

Supplementary material for Technical Note: Approximate Bayesian parameterization of a process-based tropical forest model

Florian Hartig^{1,2}, Claudia Dislich^{1,3}, Thorsten Wiegand¹, and Andreas Huth¹

¹UFZ - Helmholtz Centre for Environmental Research, Department of Ecological Modelling, Permoserstr. 15, 04318 Leipzig, Germany

²University of Freiburg, Department of Biometry and Environmental System Analysis, Tennenbacher Str. 4, 79085 Freiburg, Germany

³University of Göttingen, Department of Ecosystem Modelling, Büsgenweg 4, 37077 Göttingen, Germany

Correspondence to: Florian Hartig
(florian.hartig@biom.uni-freiburg.de)

Abstract. This document provides supplementary information for the publication "Technical Note: Approximate Bayesian parameterization of a process-based tropical forest model" to appear in Biogeosciences.

1 FORMIND model details

FORMIND is an individual-based, spatially semi-explicit, dynamic forest model. Spatially semi-explicit means that trees are assigned to a spatial 20x20m grid, but within grid cells, trees have no explicit position. Horizontally, tree crowns are therefore homogeneously distributed across their respective grid cell. Trees mostly interact within grid cells, essentially through the mechanisms that are present in all classical gap models (e.g. Shugart, 1984; Bugmann, 2001): for each tree, growth and establishment depends on the light climate at its crown top. The light climate is determined by the overtopping leaf area. Additional to this central process of competition for light on grid cells, FORMIND implements a number of other processes that act across grid cells such as tree falling and seed dispersal between grid cells. We use the FORMIND model version of Dislich et al. (2009), developed for a tropical rainforest in Ecuador, with some minor updates that have accumulated since due to the ongoing development of the model. The FORMIND scheduling is given in Alg. 1, a visual representation of the model concept in Fig. 1. In the next paragraphs, we describe the processes in the model in some more detail.

Establishment is modeled as a constant seed rain, meaning that tree regeneration is independent of relative species

abundances in the modeled tree community. Provided that species-specific light conditions are met, the number of new recruits appearing on a site is a species specific parameter of the model. There is stochasticity in the recruitment regarding the spatial distribution of recruits.

After establishment, mortality acts on all trees individuals. Mortality originates from four sources: 1) Base mortality: each tree has a species-specific base mortality rate that is independent of its age and environmental conditions. 2) Small trees have an additional species specific size-dependent mortality (see Dislich et al., 2009, Appendix A). 3) Self-thinning: when the crown area in a particular height layer exceeds the plot area, trees are randomly removed until the layer is sufficiently thinned. 4) Gap-formation: When trees larger than a threshold diameter d_f die, they are assumed to fall on a neighboring plot and produce an additional mortality proportional to their crown size on all trees that do not exceed the height of the falling tree by one meter. In that sense, FORMIND differs from more traditional gap models, where only one "representative" plot is modeled. All four mortality processes are modeled stochastically.

For each remaining tree individual, productivity and thereby growth is calculated. Productivity and growth are modeled deterministically and depend on tree size and light climate at crown top, corrected by self-shading, as well as on PFT-specific functions for light-response, photosynthesis and respiration. The light climate on the plot is derived by calculating the leaf area contributed by the trees on the plot to different height layers, and from that the light interception of the different height layers. Respiration rates are calculated according to an inverse method which takes maxi-

