

# Report 1: Soil water characteristic curves (SWCC).

Salvador Garcia, s1655274

March 12th 2017

## 1 Description of the problem

Soil water retention data are useful for constructing soil water characteristic curves, a graphical representation which describes how soil stores and releases water. Usually, for each collected sample, the data is obtained experimentally after applying different tensions levels,  $x$ , and observing  $y$ , the proportion of water (soil water content) in the soil sample. Tension levels are measured in units of atmosphere ( $atm$ ) pressure.

The observed proportion of water is limited from below by the residual soil water content and limited from above by the saturated proportion of water. The residual soil water content,  $y_{\min}$ , is the amount of water that cannot be drained from the soil even at high tension values and it is defined as the soil water content at  $x = 15atm$ . The saturated soil water content,  $y_{\max}$ , is the maximum water content able to be held in the soil before drainage takes place and it is defined as the water content at  $0atm$ . These two characteristics of the soil vary among soil samples and are different across soil depths. Therefore, for each sample, the observed data measured over several tension levels is rescaled using the residual water content value  $y_{\min}$  and the saturated water content value  $y_{\max}$  measured for this sample, so that the observed proportions of water are between 0 and 1:

$$\tilde{y} = (y - y_{\min}) / (y_{\max} - y_{\min}).$$

## Nonlinear models for conditional mean

Several analytical nonlinear expressions have been proposed in the literature for representing the soil water content  $E(\tilde{y} \mid x) = \eta(x, \beta)$ . Among the most widely used SWCCs expressions are Gardner's and van Genuchten's.

Gardner's expression is given by

$$\eta(x, \beta) = \frac{1}{1 + (\beta_1 x)^{\beta_2}}, \quad (1)$$

where  $\beta_1 > 0$  is related to the inverse air entry value of the soil and  $\beta_2 > 0$  is related to the slope of the SWCC, and  $\beta = (\beta_1, \beta_2)$ .

The van Genuchten expression is given by

$$\eta(x, \beta) = \frac{1}{\left[1 + (\beta_1 x)^{\beta_2}\right]^{\beta_3}}, \quad (2)$$

where  $\beta_1 > 0$  is related to the inverse of the air entry value,  $\beta_2, \beta_3 > 0$ , and  $\beta = (\beta_1, \beta_2, \beta_3)$ .

## Data collection

Soil water retention data is usually measured by collecting soil samples along a given region and, at each collection site, soil samples are obtained for different soil depths and the measurements on the response variable are made in replicates. The reason why soil samples are collected from different depths at the same collection site is because soil water retention at different layers of the same soil is subjected to different hydraulic properties and dynamics. Thus, there is evidence of a possible variability in the data due to an unobserved effect of soil depth.

The data set consists of soil samples collected at depths  $0 - 5cm$ ,  $15 - 20cm$  and  $60 - 65cm$ , with 3 replicate samples at each depth. For each sample, the water content was measured at eight tension levels ranging from 0.01 to 10 *atm*,  $x = (0.01, 0.03, 0.06, 0.10, 0.33, 0.80, 4, 10)$ . For each sample, the data is rescaled using the residual soil water content and limited from above by the saturated proportion of water for that sample (i.e. the observations at  $x = 0$  and at  $x = 15$ ) as described in Section “Description”. In particular, this implies that the rescaled observations  $\tilde{y}_{kji}$  are between 0 and 1,  $k = 1, \dots, 8$ ,  $j = 1, 2, 3$ ,  $i = 1, 2, 3$ .

The data is in the file DataSWCC.odc. The depths are coded as 1 for  $0 - 5cm$ , as 2 for  $15 - 20cm$ , and as 3 for  $60 - 65cm$ .

## 2 Comments about this report.

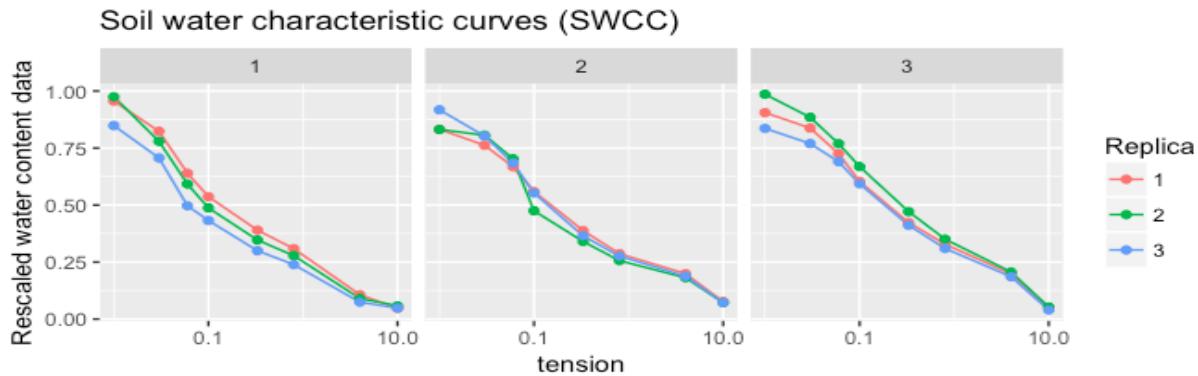
For this report, I decided to use Stan language for all the bayesian computations. The main reason is that I would like to work with bayesian data analysis in a future; in consequence, I would like to try another interface that allows me to save the models and the simulations in *R*. Stan is in an active development status, with a prominent group helping to increase and improve the language. I tried to do my best in this report, but some functions from the BUGS language have not a straightforward equivalence in Stan, so I substitute some of them with formulations in the Stan context (references are given).

At the end, I could arrive to some practical conclusions about Stan Language in comparison with BUGS, (both have advantages and disadvantages):

- Is a very versatile language that have hundreds of predefined functions that allow us to vectorize the code and make it easier to code
- The model is compiled in c++, and if you just want to make more simulations no need to recompile (making it faster)
- Used through RStudio, explanations of errors, code completion and highlighting is available.
- You can save and load the models from an R object
- It allows you to use ggplot directly for the posterior plots (so, you can change the titles, subtitles, etc)
- You have to declare the bounds of the data and parameters. In fact, the manual says that “For data and transformed parameters, the bounds are used for error checking. For parameters, the constraints are critical to sampling ...”
- It allows you to use gradient-based optimization (NUTS, HMC)
- It allows to run different chains at the same time (parallel)
- Does not compute the directly the DIC. In fact, Andrew Gelman (one of the authors), suggest in a Stan-group discussion to compute the WAIC of the LOO (Leave-one-out cross-validation).

### 3 Likelihood for the rescaled water content data.

I'll select a beta distribution as the likelihood for the rescaled water content data for three reasons. The first is that this distribution has a domain over the interval [0,1], making it ideal for the problem. The second is that the formulation of the problem is closely related to the generalized linear models, in particular the beta regression, which is well documented in the paper (Cribari-Neto & Zeileis, 2009). The idea behind the beta regression is to find a parametrization of the original beta distribution in terms of a mean and precision (or dispersion) parameters (instead of two positive shape parameters). Then, the mean parameter is modelled via a link function and some regressors. On the other hand, the precision parameter, as stated in the paper, "may be constant or depend on a set of regressions through another link function". The third reason is that a beta distribution gives a lot of flexibility in the range [0,1]. The SWCC for the rescaled given data are plotted in the below graph (The maximum water content before the drainage takes place is at 0atm):



#### Re-parameterization in terms of mean and precision

First, a beta distribution parameterized in terms of shape parameters  $\alpha$  and  $\beta$  has as mean  $\mu$  and variance  $\sigma^2$ :

$$\mu = \frac{\alpha}{\alpha + \beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

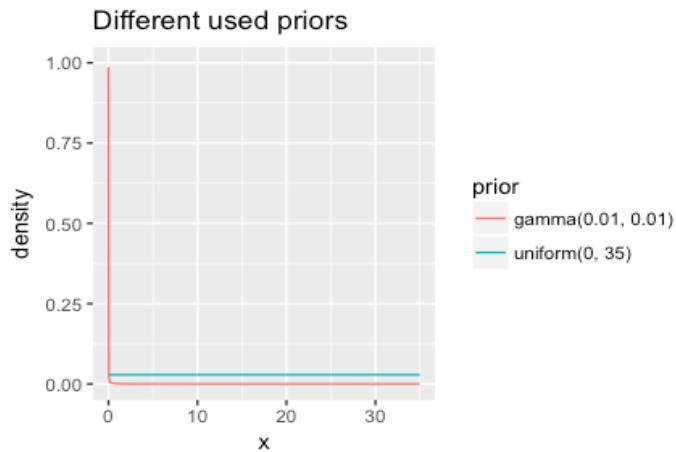
Defining  $\phi = \alpha + \beta$ , we can write  $\alpha = \phi\mu$  and  $\beta = \phi(1 - \mu)$ . Now, the variance  $\sigma^2$  can be written in terms of  $\mu$  and  $\phi$ . With this re-parameterization, we can write that  $y \sim B(\alpha, \beta)$  is equivalent to  $y \sim B(\mu, \phi)$  with  $E(y) = \mu$  and  $Var(y) = \frac{\mu(1-\mu)}{\phi+1}$ . As  $\phi$  is the sum of two positive numbers, we have that  $\phi > 0$  and that  $\mu$  is in the interval (0,1). In his paper, (Cribari-Neto & Zeileis, 2009) justifies that  $\phi$  can be seen as a precision parameter. This is because, for a fixed  $\mu$ , if we increase  $\phi$ , the variance of  $y$  will decrease.

