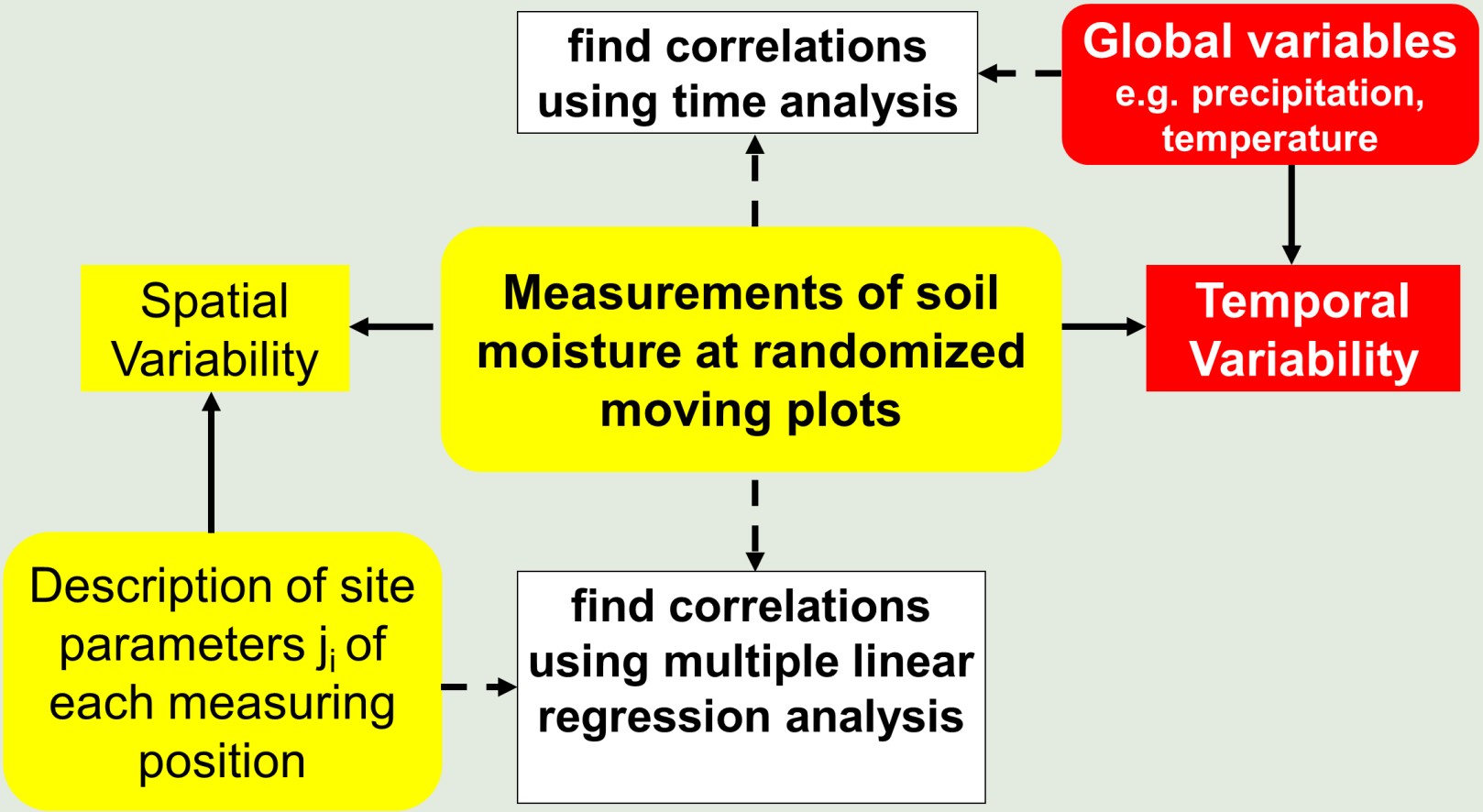
Soil moisture measurements:

- 31 TDR-positions
- Undisturbed installation
- 14-days periods
- Half-hour value
- 2 measurement periods per plot

- Selection of plots: forest stand, aspect, slope and topographic wetness index
- 2 intensive measurement plots (soil moisture and tension, precipitation) and 1 climate station

# Idea of RMP concept

$$\theta(x,y,t) = \theta_{temporal}(t) + \theta_{spatial}(x,y,t) + \varepsilon_{\theta}(x,y,t)$$



# Possible Predictors

Information at each TDR sensor position:

- To level exact position (easting, northing)
- Morphologic parameters from 1m grid DEM
- Vegetation parameters
- Soil profile description