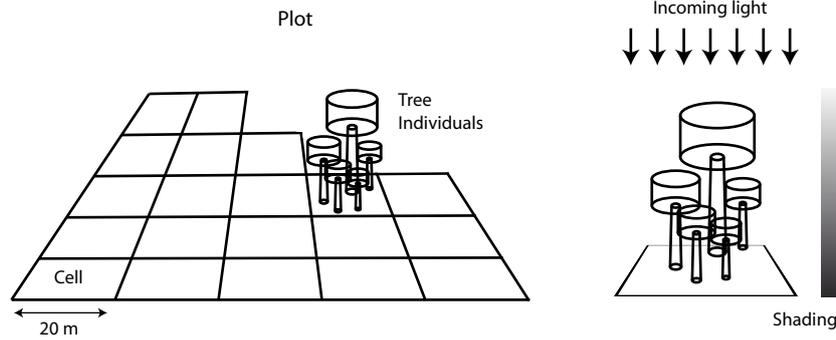


Fig. 1. Formind model principles. Crowns are illustrated as cylinders for the purpose of illustration; in the model there is no explicit position of the individual tree within the cell, and crowns are therefore distributed horizontally across the whole cell.

maximum growth rates under full light as an input (Dislich et al., 2009, Appendix A). Maximum growth rates can be derived from observations, but in this study, we treat growth rates as parameters that are fit to observed community data.

For reasons of computability, seedlings of the same PFT and age within one plot are grouped into cohorts, which is mathematically identical to calculating individual trees as there are no stochastic effects on existing tree individuals except for mortality.

Algorithm 1 FORMIND scheduling

```

1: Read initial tree configuration
2: for  $t = 1$  to  $t_{\text{end}}$  do
3:   for all Plots do
4:     if Light conditions allow establishment then
5:       Establishment of new seedlings from seed bank
6:     end if
7:     for all Trees do
8:       Mortality
9:     end for
10:    Update light climate
11:    for all Trees do
12:      Growth
13:    end for
14:  end for
15: end for

```

2 Statistical Algorithm

The algorithm for the Bayesian parameter estimation was implemented in Python 2.6, using Scipy (Jones et al., 2001), Numpy and parallel python (<http://www.parallelpython.com>). Parallel python was used to speed up the MCMC algorithm - instead of calculating the posterior value of one new parameter proposal, we always propose n values in parallel (here, n was 6 or 12). If the first value was rejected, the algorithm goes on

to check for acceptance of the second value and so on. If one value was accepted, the other values were discarded. The acceptance check was done strictly ordered, so that the order of steps within this algorithm is identical to that of a usual MCMC. The advantage, however, is that time is saved in the case of rejection because practically all our computational costs are for the FORMIND calculations (the estimation of mean and covariance of one parameter combination for a typical situation of 7 PFTs on 1 ha over 10.000 yrs with 5yr time steps takes around 20 seconds). Therefore, parallel proposals create a considerable speed-up (maximally by a factor n) when there are high rejection rates. Based on the observed rejection rates, we estimate that the average speed-up through parallelization with 6 cores was approximately a factor 3-4.

Algorithm 2 MCMC-SPL (parallel version)

```

1: Choose initial condition
2: Calculate initial unnormalized posterior value (eqs. 1,2, main text)
3: repeat
4:   Propose  $n$  new  $\phi$  according to proposal function  $q(\phi \rightarrow \phi')$ 
5:   Create  $n$  proposal  $\phi_i$  and run the model with those in parallel
6: repeat
7:   Estimate  $p(\phi'_i | D_{\text{obs}})$  according to eq. 1,2, main text
8:   Accept  $\phi'_i$  with probability  $\frac{p(\phi'_i | D_{\text{obs}})q(\phi \rightarrow \phi'_i)}{p(\phi | D_{\text{obs}})q(\phi'_i \rightarrow \phi)}$ , else stay at  $\phi$ 
9: until Acceptance of one  $\phi'_i$  or all  $n$  runs tested
10: If applicable, adjust  $q(\phi \rightarrow \phi')$  according to (Haario et al., 2001)
11: until Convergence

```

The algorithm was started with random initial values ϕ^{init} that were generated by adding a random parameter vector ϕ^r to the prior best estimate ϕ^* according to $\phi^{\text{init}} = 0.9 \cdot \phi^* + 0.1 \cdot \phi^r$. The best estimate ϕ^* was the "true" value for the virtual tropical forest, and the value from Dislich et al. (2009)