More discussion should be done about the parameter  $\phi$ . It is easier to assume that this parameter is constant for all the observations, but in general this is not true. In fact, as (Cribari-Neto & Zeileis, 2009) states, *data in the interval [0,1] such as rates and proportions are usually heteroscedastic*. But, because of the objective of this project, a constant  $\phi$  will be considered. Also, a consideration about the parameter  $\mu$  should be made. The question about the parameter  $\mu$ , is: should we consider one parameter per observation ( $\mu_{ijk}$ ) or is it better to consider  $\mu_{kj}$  or  $\mu_k$ ? (i for replica, j for depth soil and k for the levels of the regressors x). Different experiments will be done in this report. Assuming that  $\phi$  is common for every observation (i.e. Homoscedasticity), we can write the basic model of the report as:

$$\tilde{y}_{kji} | \mu_{ijk}, \phi \sim B(\phi(\mu_{ijk}), \phi(1 - \mu_{ijk})) \quad (3)$$

## 4 Priors for the parameters $\phi$ and $\mu_{ijk}$

In this problem we do not have prior information for the apriori distribution, so uninformative prior distributions will be used. The restrictions of the problem states that  $\beta_1, \beta_2$  and  $\beta_3 > 0$ . Then, as stated in (Costa Mota Paraíba, Bochkina, & Diniz, 2017), “the parameters  $\beta_1, \beta_2$  and  $\beta_3$  have a broad interpretation of scale (possibly under a re-parameterization)”. Then, a prior gamma distribution with small shape and scale parameters can be used. On the other hand, the  $\phi$  parameter can be considered as a precision parameter, then an uninformative gamma distribution with shape  $\alpha = 0.01$  and rate  $\beta = 0.01$  will be used.



### Including soil depth variability

The first model is a pooled effects model. After this, to include the variability associated with the the soil depth, an independent effects and a hierarchical model will be tried. In the independent effects and hierarchical model, each  $\beta_1, \beta_2$  and  $\beta_3$  is a three dimensional vector, with one dimension for each soil depth. This way, in total nine  $\beta$ 's parameters will be estimated. In the case of the hierarchical model, the  $\beta$ 's associated at the same soil depth will have common uninformative hyperparameters (with distribution  $\text{gamma}(0.01, 0.01)$ ). Finally, for the last two approaches,  $\mu_{kj}$  will be considered (i.e. it will depend on the regressor value and on the soil depth).

### Number of iterations, chains, burning and thinning

Stan is relatively faster than Bugs. For this reason, more iterations can be done in the same time. In this report, two chains were computed with 150,000 iterations each one. To reduce the autocorrelation, a thinning of 10 was done. Also, by default Stan takes a burning of the 50% of the iterations, so at the end 7,500 iterations were taken into account.

### Convergence and autocorrelation

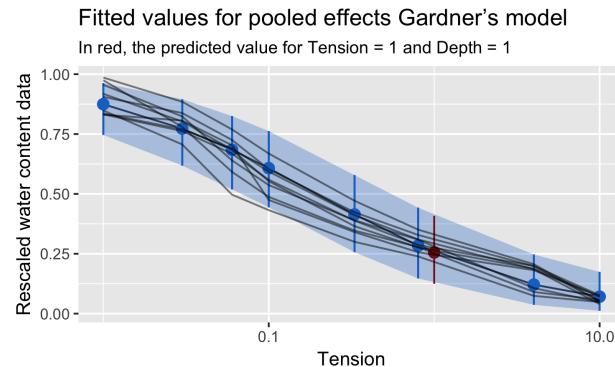
The Stan default output shows the  $\hat{R}$  statistic. As stated in the documentation, we should be careful of values larger than 1.1 or smaller than .9. Also, they suggest to make an histogram of the  $\hat{R}$  statistics, to see if some parameter of the model has convergence problems. The autocorrelation plots also were analyzed (The convergence and autocorrelation plots are in the appendix section).

## 5 Pooled effects for the soil depth variability

For the first model,  $\mu$  will consider the effect of the values in the regressor  $x$  (i.e. will use the parameter  $\mu_k$ , with  $k = 1, \dots, 8$ ), with  $\mu_k = \frac{1}{1+(\beta_1 x_k)^{\beta_2}}$  and each  $\beta_1, \beta_2$  distributed  $Gamma(0.01, 0.01)$  for the Gardners expression. And with  $\mu_k = \frac{1}{[1+(\beta_1 x_k)^{\beta_2}]^{\beta_3}}$  with  $\beta_1, \beta_2, \beta_3$  distributed  $Gamma(0.01, 0.01)$  for the van Genuchten expression.

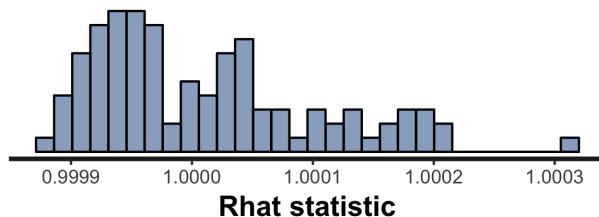
### Gardner's model

#### Fitted Values



The fitted values are shown in the left graph. As this model does not consider the depth levels, the plot just shows the model at different levels of the regressor  $x$  (Tension level). The blue points are the mean value of the model given regressor  $x$  (with a 95% credible interval). The gray lines show the different replicas taken at the depth levels. Finally, the red line and point, represent the predictive distribution given the regressor  $x = 1$ .

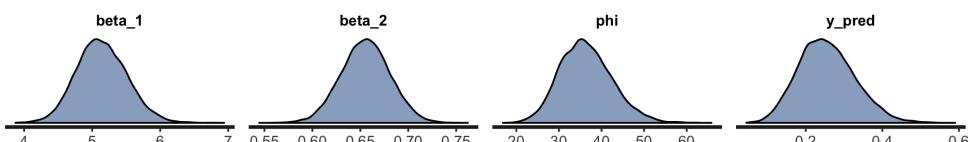
#### Convergence and autocorrelation



Here is the histogram of the  $\hat{R}$  statistic. All values are close to 1: no problem of convergence. In addition, the autocorrelation plots in the Appendix A does not show any problem.

#### Posterior summaries

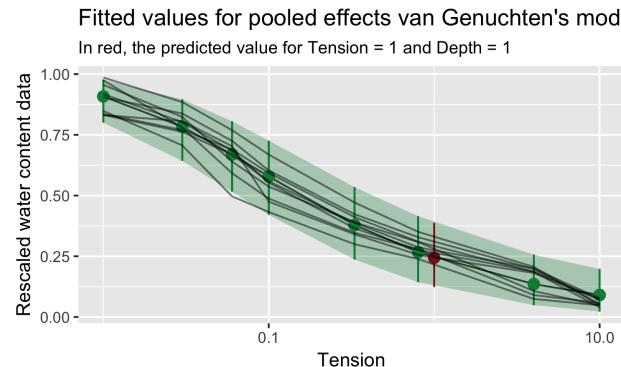
The below graph show the posterior distribution density for the parameters  $\beta_1, \beta_2, \phi$  and  $y_{pred}$ . In addition, the posterior summaries are given.



Inference for Stan model: 1_pooled_2pars.										
2 chains, each with iter=150000; warmup=75000; thin=10;										
post-warmup draws per chain=7500, total post-warmup draws=15000.										
	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
beta_1	5.14	0.00	0.37	4.45	4.89	5.13	5.39	5.90	15000	1
beta_2	0.66	0.00	0.03	0.61	0.64	0.66	0.67	0.71	14311	1
phi	36.39	0.05	6.11	25.44	32.01	36.01	40.28	49.35	14573	1
y_pred	0.25	0.00	0.07	0.12	0.20	0.25	0.30	0.41	15000	1

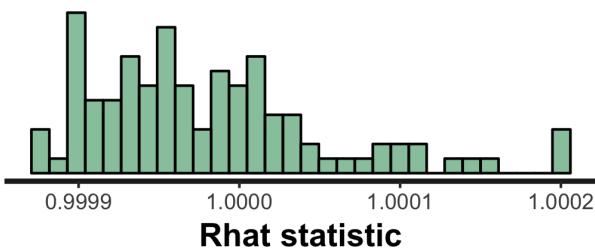
## van Genuchten model

### Fitted Values



The fitted values are shown in the left graph. As this model does not consider the depth levels, the plot just shows the model at different levels of the regressor  $x$  (Tension level). The green points are the mean value of the model given regressor  $x$  (with a 95% credible interval). The gray lines show the different replicas taken at the depth levels. Finally, the red line and point, represent the predictive distribution given the regressor  $x$ .

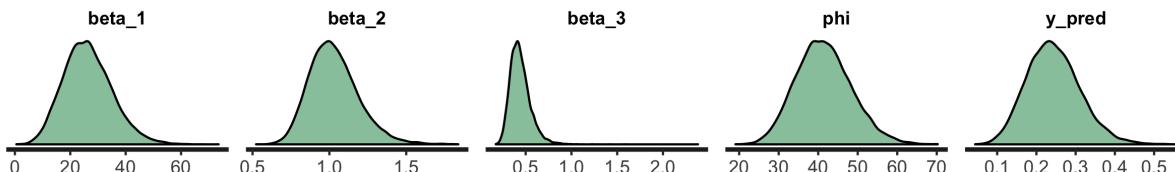
### Convergence and autocorrelation



Here is the histogram of the  $\hat{R}$  statistic. All values are close to 1: no problem of convergence. In addition, the autocorrelation plots in the Appendix A does not show any problem.

### Posterior summaries

The below graph show the posterior distribution density for the parameters  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ ,  $\phi$  and  $y_{pred}$ . In addition, the posterior summaries are given.



```
Inference for Stan model: 1_pooled_3pars.
2 chains, each with iter=150000; warmup=75000; thin=10;
post-warmup draws per chain=7500, total post-warmup draws=15000.

          mean   se_mean    sd   2.5%   25%   50%   75% 97.5% n_eff Rhat
beta_1    26.36     0.07   8.75  11.06  20.22  25.81  31.90  45.21 15000    1
beta_2     1.03     0.00   0.16    0.76   0.92   1.01   1.12   1.39 14996    1
beta_3     0.44     0.00   0.11    0.27   0.36   0.43   0.50   0.70 15000    1
phi        41.44     0.06   6.98  28.78  36.55  41.15  45.90  56.21 14270    1
y_pred     0.24     0.00   0.07   0.12   0.20   0.24   0.29   0.39 15000    1
```

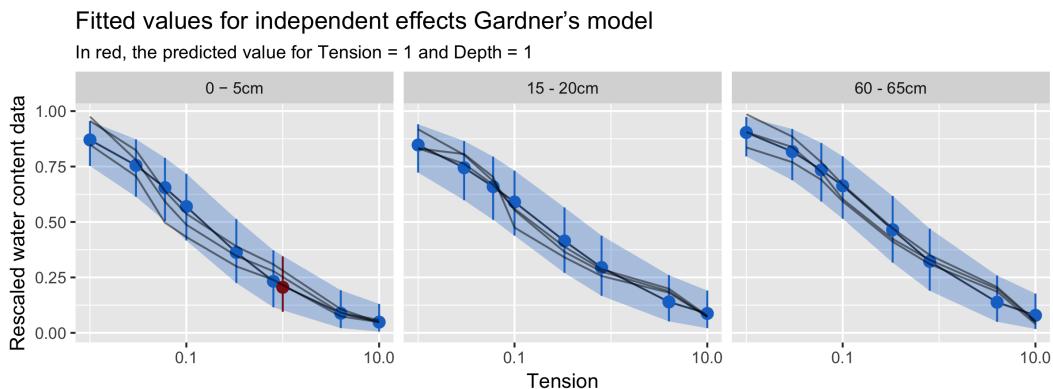
## 6 Independent effects for the soil depth variability

For this second model,  $\mu$  will just consider the effect of the values of the regressor  $x$  and the effect of soil depth as a hierarchical effect (i.e.  $\mu_{kj}$ , with  $k = 1, \dots, 8$  and  $j = 1, 2, 3$ ), with  $\mu_{kj} = \frac{1}{1+(\beta_{1j}x_k)^{\beta_{2j}}}$  and each  $\beta_{1j}, \beta_{2j}$  distributed  $Gamma(0.01, 0.01)$  for  $j = 1, 2, 3$  for the Gardners expression. And with  $\mu_{kj} = \frac{1}{[1+(\beta_{1j}x_k)^{\beta_{2j}}]^{\beta_{3j}}}$  for  $j = 1, 2, 3$  and each  $\beta_{1j}, \beta_{2j}, \beta_{3j}$  distributed  $Gamma(0.01, 0.01)$  for  $j = 1, 2, 3$  for the van Genuchten expression.