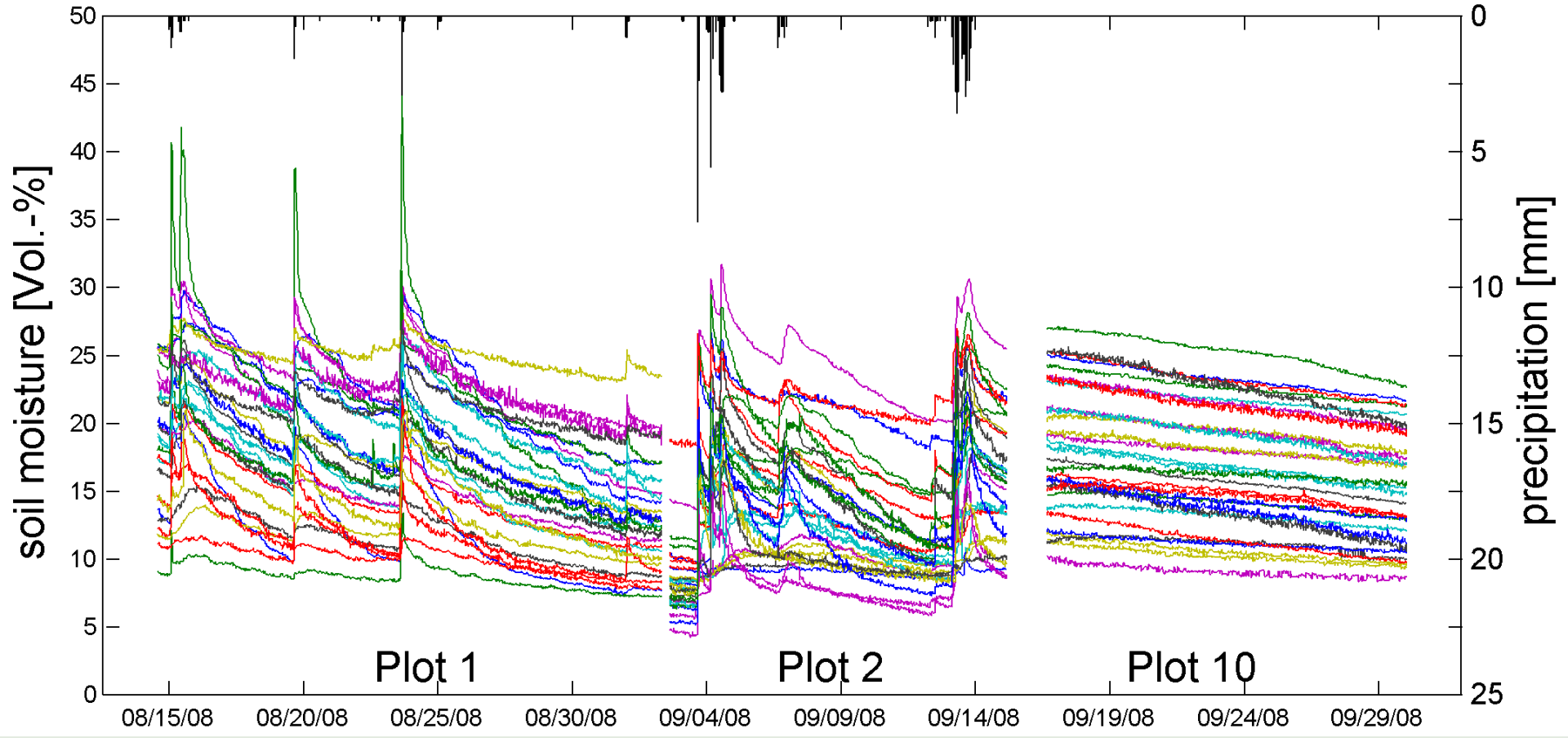
→ 150 site factors – spatial variability

Information for each measurement time step per area:

- Precipitation P, air humid RF, air temperature T, radiation …
- cumulated climate factors (sum of P 48 hour, mean RF30 hour…)

→ 17 climate factors – temporal variability

# Example of Measurements




very hight spatial and temporal variability of soil moisture

| Plot | 1 | 2 | 10 |
|---|---|---|---|
| Max | 44.1 | 31.7 | 27.1 |
| Mean | 17.3 | 13.8 | 16.8 |
| Min | 7.2 | 4.2. | 8.3 |
| CV [%] | **28.2** | **31.6** | **27.4** |
| Precip. [mm] | 22 | 103 | 0 |

# Developing a GAM for soil moisture

**Dataset:** 1.271.984 measurements in half an hour resolution

Measurement variable: soil moisture SW

| ID | plot | area | date | SW | relSWC | relAWC | easting | northing | Slope …… |
|---|---|---|---|---|---|---|---|---|---|
| 2110724 | 1..12 | 1...4 | 11/03/08 19:30:00 | 28.0 | 57.2 | 58.6 | 3572052 | 5320166 | 11.734 |

**Response variable for GAM**

Transformation of soil moisture:

relative saturated water content = **rel. SWC**

relative available water content = rel. AWC

**Better comparison of measurements**

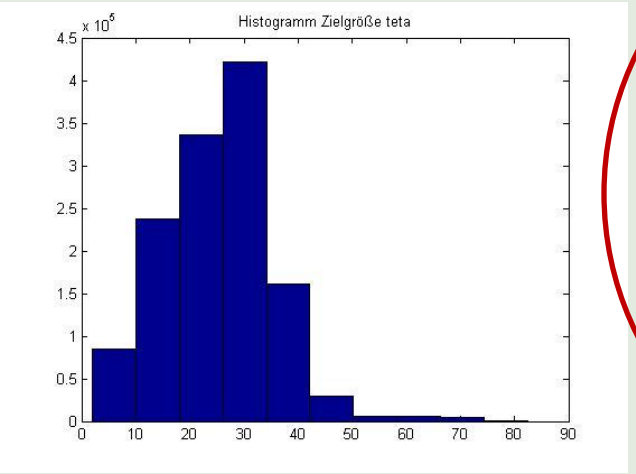
# Developing a GAM for soil moisture

1. Build a possible GAM-structure
2. Plot histogram of the response variable
3. Make scatterplots of target value and variables
4. Select some variables with a high correlation or trend to relative saturation water content
5. Build a complete data matrix with spatial and temporal information to each measurement
6. Build a GAM step by step, beginning with one variable as fixed effect
7. Compare models by AIC, BIC, p-values, adj. $R^2$, degrees of freedom