PARAMETER	ABBR.	LOWER-UPPER (TRUE)	UNITS
light extinction coefficient	li-ext	0.4-0.8 (0.6)	$[m^2 \cdot m^{-2}]$
recruitment rate pft1	recr1	50-200 (100)	$[ind/ha/yr]$
recruitment rate pft2	recr2	15-50 (30)	$[ind/ha/yr]$
recruitment rate pft3	recr3	10-40 (20)	$[ind/ha/yr]$
min light for establishment pft1	estab1	0-0.3 (0.1)	[-]
min light for establishment pft2	estab2	0-0.15 (0.05)	[-]
min light for establishment pft3	estab3	0-0.1 (0.01)	[-]
mortality pft1	mort1	0-0.25 (0.05)	$[yr^{-1}]$
mortality pft2	mort2	0-0.1 (0.15)	$[yr^{-1}]$
mortality pft3	mort3	0-0.05 (0.005)	$[yr^{-1}]$
falling probability of trees	fall	0.2-0.7 (0.4)	[-]
leaf area index per tree	LAI	1.5-2.5 (2)	$[m^2 \cdot m^{-2}]$
crown diameter	cr-d	0.12-0.2 (0.15)	[-]
crown length	cr-l	0.1-0.35 (0.25)	[-]
max dbh growth pf1	gro1	20-80 (41)	$[mm/yr]$
max dbh growth pf2	gro2	5-15 (9.2)	$[mm/yr]$
max dbh growth pf3	gro3	2-6 (3.5)	$[mm/yr]$
start growth pf1 (% of max)	start1	0-100 (40)	[-]
start growth pf2 (% of max)	start2	0-100 (40)	[-]
start growth pf3 (% of max)	start3	0-100 (40)	[-]
end growth pft1 (% of max)	end1	0-100 (10)	[-]
end growth pft2 (% of max)	end2	0-100 (10)	[-]
end growth pft3 (% of max)	end3	0-100 (10)	[-]
max growth diameter pft1	dia1	0.0-1.0 (1/3)	[-]
max growth diameter pft2	dia2	0.0-1.0 (1/3)	[-]
max growth diameter pft3	dia3	0.0-1.0 (1/3)	[-]

Table 1. Ranges for the uniform priors used for fitting the model to the virtual data. "Abbr." refers to the parameter abbreviation used in the figures. Lower, upper refers to the lower and upper bound of the uniform prior distributions. True refers to the values used to create the virtual data.

for the Ecuadorian parameterization. The proposal function $q(\phi \rightarrow \phi')$ was chosen multivariate normal, with a covariance adaptation according to (Haario et al., 2001) for the parameterizations to the virtual data, and a fixed proposal function for the parameterization to the Ecuadorian data. The covariance adaptation of Haario et al. (2001) sets the covariance of the proposal function as $\Sigma_i = c * cov_i(p(\Phi|D))$, where i is the i -th step of the algorithm, and the scaling parameter c is a constant whose optimal choice depends on the target function (we used $c = 2.38^2/d$, where d is the number of dimensions of the parameter space). Although the adaptive algorithm leads to more efficient proposal generation under correlations in the posterior, we noted that there were some remaining inefficiencies in the proposal generation that were probably due to the observed nonlinear and higher-order correlations in the posterior. To minimize the effect of those, we did not vary all parameters at once in one step of the MCMC, but first drew two random parameters, and then drew a proposal for those according to Alg. 2.

3 Prior ranges

Tables 1,2 show prior ranges and additional information of the parameters estimated in V1-V5 and E1, respectively.