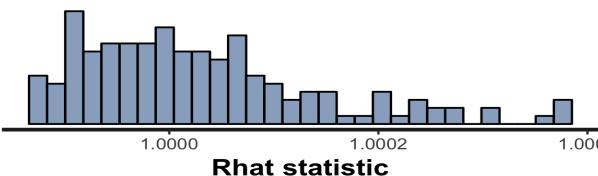
### Gardner's model

#### Fitted Values

The fitted values are shown in the below graph. Now, the model considers the depth levels and there is one plot for each one. The blue points are the mean value of the model given regressor  $x$  (with a 95% credible interval). The gray lines show the different replicas taken at the depth levels. Finally, the red line and point, represent the predictive distribution given the regresor  $x = 1, depth = 1$ .



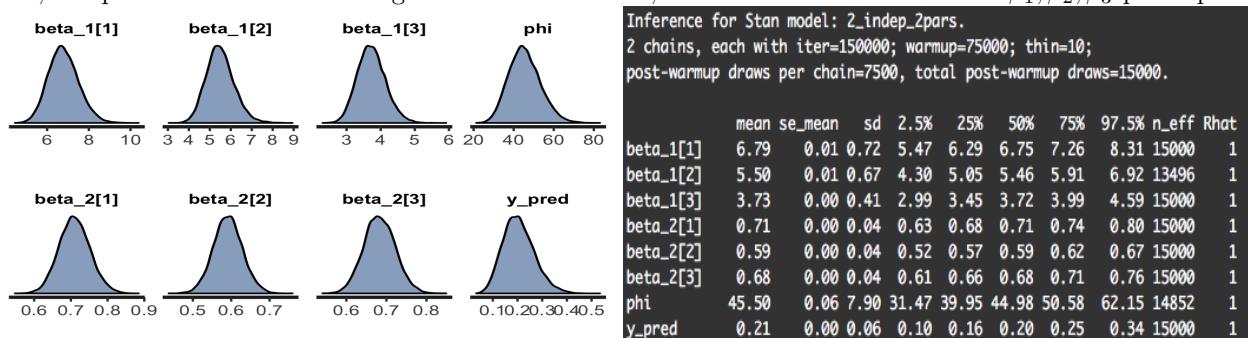
### Convergence and autocorrelation



Here is the histogram of the  $\hat{R}$  statistic. All values are close to 1: no problem of convergence. In addition, the autocorrelation plots in the Appendix A does not show any problem.

### Posterior summaries

The below graph show the posterior distribution density for the parameters  $\beta_1, \beta_2, \phi$  and  $y_{pred}$ . In addition, the posterior summaries are given. As can be seen, now there is one estimate of  $\beta_1, \beta_2, \beta_3$  per depth.



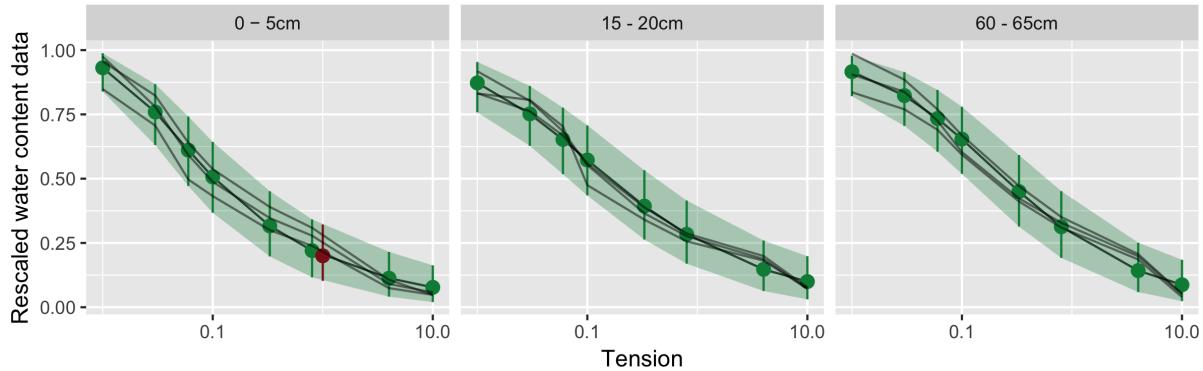
## van Genuchten model

### Fitted Values

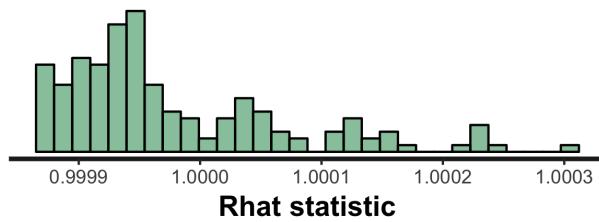
The fitted values are shown in the below graph. Now, the model considers the depth levels and there is one plot for each one. The green points are the mean value of the model given regressor  $x$  (with a 95% credible interval). The gray lines show the different replicas taken at the depth levels. Finally, the red line and point, represent the predictive distribution given the regresor  $x = 1, \text{depth} = 1$ .

Fitted values for the independent effects van Genuchten's model

In red, the predicted value for Tension = 1 and Depth = 1



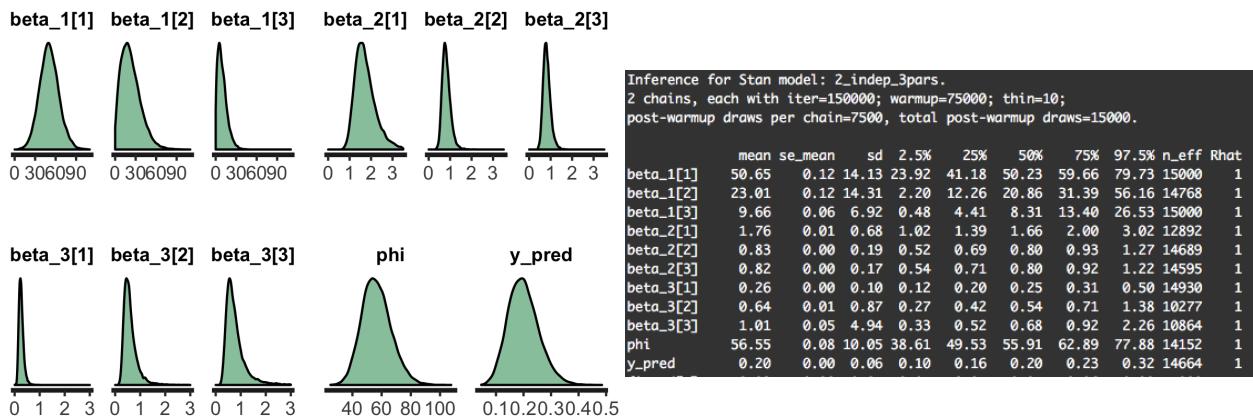
### Convergence and autocorrelation



Here is the histogram of the  $\hat{R}$  statistic. All values are close to 1: no problem of convergence. In addition, the autocorrelation plots in the Appendix A does not show any problem.

### Posterior summaries

The below graph show the posterior distribution density for the parameters  $\beta_1, \beta_2, \beta_3, \phi$  and  $y_{pred}$ . In addition, the posterior summaries are given. As can be seen, now there is one estimate of  $\beta_1, \beta_2, \beta_3$  per depth.



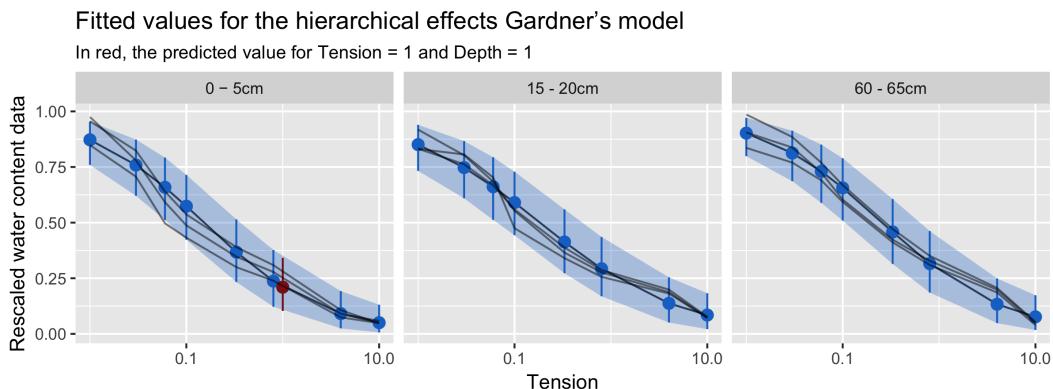
## 7 Hierarchical modelling to include soil depth variability

For the third model,  $\mu$  will just consider the effect of the values of the regressor  $x$  and the effect of soil depth variability as an independent effect (i.e.  $\mu_{kj}$ , with  $k = 1, \dots, 8$  and  $j = 1, 2, 3$ ), with  $\mu_{kj} = \frac{1}{1+(\beta_{1j}x_k)^{\beta_{2j}}}$  and each  $\beta_{1j}, \beta_{2j}$  distributed  $Gamma(0.01, 0.01)$  for  $j = 1, 2, 3$  for the Gardners expression. And with  $\mu_{kj} = \frac{1}{[1+(\beta_{1j}x_k)^{\beta_{2j}}]^{\beta_{3j}}}$  for  $j = 1, 2, 3$  each  $\beta_{1j}, \beta_{2j}, \beta_{3j}$  distributed  $Gamma(0.01, 0.01)$  for  $j = 1, 2, 3$  for the van Genuchten expression.