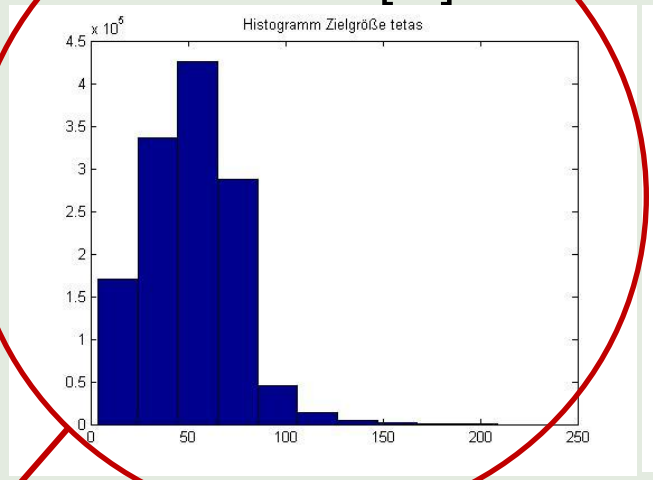
# Developing a GAM for soil water

Distribution of

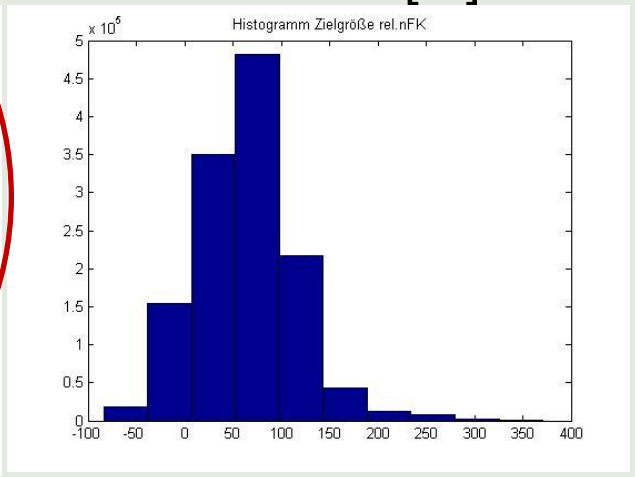
Measurements




rel. SWC [%]




rel. AWC [%]