4 Additional analyses of the parameter estimates

In the last part of this supplementary material, we show more detailed analyses for the results presented in the main article (in particular Figs. 3,5), and additional analyses that replicates the setup of Figs. 3,5 with differences in the number of parameters estimated, and in the aggregation type (summary statistics) that are used to compare model results and field data. A summary of the cases considered is provided in Table 3. For all these cases, we show a) a histogram of the marginal posterior density, which allows gaining a better picture of the distribution represented by the violin plots in the main paper, and b) a plot of posterior pair correlation density. The width of the marginal distributions was scaled to the prior width (denoted by the green lines at the sides of the plot). For the virtual case, the red line depicts the "true" parameter value that was used to create the virtual

PARAMETER	ABBR.	LOWER-UPPER (DISLICH ET AL.)	UNITS
recruitment rate pft1	recr1	5-100 (50)	[ind/ha/yr]
recruitment rate pft2	recr2	100-300 (180)	[ind/ha/yr]
recruitment rate pft3	recr3	50-250 (130)	[ind/ha/yr]
recruitment rate pft4	recr4	10-100 (50)	[ind/ha/yr]
recruitment rate pft5	recr5	50-200 (120)	[ind/ha/yr]
recruitment rate pft6	recr6	100-500 (310)	[ind/ha/yr]
recruitment rate pft7	recr7	20-100 (50)	[ind/ha/yr]
mortality pft1	mort1	0-0.25 (0.05)	[yr ⁻¹]
mortality pft2	mort2	0-0.1 (0.09)	[yr ⁻¹]
mortality pft3	mort3	0-0.1 (0.05)	[yr ⁻¹]
mortality pft4	mort4	0-0.25 (0.05)	[yr ⁻¹]
mortality pft5	mort5	0-0.1 (0.06)	[yr ⁻¹]
mortality pft6	mort6	0-0.05 (0.018)	[yr ⁻¹]
mortality pft7	mort7	0-0.25 (0.008)	[yr ⁻¹]
max dbh growth type1	gro1	10-40 (20)	[mm/yr]
max dbh growth type2	gro2	5-30 (10)	[mm/yr]
max dbh growth type3	gro3	2-30 (6)	[mm/yr]
max dbh growth type4	gro4	1-15 (2)	[mm/yr]
max growth diameter pft1	dia1	0.0-1.0 (0.33)	[-]
max growth diameter pft2	dia2	0.0-1.0 (0.33)	[-]
max growth diameter pft3	dia3	0.0-1.0 (0.25)	[-]
max growth diameter pft4	dia4	0.0-1.0 (0.33)	[-]
max growth diameter pft5	dia5	0.0-1.0 (0.2)	[-]
max growth diameter pft6	dia6	0.0-1.0 (0.33)	[-]
max growth diameter pft7	dia7	0.0-1.0 (0.33)	[-]

Table 2. Ranges for the uniform priors for the Ecuadorian fit. Note that the grouping is for 7 PFTs, but there are also 4 broader growth types to which the 7 PFTs belong (see [Dislich et al., 2009](#), for details). "Abbr." refers to the parameter abbreviation used in the figures. Lower, upper refers to the lower and upper bound of the uniform prior distributions. Dislich et al. refers to the parameter values used in [Dislich et al. \(2009\)](#).

field data. For the dataset from Ecuador, the red line depicts the parameter values chosen by ([Dislich et al., 2009](#)). However, as we note in the main text, the model setup was not completely identical, so there are limits in the comparability for the Ecuadorian case. Prior and true parameter values are also provided in Tables 1,2. In the caption of the marginal density plots, we provide some additional information for the runs such as sample size, convergence diagnostics (using Geldman-Rubin, see [Gelman and Rubin, 1992](#); [Brooks and Gelman, 1998](#)) and runtime. In all cases, we removed 100.000 samples as burn-in from the chains.

4.1 V1: Parameterization to virtual data, details for results from the main paper

Figs. 2,3 show detailed plots for Fig. 3 of the main text, which allows a better assessment of the shape of the distributions, and of the parameter correlations.