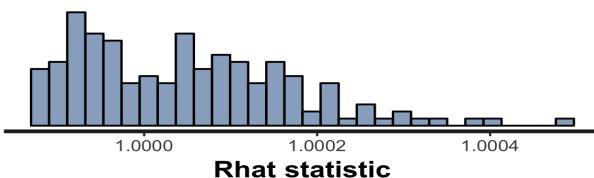
### Gardner's model

#### Fitted Values

The fitted values are shown in the below graph. Now, the model considers the depth levels and there is one plot for each one. The blue points are the mean value of the model given regressor  $x$  (with a 95% credible interval). The gray lines show the different replicas taken at the depth levels. Finally, the red line and point, represent the predictive distribution given the regresor  $x = 1, depth = 1$ .



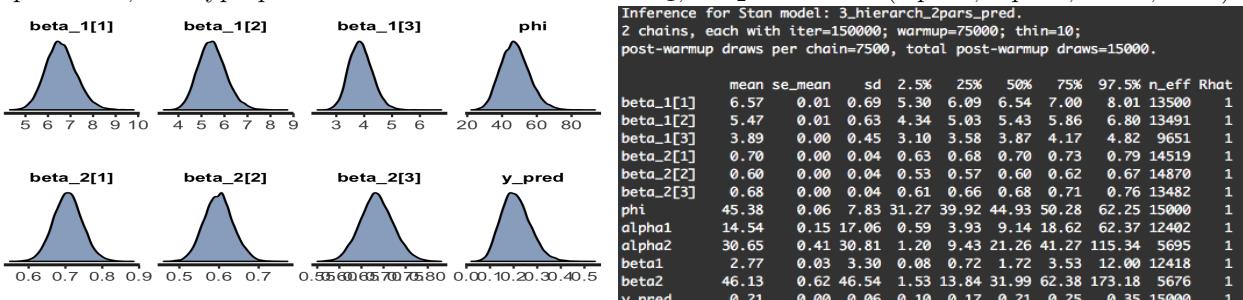
### Convergence and autocorrelation



Here is the histogram of the  $\hat{R}$  statistic. All values are close to 1: no problem of convergence. In addition, the autocorrelation plots in the Appendix A does not show any problem.

### Posterior summaries

The below graph show the posterior distribution density for the parameters  $\beta_1, \beta_2, \phi$  and  $y_{pred}$ . In addition, the posterior summaries are given. As can be seen, now there is one estimate of  $\beta_1, \beta_2$  per depth. Also, the hyperparameters estimates for each  $\beta_1, \beta_2$  are shown ( $\alpha_1, \alpha_2, \beta_1, \beta_2$ ).



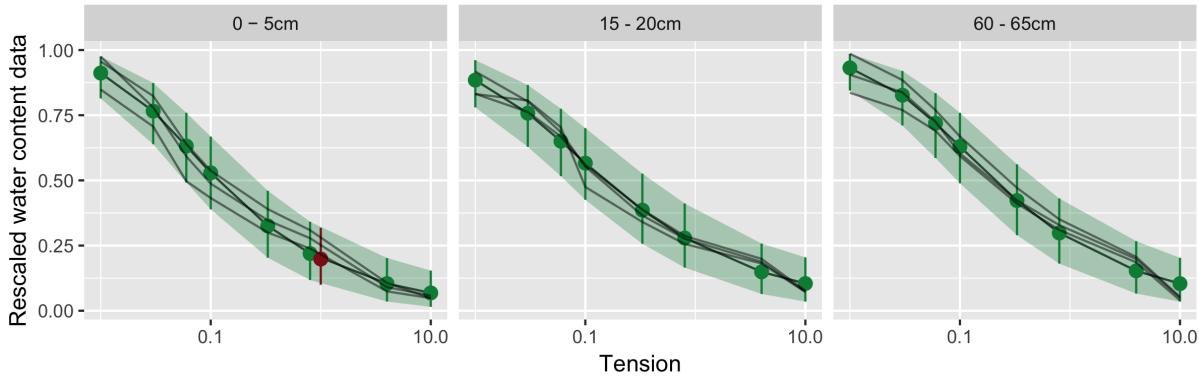
## van Genuchten model

### Fitted Values

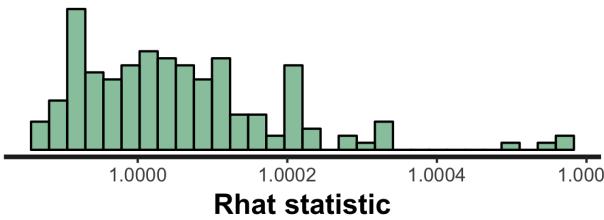
The fitted values are shown in the below graph. Now, the model considers the depth levels and there is one plot for each one. The green points are the mean value of the model given regressor  $x$  (with a 95% credible interval). The gray lines show the different replicas taken at the depth levels. Finally, the red line and point, represent the predictive distribution given the regresor  $x = 1, \text{depth} = 1$ .

Fitted values for the hierarchical effects van Genuchten's model

In red, the predicted value for Tension = 1 and Depth = 1



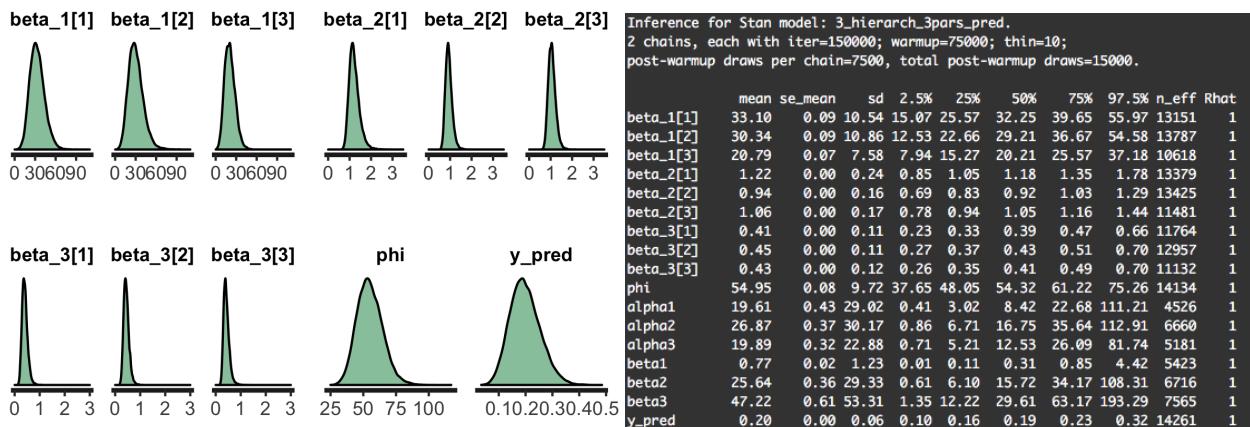
### Convergence and autocorrelation



Here is the histogram of the  $\hat{R}$  statistic. All values are close to 1: no problem of convergence. In addition, the autocorrelation plots in the Appendix A does not show any problem.

### Posterior summaries

The below graph show the posterior distribution density for the parameters  $\beta_1, \beta_2, \beta_3, \phi$  and  $y_{pred}$ . In addition, the posterior summaries are given. As can be seen, now there is one estimate of  $\beta_1, \beta_2, \beta_3$  per depth. Also, the hyperparameters estimates for each  $\beta_1, \beta_2, \beta_3$  are shown ( $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$ ).



## 8 Model selection

Stan does not provide the DIC directly. The reasons about this decision are written on the STAN-users mailing group <sup>1</sup>. There is an open discussion about the benefits of different measures, but until now, they prefer other measures like WAIC and LOO (Leave-One-Out Cross-Validation).

As these measures require to compute the log point-wise predictive density, a piece of code is included in each stan file. After that, following the example of (Vehtari, Gelman, & Gabry, 2016b), I compute the WAIC for each model (using the package `loo` (Vehtari, Gelman, & Gabry, 2016a)):

Table 1: Comparison of WAIC

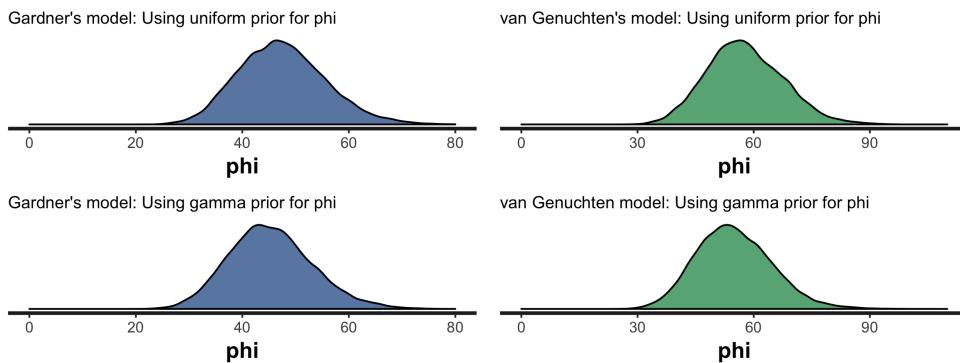
Model	Estimate	Estimate
	$\text{elpd}_{\text{waic}}$	$p_{\text{waic}}$
Pooled-effects Gardner model	95.7	3.7
Pooled-effects van Genuchten model	100.0	4.9
Independent-effects Gardner model	101.0	8.1
Independent-effects van Genuchten model	105.3	13.5
Hierarchical-effects Gardner model	101.1	7.3
Hierarchical-effects van Genuchten model	106.6	10.3

We have the equality  $\widehat{\text{elpd}}_{\text{waic}} = \widehat{\text{lpd}} - \widehat{p}_{\text{waic}}$ , with  $\widehat{\text{lpd}}$  a “measure of predictive accuracy”, and  $\widehat{p}_{\text{waic}}$  the “estimated effective number of parameters.”. Then, for these model and the data provided, the Hierarchical-effects van Genuchten model is the preferred.