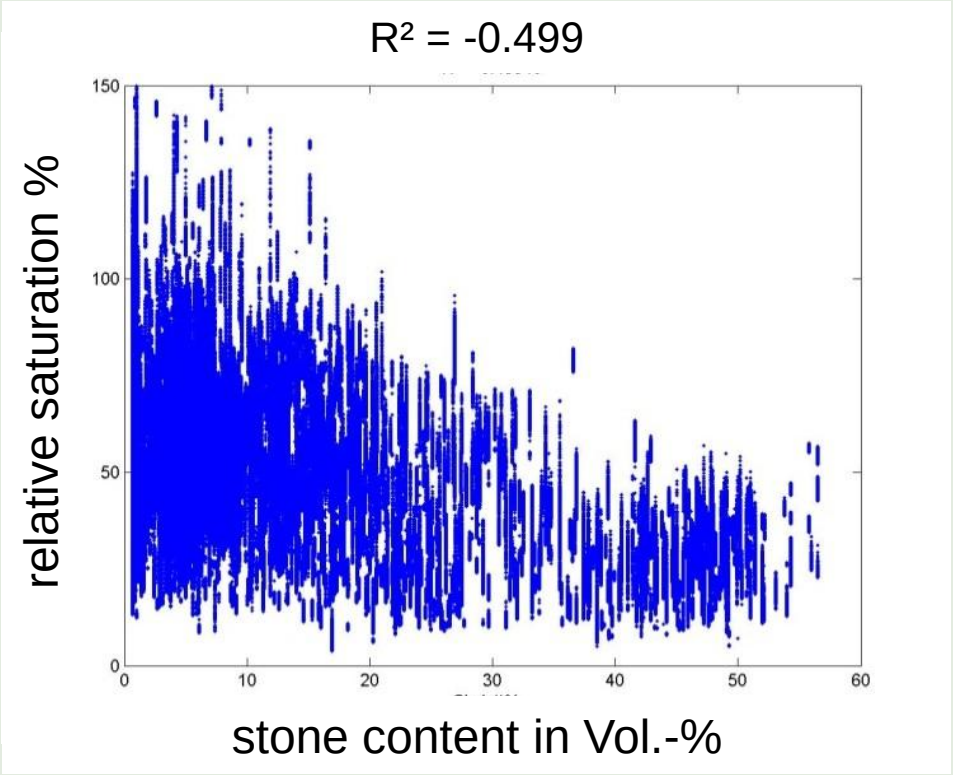


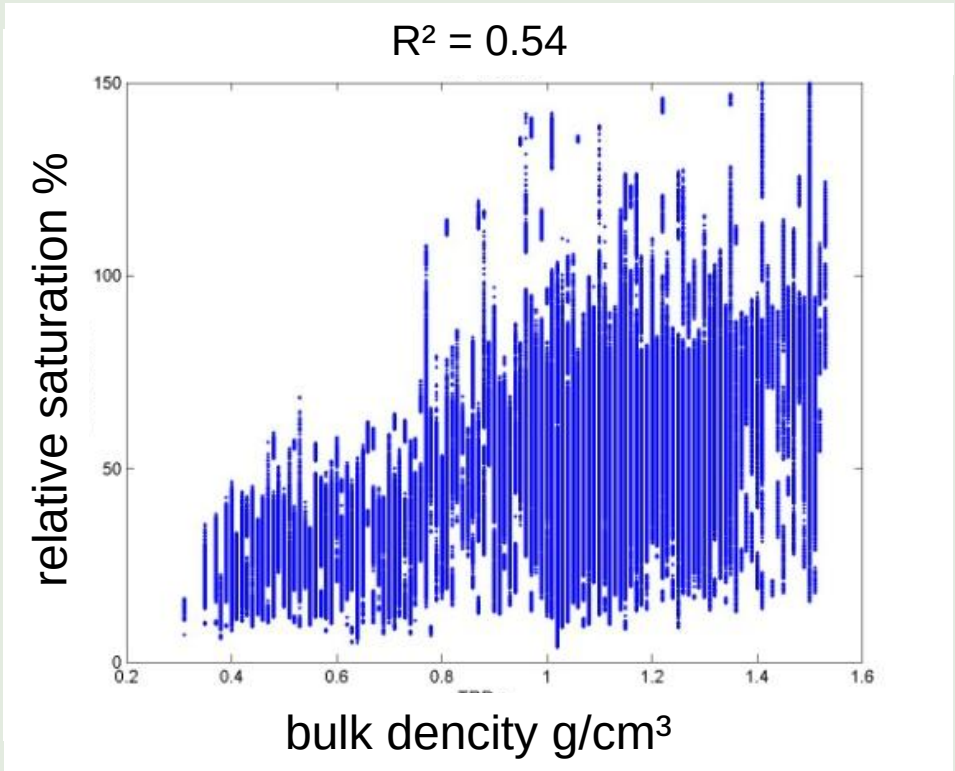

$$\log(y) = x\beta_i + b(s) + d(t) + \varepsilon$$

fixed effects + spatial correlation + temporal correlation + random effects $N(0,\sigma^2)$

# Scatterplots of variables and target value

select variables with a correlation to target value




- slope, vertical distance to water bed, sand content, loamy content, silt content, altitude, coordinates, slope-length factor, stream power index
- precipitation, rel. humidity, global radiation, air temperature
- sum of precipitation/radiation of last 48 hour, mean air temperature/rel. humidity last 30 hours

# Build a GAMM

**Fit.GAMM.53 = gamm(relSWC~s(bulk dencity), family=Gamma(link=log), data=daten)**

| AIC | BIC | adj.$R^2$ | significant |
|---|---|---|---|
| 89368.18 | 89406.24 | 0.255 | <2e-16 *** |

summary(Fit.GAMM.53$gam)
Family: Gamma Link function: log
Formula:tetas ~ s(bulk_dencity)

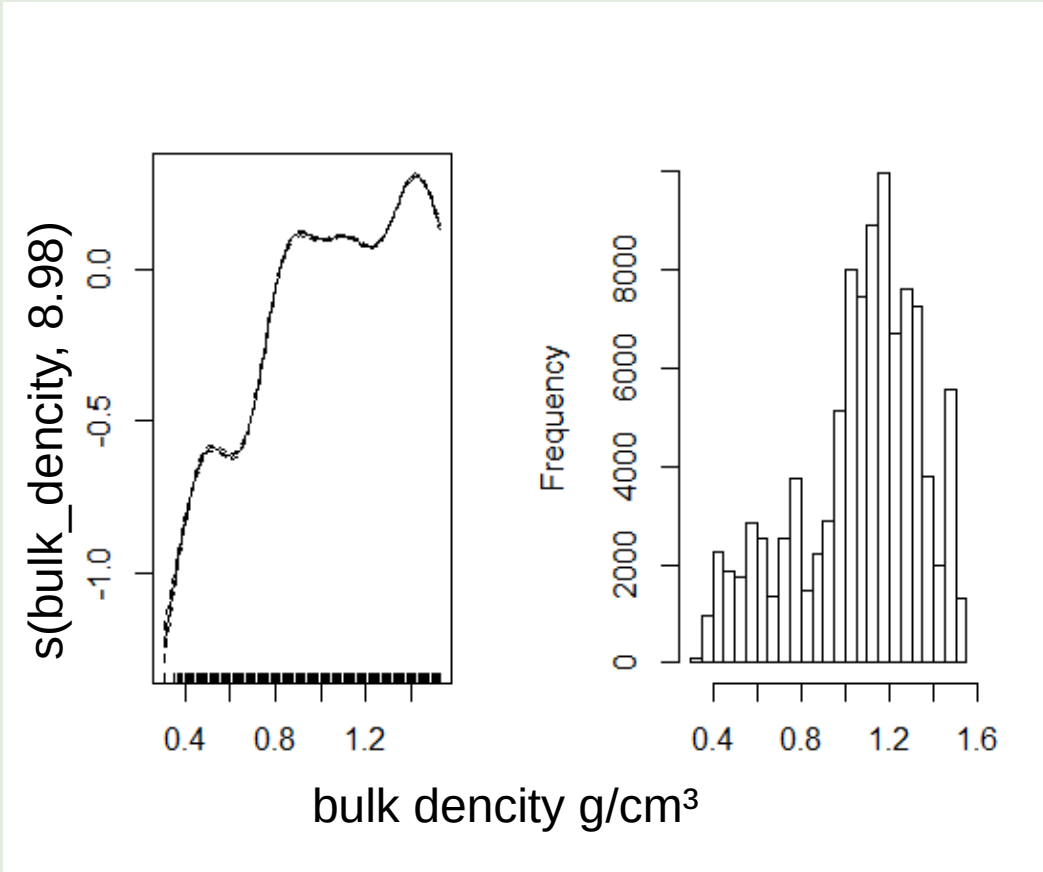
Parametric coefficients:

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 3.903980 | 0.001196 | 3265 | <2e-16 |

Approximate significance of smooth terms:

| | df | Ref.df | F | p-value |
|---|---|---|---|---|
| s(b.d.) | 8.978 | 8.978 | 6218 | <2e-16 |

Scale est. = 0.14299
n = 100000



# Build a GAM

**Fit.GAMM.7 = gamm(rel.SWC ~s(easting,northing), family=Gamma(link=log), data=daten)**

| AIC | BIC | adj.$R^2$ | significant |
|---|---|---|---|
| 80041.23 | 80088.8 | 0.362 | <2e-16 *** |

summary(Fit.GAMM.7$gam)

Family: Gamma
Link function: log
Formula: rel.SWC ~ s(easting, northing)

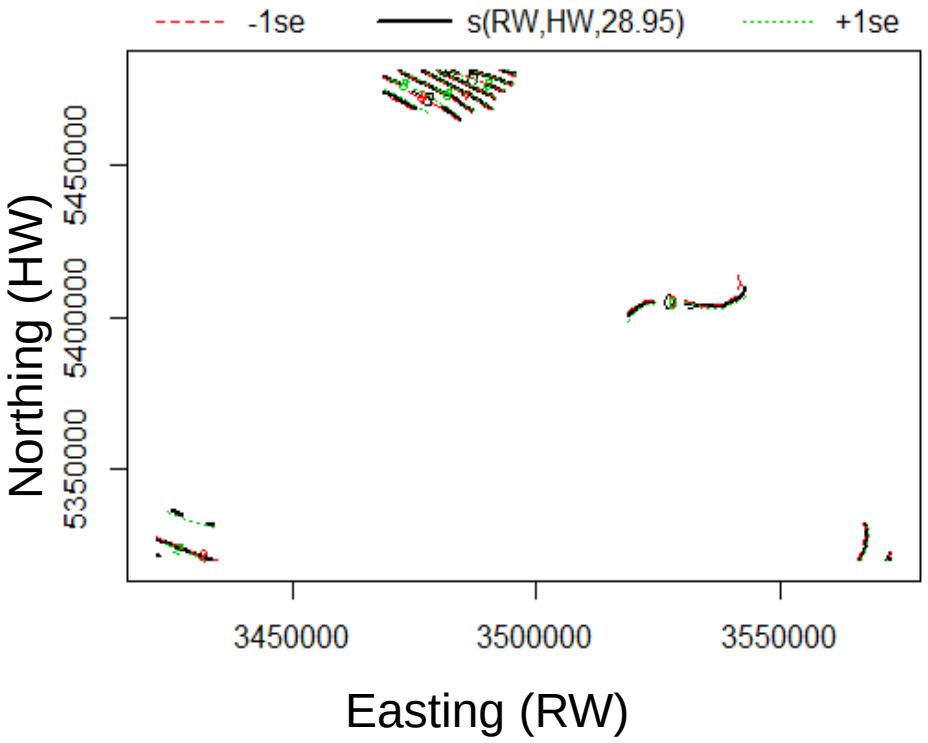
Parametric coefficients:

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 3.89347 | 0.00114 | 3414 | <2e-16 |

Approximate significance of smooth terms:

| | edf | Ref.df | F | p-value |
|---|---|---|---|---|
| s(east.,north.) | 28.95 | 28.95 | 2646 | <2e-16 |

Scale est. = 0.14299
n = 100000



# Build a GAM

**Fit.GAMM.11 = gamm(rel.SWC ~s(Psum48h)+s(Tmean30h)+s(bulk density)+s(gravel), family=Gamma(link=log), data=daten)**

summary(Fit.GAMM.11$gam)