## 9 Prior sensitivity analysis

### Change in the prior of $\phi$

To analyze the prior sensitivity for this parameter, a gamma prior distribution with shape and scale parameters equal to 0.01, and a uniform distribution between in the interval [0, 100] for the parameter  $\phi$  will be compared:

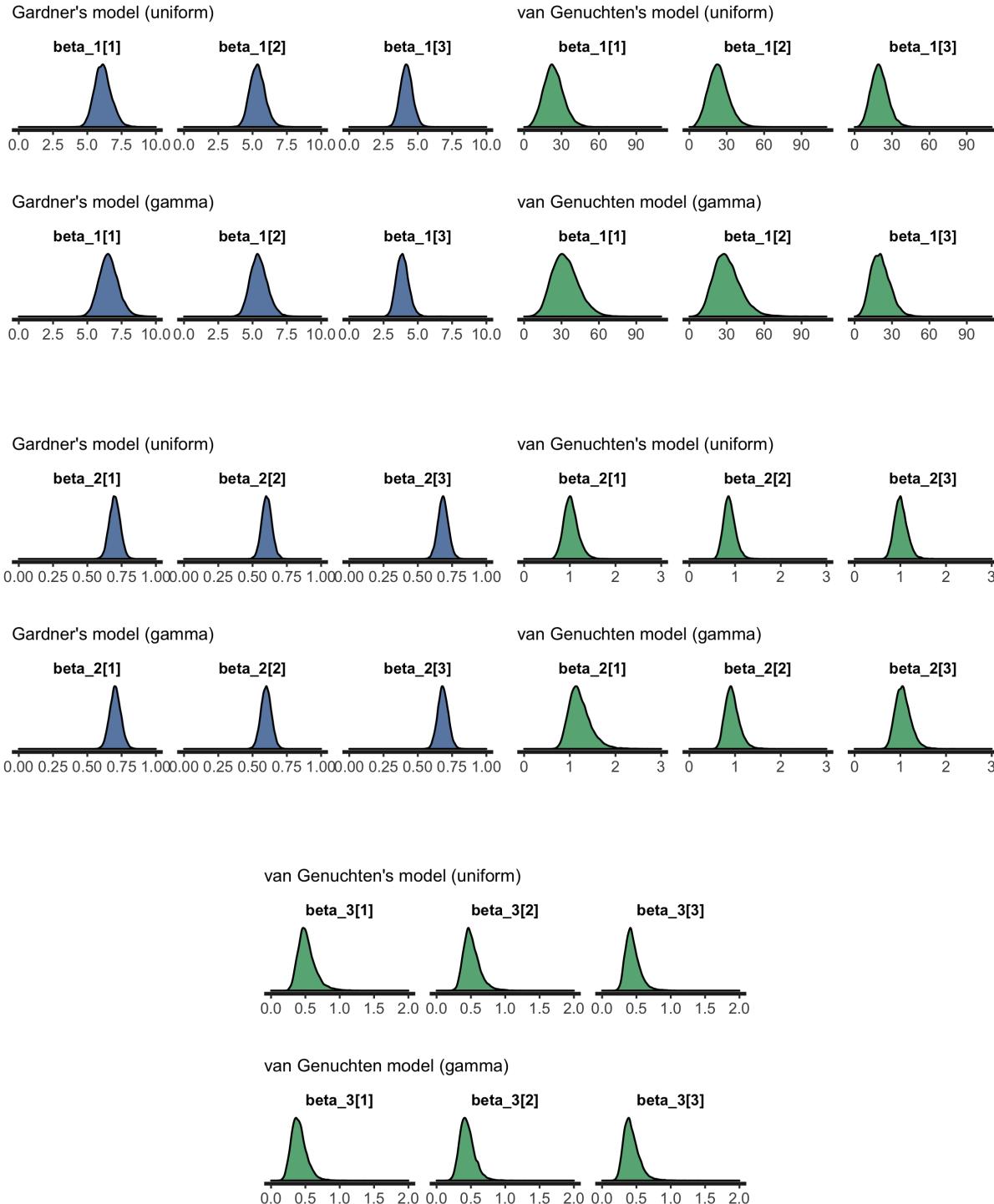


Observing the above graph, for this change of priors, there is no important change in the posterior distribution of  $\phi$ .

<sup>1</sup> <https://groups.google.com/forum/#topic/stan-users/SbdA47xC4Fw>

**Change in the prior of the hyperparameters of each  $\beta_i$ ,  $i = 1, 2, 3$ .**

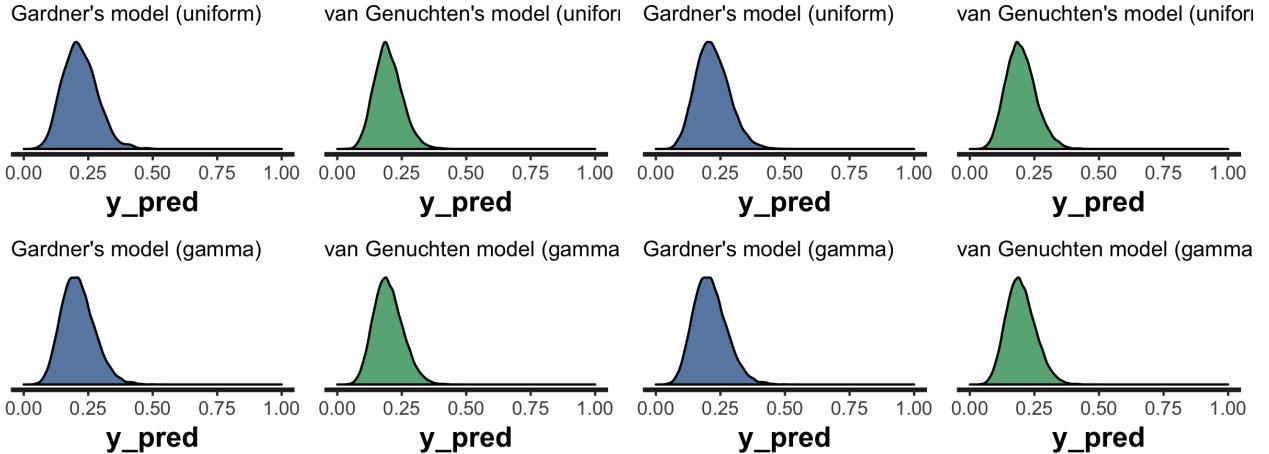
Now, the selected model uses prior distributions in the hyperparameters of each  $\beta$ . In the first experiment, these a priori distributions were uninformative gamma distributions with parameters  $\alpha = 0.01, \beta = 0.01$ . Now, uniform distributions from 0 to 100 were used.



In these graphs we can see a slightly movement of the posterior distributions, but the shape and domain is almost the same for all of them.

## Change in the predictive posterior distribution

To assess if there is a change in the predictive posterior distribution when specifying distinct prior distributions, the posterior plots are shown in the graphs below:



The four plots in the left shows the posterior plots of  $y_{pred}$  when prior distribution for  $\phi$  is changed. On the other hand, the four right plots shows the posterior plots of  $y_{pred}$  when the prior distribution of the hyperparameters of  $\beta_1, \beta_2, \beta_3$  are changed. As can be seen, the use of both priors provide similar posterior distributions.

## 10 Conclusions and improvements to the model

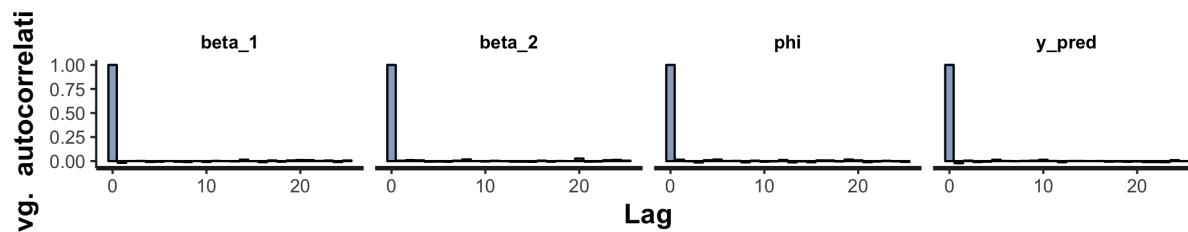
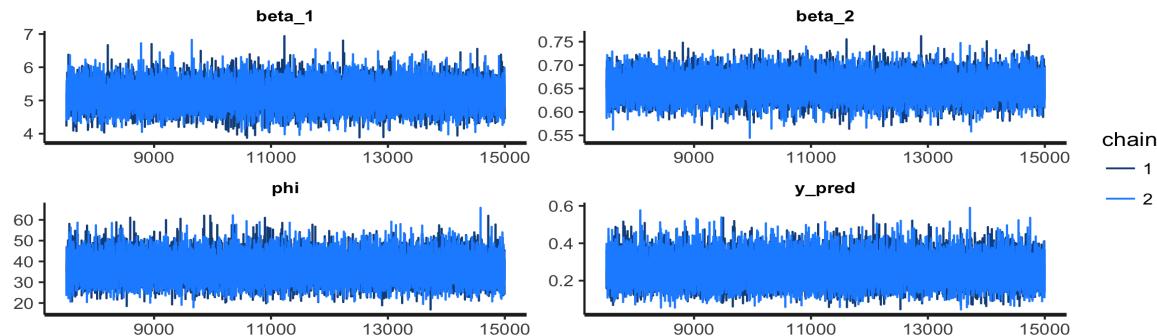
I have two ideas to improve this model. The first is to consider a hierarchical-effect of a subgroup of  $\beta_1, \beta_2, \beta_3$ , not of the three of them (as is made in the last section). As we are penalizing the WAIC with the effective number of parameters, less parameters can yield to a better result. The second idea is to try different models for  $\phi$ . In this report, I treated the  $\phi$  as a constant, but it can depend on the level of the depth (as the  $\beta$ 's parameters), or even as (Cribari-Neto & Zeileis, 2009) says, can have its own model.