Family: Gamma
Link function: log
Formula: relSWC ~ s(Psum48h)+ s(TMean30h)
+ s(skelett_r) + s(TRD_r)
Scale est. = 0.13928 n = 94395

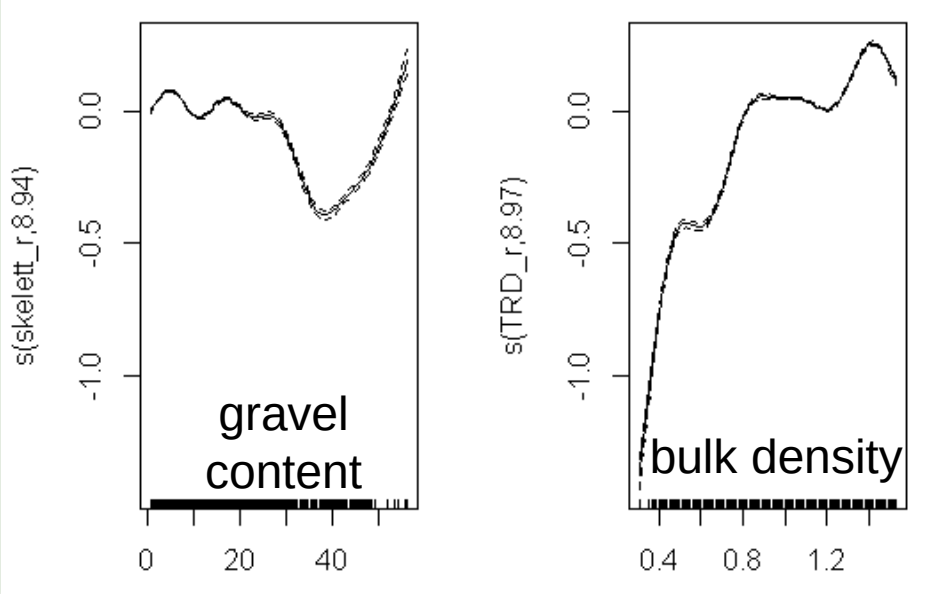
Parametric coefficients:

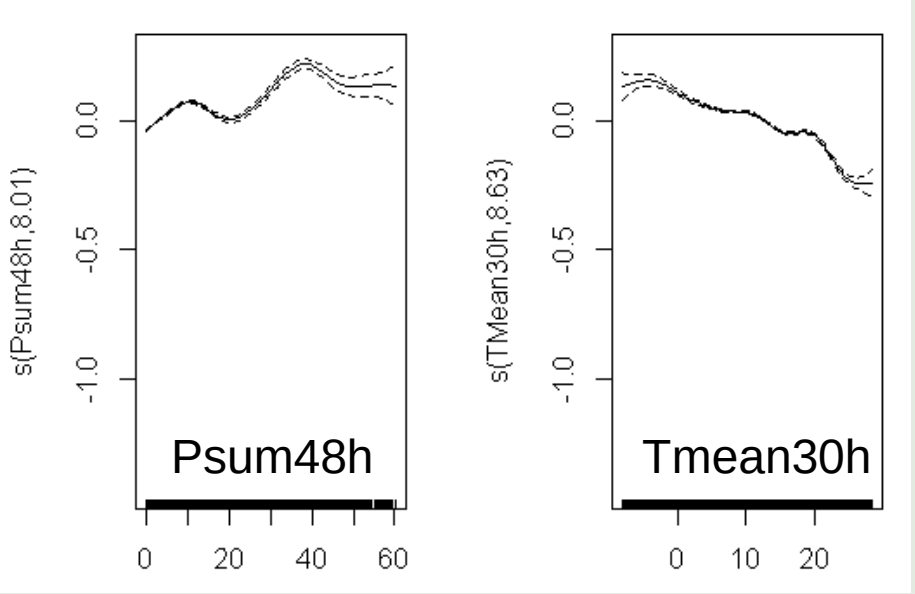
| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 3.923059 | 0.001215 | 3230 | <2e-16 *** |

| AIC | BIC | adj.R² | significant |
|---|---|---|---|
| 82020.22 | 82114.77 | 0.292 | <2e-16 *** |

Approximate significance of smooth terms:

| | edf | Ref.d f | F | p-value |
|---|---|---|---|---|
| s(Psum48h) | 8.008 | 8.008 | 205.6 | <2e-16 *** |
| s(TMean30h) | 8.634 | 8.634 | 251.6 | <2e-16 *** |
| s(skelett_r) | 8.939 | 8.939 | 403.9 | <2e-16 *** |
| s(TRD_r) | 8.971 | 8.971 | 1342.7 | <2e-16 *** |



# Questions

- Make the transformation of target value sense?
- Large dataset, better to reduce it, e.g. mean day values?
- How to select the important variables ?
- Transformation of variables to get better distribution?
- How can I respect  autocorrelation in time series?
- Time format- date or numbers?