As the conclusions, the model that was the better in terms of WAIC was the hierarchical model with three parameters. In the final model there is evidence of variability in the data according to the depth level, for example, the mean of  $\beta_1$  when depth = 1 is 33.10 and the mean of  $\beta_1$  when depth = 3 is 20.79, a difference of more than 12 units (and this number is large compared with the sd of each parameter). In fact, a line of code can be added to the model to estimate the difference between the  $\beta_1$  for different depths. This idea can help us to decide for which variables we should make pooling or consider a hierarchical model (In addition to the WAIC measure).

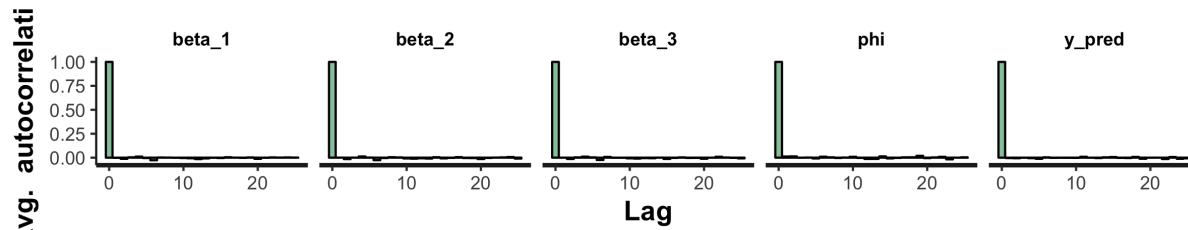
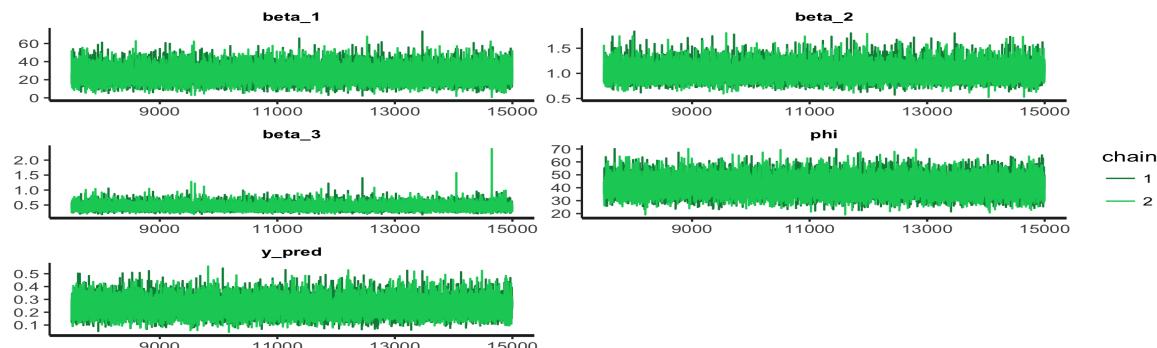
## 11 Appendix

### Appendix A: Pooled model

Gardner's model

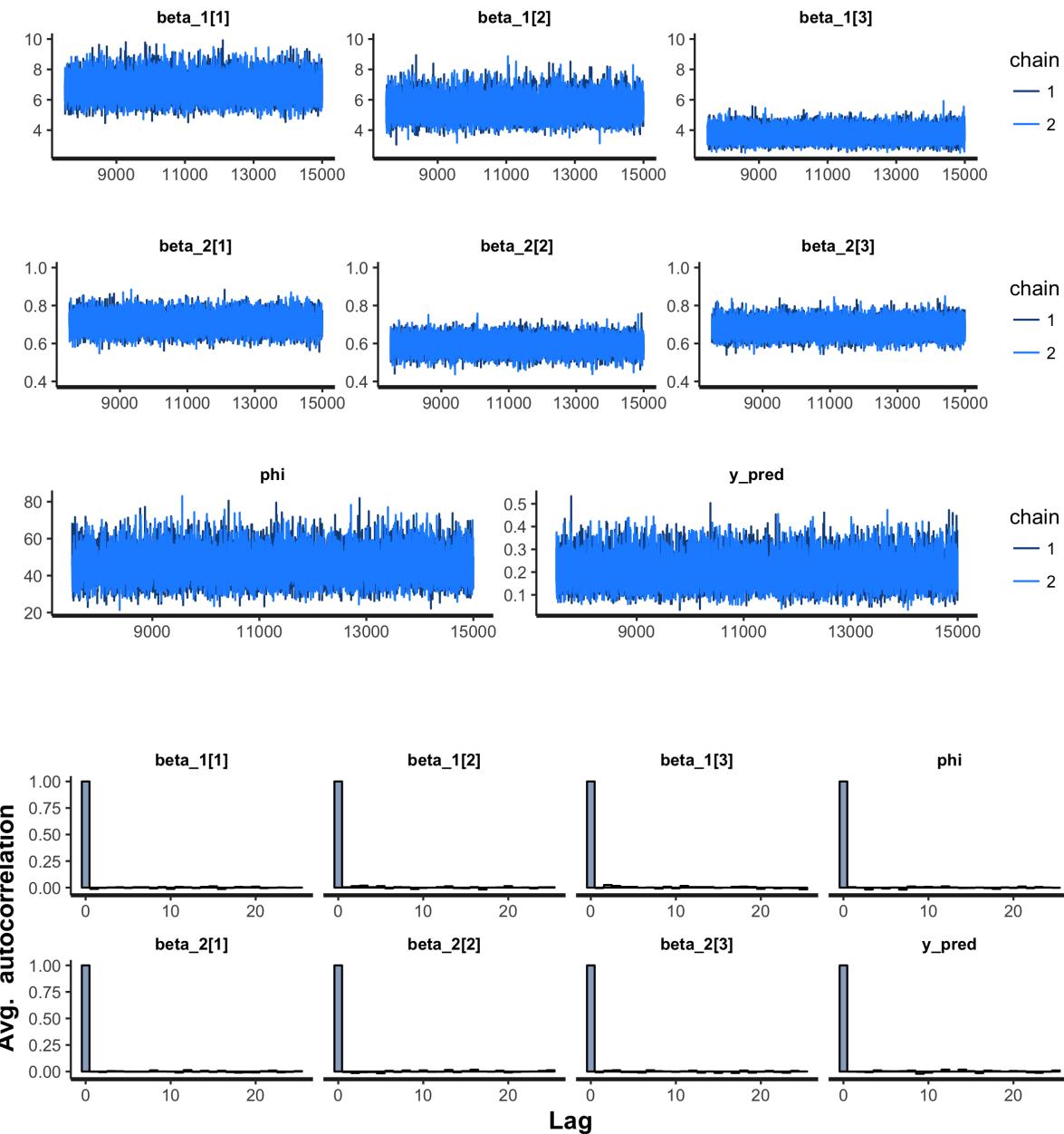


van Genuchten model

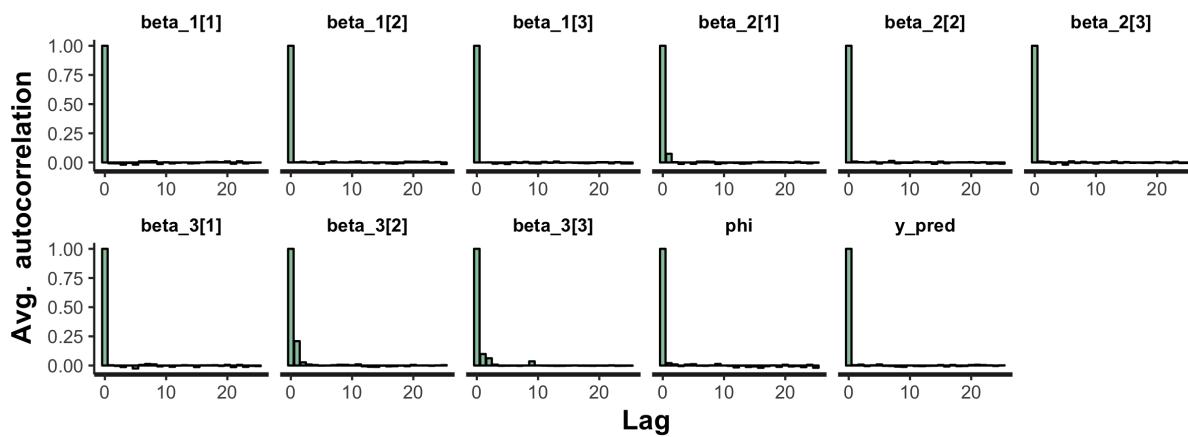
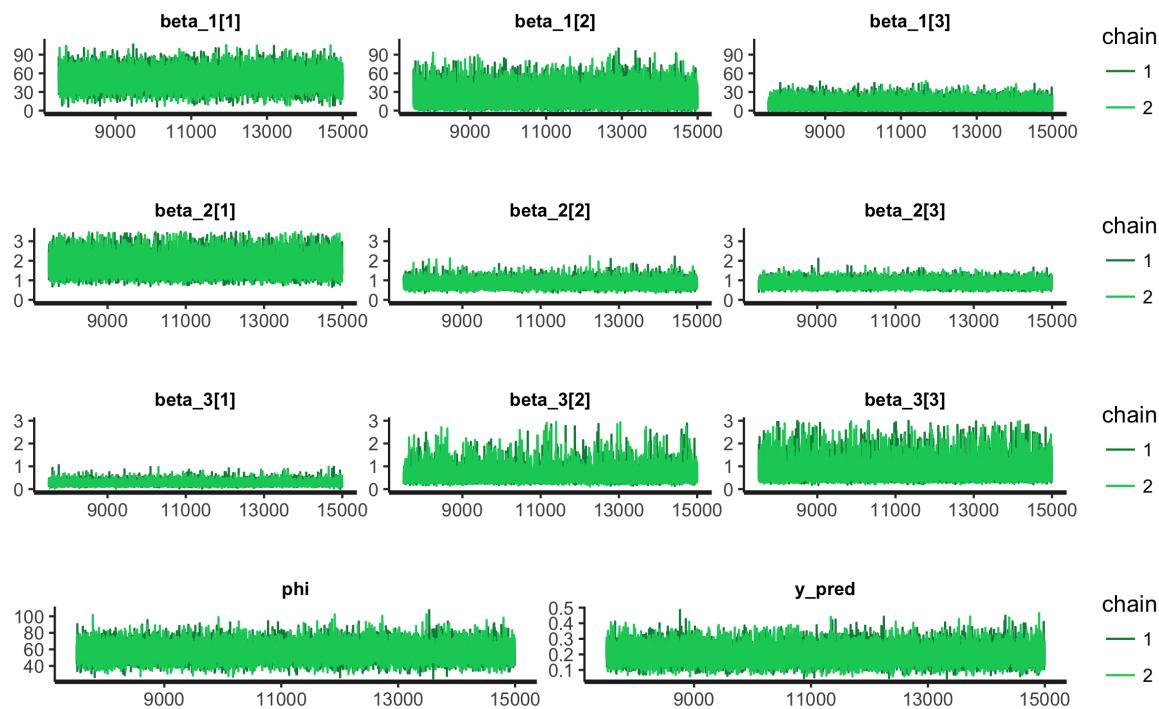