# Thank you for your attention


Thanks to my colleagues for field measurements, laboratory and data analysis

# Example of spatial variability

Use geostatistical analysis and kriging method to get a dynamic map of soil moisture

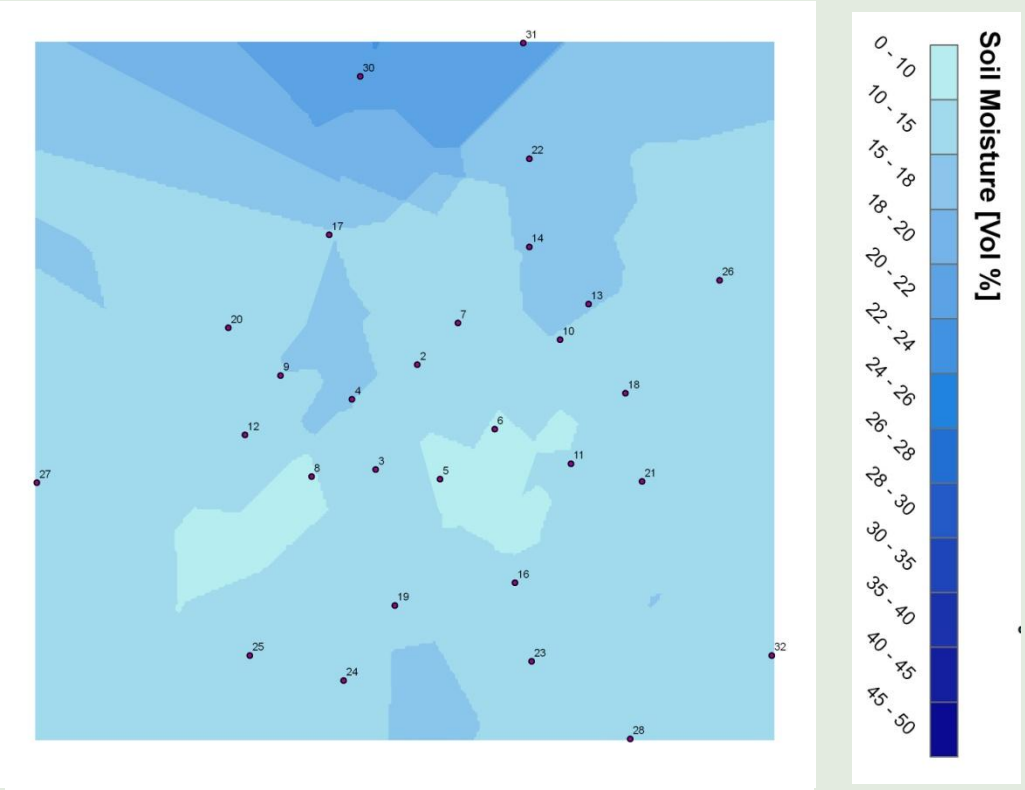
- Mean value per day
- Create variograms

e.g. RMP8

Lag = 2.3-7.5 m

Sill = 28-35 Vol.%

Nugget = 0 - 20 Vol. %




→ Get data for validating the regression model

# Build a GAM

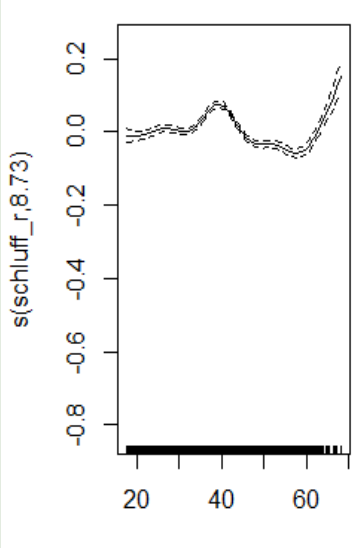
**Fit.GAMM.8 = gamm(relSWC~**s(RW,HW,bs="ts")+s(slope)+s(skelett_r)s(TRD_r)+s(schluff_r)**, family=Gamma(link=log), data=daten)**

| AIC | BIC | Adj.R² | significant |
|---|---|---|---|
| 77141.28 | 77245.92 | 0.385 | <2e-16*** all |

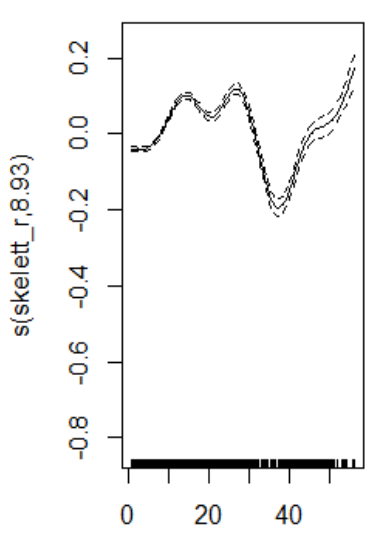
Approximate significance of smooth terms:

| | edf | Ref.df | F | p-value |
|---|---|---|---|---|
| s(RW,HW) | 28.893 | 29.000 | 501.81 | <2e-16*** |
| s(slope) | 8.676 | 8.676 | 55.38 | <2e-16 *** |
| s(skelett_r) | 8.931 | 8.931 | 164.11 | <2e-16 *** |
| s(TRD_r) | 8.854 | 8.854 | 193.52 | <2e-16 *** |
| s(schluff_r) | 8.727 | 8.727 | 45.12 | <2e-16 *** |