## Appendix B: Independent effects model

### Gardner's model

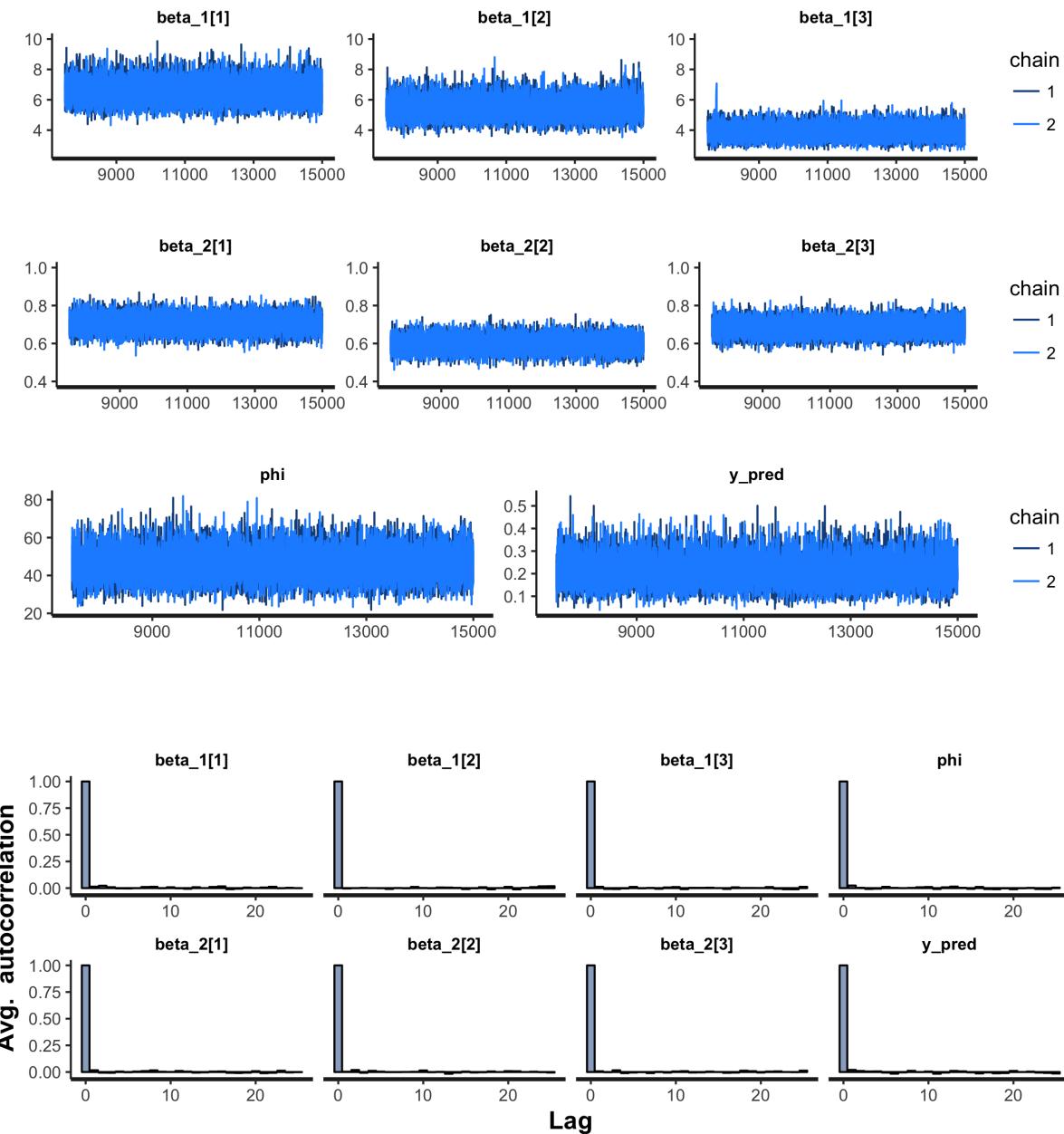


## van Genuchten model

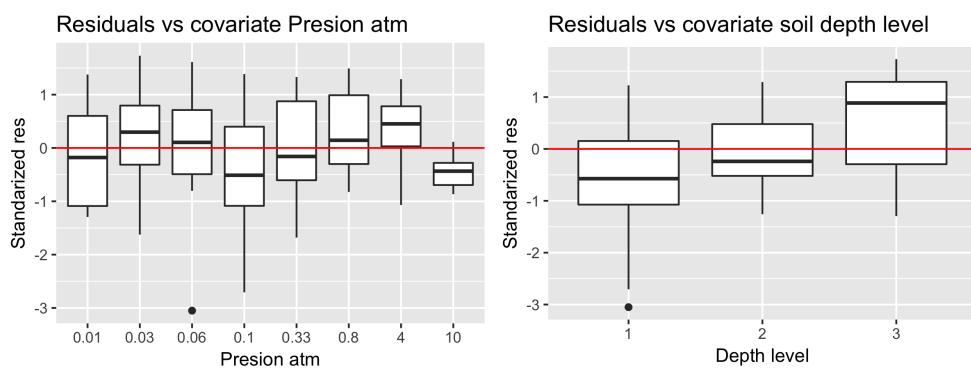
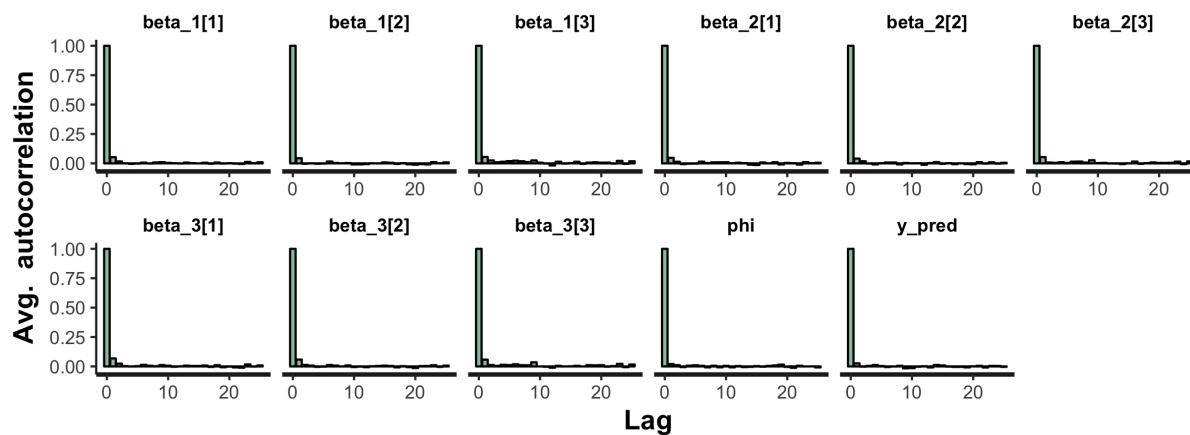
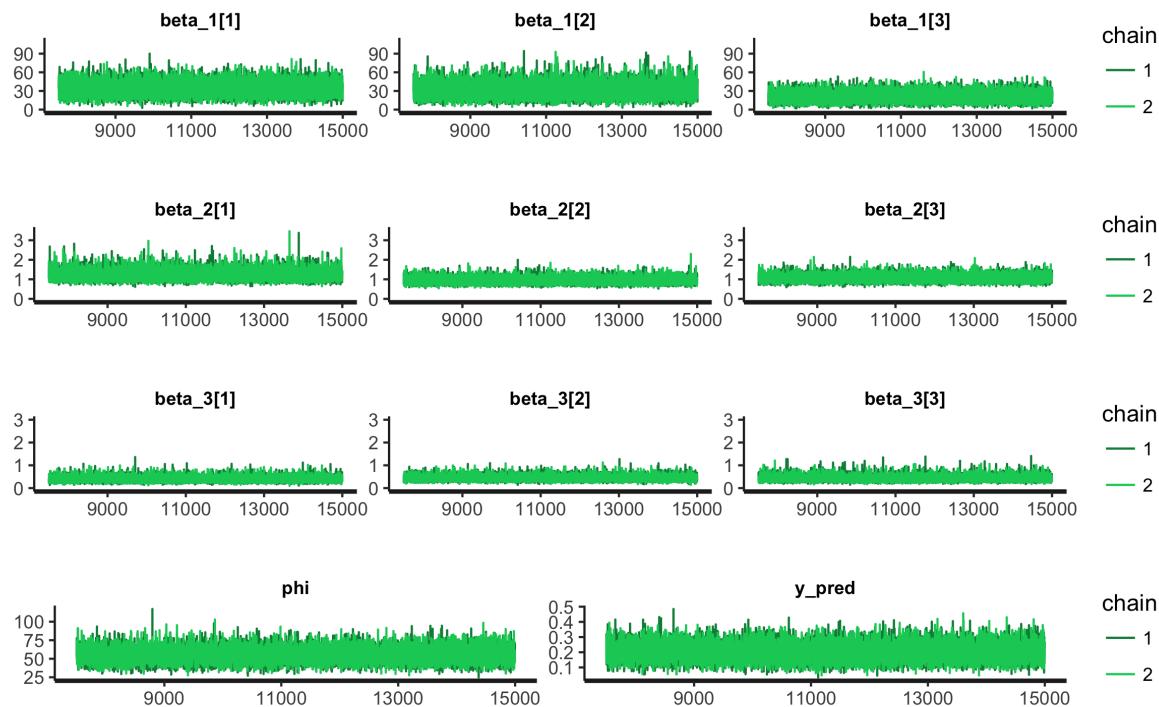


## Appendix C: Hierarchical model

### Gardner's model



## van Genuchten model



## 12 Code

### Hierarchical-effects model - van Genuchten

As the modification for the Gardner's model is trivial, just the van Genuchten code is provided for each of the three models.

```

data {
  int<lower=0> n_tot;
  int<lower=0> Depth[n_tot];
  vector<lower=0>[n_tot] x; # Regressor x, (Atm)
  vector<lower=0>[n_tot] ymax;
  vector<lower=0>[n_tot] ymin;
  vector<lower=0>[n_tot] y;
  int<lower=0> x_ind [n_tot]; # A new variable was added to indicate the level of regressor x (from 1 to 8)
  int<lower=0> x_new; # New data
  int<lower=0> Depth_new; # New data
}
transformed data {
  vector<lower=0, upper=1>[n_tot] y_hat;
  for (i in 1:n_tot)
    y_hat[i] = (y[i]-ymin[i])*pow(ymax[i]-ymin[i], -1); # Reescaled dependent variable
}
parameters {
  real<lower=0> beta_1[3];
  real<lower=0> beta_2[3];
  real<lower=0> beta_3[3];
  real<lower=0> phi;
  real<lower=0> alpha1;
  real<lower=0> alpha2;
  real<lower=0> alpha3;
  real<lower=0> beta1;
  real<lower=0> beta2;
  real<lower=0> beta3;
  real<lower=0> y_pred; # variable of prediction given x_new and depth_new
  real<lower=0, upper= 1> fitted[24]; # fitted values for the model
}
transformed parameters {
  vector<lower=0>[24] mu; # in total 24 mu's (8 levels of regressor x, and 3 levels of depth)
  vector<lower=0>[24] sigma2;
  real<lower=0> mu_pred;
  real<lower=0> sigma2_pred;
  for (j in 0: 2){
    for (i in 1:8){
      mu[j*8+i] = 1*pow((1 + pow(beta_1[Depth[j*24+i]] * x[i],
        beta_2[Depth[j*24+i]]), -beta_3[Depth[j*24+i]]));
      sigma2[j*8+i] = mu[j*8+i] .*(1-mu[j*8+i]).*pow((phi+1), -1);
    }
  }
}

```

```

        }
    }
    mu_pred = 1*pow((1 + pow(beta_1[Depth_new] * x_new, beta_2[Depth_new])), -beta_3[Depth_new]);
    sigma2_pred = mu_pred .*(1-mu_pred).*pow((phi+1), -1);
}
model {
  for ( i in 0:2 ){ # total 3*3*8 observations correspond to depth 1
    for ( j in 0:2 ){ # first 3*8 observations correspond to depth 1
      for ( k in 1:8 ){ # level of regressor x, takes 8 values (same mean for each level)
        y_hat[i*24+j*8+k] ~ beta(mu[i*8+x.ind[k]].*phi, (1-mu[i*8+x.ind[k]]).*phi); # van Genuchten
        model
      }
    }
  }
  for( i in 1:3){ # hierarchical model
    beta_1[i] ~ gamma(alpha1, beta1);
    beta_2[i] ~ gamma(alpha2, beta2);
    beta_3[i] ~ gamma(alpha3, beta3);
  }
  alpha1 ~ gamma(0.01, 0.01);
  alpha2 ~ gamma(0.01, 0.01);
  alpha3 ~ gamma(0.01, 0.01);
  beta1 ~ gamma(0.01, 0.01);
  beta2 ~ gamma(0.01, 0.01);
  beta3 ~ gamma(0.01, 0.01);
  phi ~ gamma(.01, .01);
  y_pred ~ beta(mu_pred*phi, (1-mu_pred)*phi); # prediction distribution
  for (j in 0: 2){
    for (i in 1:8){
      fitted [j*8+i] ~ beta(mu[j*8+i]*phi, (1-mu[j*8+i])*phi); # fitted values
    }
  }
}
generated quantities{
  vector[n_tot] log_li; # pointwise log likelihood
  vector[n_tot] stres; # standarized residuals
  for ( i in 0:2 ){
    for ( j in 0:2 ){
      for ( k in 1:8 ){
        log_li [i*24+j*8+k] = beta_lpdf(y_hat[i*24+j*8+k] | mu[i*8+x.ind[k]].*phi,
          (1-mu[i*8+x.ind[k]]).*phi);
        stres [i*24+j*8+k] = (y_hat[i*24+j*8+k]-mu[i*8+x.ind[k]]*pow(sigma2[i*8+x.ind[k]], -0.5));
      }
    }
  }
}

```

## Independent-effects model - van Genuchten

```

data {
  int<lower=0> n_tot;
  int<lower=0> Depth[n_tot];
  vector<lower=0>[n_tot] x;
  vector<lower=0>[n_tot] ymax;
  vector<lower=0>[n_tot] ymin;
  vector<lower=0>[n_tot] y;
  int<lower=0> x_ind [n_tot];
  int<lower=0> x_new;
  int<lower=0> Depth_new;
}

transformed data {
  vector<lower=0, upper=1>[n_tot] y_hat;
  for (i in 1:n_tot)
    y_hat[i] = (y[i]-ymin[i])*pow(ymax[i]-ymin[i], -1);
}

parameters {
  real<lower=0> beta_1[3];
  real<lower=0> beta_2[3];
  real<lower=0> beta_3[3];
  real<lower=0> phi;
  real<lower=0> y_pred;
  real<lower=0, upper= 1> fitted[24];
}

transformed parameters {
  vector<lower=0>[24] mu;
  vector<lower=0>[24] sigma2;
  real<lower=0> mu_pred;
  real<lower=0> sigma2_pred;

  for (j in 0: 2){
    for (i in 1:8){
      mu[j*8+i] = 1*pow((1 + pow(beta_1[Depth[j*24+i]] * x[i],
        beta_2[Depth[j*24+i]]),-beta_3[Depth[j*24+i]]));
      sigma2[j*8+i] = mu[j*8+i] .*(1-mu[j*8+i]).*pow((phi+1), -1);
    }
  }
  mu_pred = 1*pow((1 + pow(beta_1[Depth_new] * x_new, beta_2[Depth_new])), -beta_3[Depth_new]);
  sigma2_pred = mu_pred .*(1-mu_pred).*pow((phi+1), -1);
}

model {

```

```

for ( i in 0:2 ){
  for ( j in 0:2 ){
    for ( k in 1:8 ){
      y_hat[i*24+j*8+k] ~ beta(mu[i*8+x.ind[k]].*phi, (1-mu[i*8+x.ind[k]]).*phi);
    }
  }
}
for( i in 1:3){
  beta_1[i] ~ gamma(.01, .01);
  beta_2[i] ~ gamma(.01, .01);
  beta_3[i] ~ gamma(.01, .01);
}
phi ~ gamma(.01, .01);
y_pred ~ beta(mu_pred*phi, (1-mu_pred)*phi);

for (j in 0: 2){
  for (i in 1:8){
    fitted [j*8+i] ~ beta(mu[j*8+i]*phi, (1-mu[j*8+i])*phi);
  }
}
generated quantities{
  vector[n_tot] log_lik ;
  for ( i in 0:2 ){
    for ( j in 0:2 ){
      for ( k in 1:8 ){
        log_lik [i*24+j*8+k] = beta_lpdf(y_hat[i*24+j*8+k] | mu[i*8+x.ind[k]].*phi,
          (1-mu[i*8+x.ind[k]]).*phi);
      }
    }
  }
}

```

## Pooled-effects model - van Genuchten

---

```

data {
  int<lower=0> n_tot;
  vector<lower=0>[n_tot] Depth;
  vector<lower=0>[n_tot] x;
  vector<lower=0>[n_tot] ymax;
  vector<lower=0>[n_tot] ymin;
  vector<lower=0>[n_tot] y;
  int<lower=0> x.ind [n_tot];
  int<lower=0> x_new;
  int<lower=0> Depth_new;

```

```

}

transformed data {
    vector<lower=0, upper=1>[n_tot] y_hat;
    for (i in 1:n_tot)
        y_hat[i] = (y[i]-ymin[i])*pow(ymax[i]-ymin[i], -1);
}

parameters {
    real<lower=0> beta_1;
    real<lower=0> beta_2;
    real<lower=0> beta_3;
    real<lower=0> phi;
    real<lower=0> y_pred;
    real<lower=0, upper= 1> fitted[8];
}

transformed parameters {
    vector<lower=0>[8] mu;
    vector<lower=0>[8] sigma2;
    real<lower=0> mu_pred;
    real<lower=0> sigma2_pred;

    for (i in 1:8){
        mu[i] = 1*pow((1 + pow(beta_1 * x[i], beta_2)), -beta_3);
        sigma2[i] = mu[i] .* (1-mu[i]).*pow((phi+1), -1);
    }
    mu_pred = 1*pow((1 + pow(beta_1 * x_new, beta_2)), -beta_3);
    sigma2_pred = mu_pred .* (1-mu_pred).*pow((phi+1), -1);
}

model {
    for (i in 1:72){
        y_hat[i] ~ beta(mu[x.ind[i]]. *phi, (1-mu[x.ind[i]]). *phi);
    }
    beta_1 ~ gamma(.01, .01);
    beta_2 ~ gamma(.01, .01);
    beta_3 ~ gamma(.01, .01);
    phi ~ gamma(.01, .01);

    y_pred ~ beta(mu_pred*phi, (1-mu_pred)*phi);

    for (i in 1:8){
        fitted [i] ~ beta(mu[i]*phi, (1-mu[i])*phi);
    }
}

```

```
generated quantities{  
    vector[n_tot] log_lik;  
    for ( i in 1:72 ){  
        log_lik [i] = beta_lpdf(y_hat[i] | mu[x_ind[i]].*phi, (1-mu[x_ind[i]]).*phi);  
    }  
}
```

## References

- Costa Mota Paraíba, C., Bochkina, N., & Diniz, C. A. R. (2017). Bayesian truncated beta nonlinear mixed-effects models. *Journal of Applied Statistics*, 1–27.
- Cribari-Neto, F., & Zeileis, A. (2009). Beta regression in r.
- Vehtari, A., Gelman, A., & Gabry, J. (2016a). loo: Efficient leave-one-out cross-validation and waic for bayesian models [Computer software manual]. Retrieved from <https://github.com/stan-dev/loo> (R package version 1.0.0)
- Vehtari, A., Gelman, A., & Gabry, J. (2016b). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 1–20.