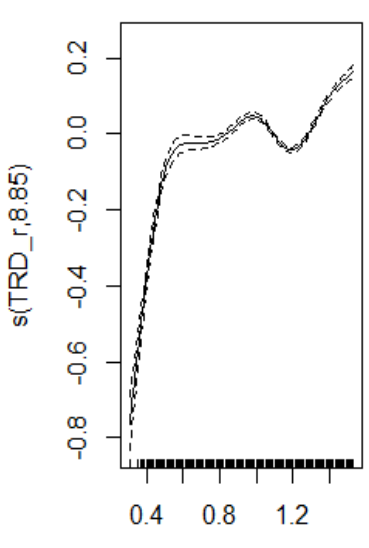
summary(Fit.GAMM.8$gam)
Family: Gamma
Link function: log
Scale est. = 0.12604
n = 100000



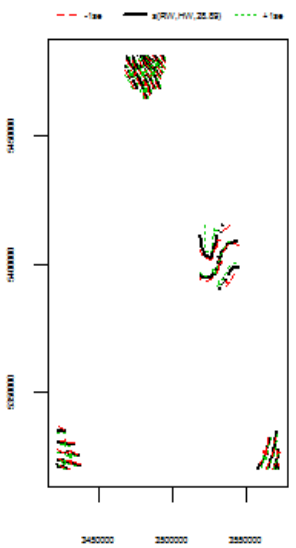

silt content



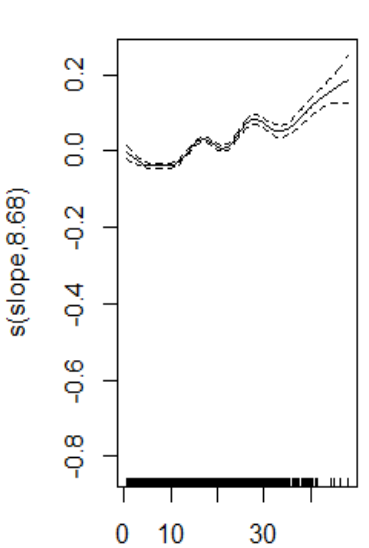

gravel content



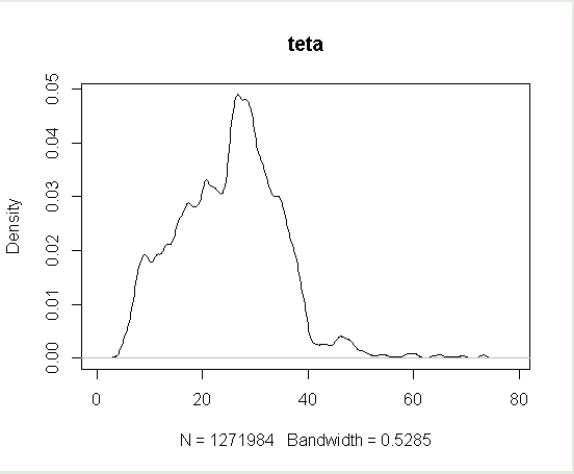

bulk density

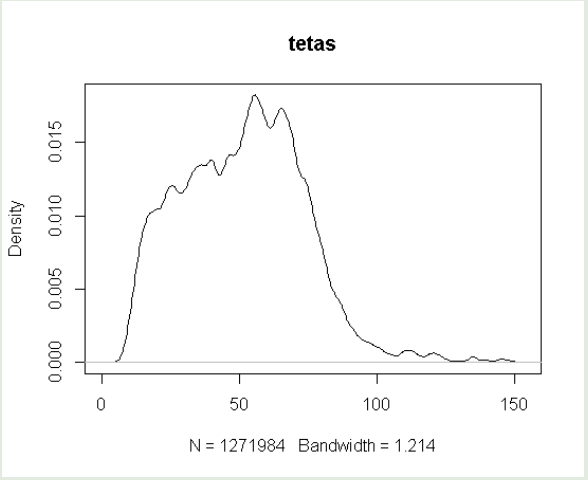


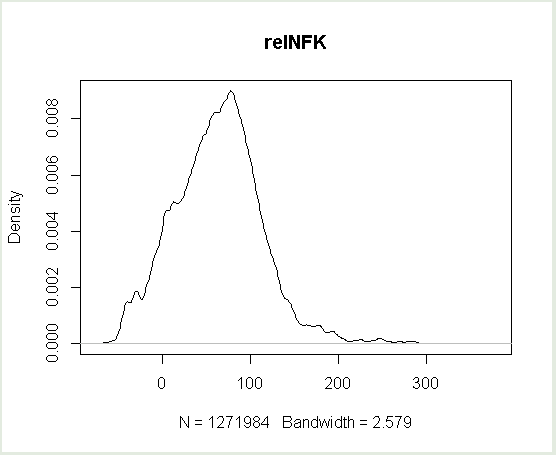
easting RW




slope [ ]


teta
Density
0.00
0.01
0.02
0.03
0.04
0.05
0
20
40
60
80
N = 1271984 Bandwidth = 0.5285


tetas
Density
0.000
0.005
0.010
0.015
0
50
100
150
N = 1271984 Bandwidth = 1.214


relNFK
Density
0.000
0.002
0.004
0.006
0.008
0
100
200
300
N = 1271984 Bandwidth = 2.579