4.2 V2: Parameterization to virtual data, full parameter set

Figs. 4,5 uses the same data as V1, but with a larger number of parameters estimated. Those additional parameters are parameters for the crown geometry, for specific leaf area, the light extinction coefficient, the leaf area index (LAI) per tree, and probability of gap formation after tree mortality (Table 1). As can be seen, marginal parameter uncertainty considerably increases when fitting a larger number of parameters, which must be due to additional trade-offs between the old and the new parameters with respect to the data. The lower amount of strong pair correlation in Fig. 5 as compared to Fig. 2 suggests that those newly introduced trade-offs are predominantly of higher order and therefore not picked up by the pair correlation plots. Again, we stress that this is neither a fundamental problem, nor specific to the simulation-based likelihood approximation, but simply a result from interactions between parameters with respect to the data used for the fit - we would most probably find the same results in a conventional Bayesian analysis. However, the wide marginal distributions that result from these correlations make it dif-

CASE	EXPLANATION	DIMENSIONS
<u>Parameterization to virtual data, 3 PFTs:</u>		
V1	Data: SSD, GRO, reduced parameters	12, 96
V2	Data: SSD, GRO, full parameters	26, 96
V3	Data: SSD, reduced parameters	12, 48
V4	Data: total SSD, reduced parameters	12, 16
V5	Data: BM, reduced parameters	12, 3
<u>Parameterization to Ecuadorian field data, 7 PFTs:</u>		
E1	Data: SSD	18, 112

Table 3. Overview of parameterizations for different models, parameters and summary statistics. Abbreviations for the data: SSD = stem size distribution (16 10-cm classes), GRO = size-specific average growth distribution (16 10-cm classes), BM = biomass. If not stated otherwise, the data type was available for each PFT separately. If we use the mean over all PFTs, we label the data with "total". Full parameters means that the parameters given in Tables 1,2 are under calibration. Reduced parameters means that from Table 1, only recruitment, mortality, maximum growth and maximum growth diameter are estimated. Dimensions gives the number of estimated parameters and the number of data points in that order.

difficult to see how parameter uncertainty is affected by the simulation-based approximation and by the choice of model output. The latter is the reason why we estimated only a reduced set of parameters for the main results.

4.3 V3-V5: Parameterization to virtual data with more aggregated model outputs

The next plots Figs. 6-11 show results from fitting the same parameters as in V1 (Fig. 3, main text), but using more aggregated model outputs (i.e. coarser summary statistics) for the comparison between model and observed data. One can see that width of the posterior distribution generally increases when going to more aggregated representations of the data. Also some new trade-offs between parameters appear while going to more aggregated outputs, while others disappear, potentially indicating higher-order interactions between parameters with respect to the more aggregated data.

4.4 E1: Parameterization to Ecuadorian data

Finally, Figs. 12,13 show details of the parameter estimation with field data from Ecuador. Our data consisted of size-abundance distributions only. From our results for the virtual dataset (Fig 7), we know that this data type leads to relatively strong correlations when fitting similar parameters as in V3, which makes the result difficult to interpret. To avoid these correlations, we estimated a lower number of parameters per plant functional type than for the virtual case. However, note that the number of parameters is still larger than for V3, due to the higher number of plant function types. For the model parameters that were not fit to data, we used the values from Dislich et al. (2009).

References

- Brooks, S., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph Stat.*, 434–455.
- Bugmann, H., 2001. A Review of Forest Gap Models. *Clim. Change* 51 (3), 259–305.
- Dislich, C., Günter, S., Homeier, J., Schröder, B., Huth, A., 2009. Simulating forest dynamics of a tropical montane forest in South Ecuador. *Erdkunde* 63, 347–364. doi: 10.1098/rstb.2003.1425
- Gelman, A., Rubin, D., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7 (4), 457–472.
- Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. *Bernoulli* 7 (2), 223–242.
- Jones, E., Oliphant, T., Peterson, P., et al., 2001. SciPy: Open source scientific tools for Python. <http://www.scipy.org/>
- Shugart, H. H., 1984. *A Theory of Forest Dynamics: The Ecological Implications of Forest Succession Models*. Springer, New York, USA, New York.

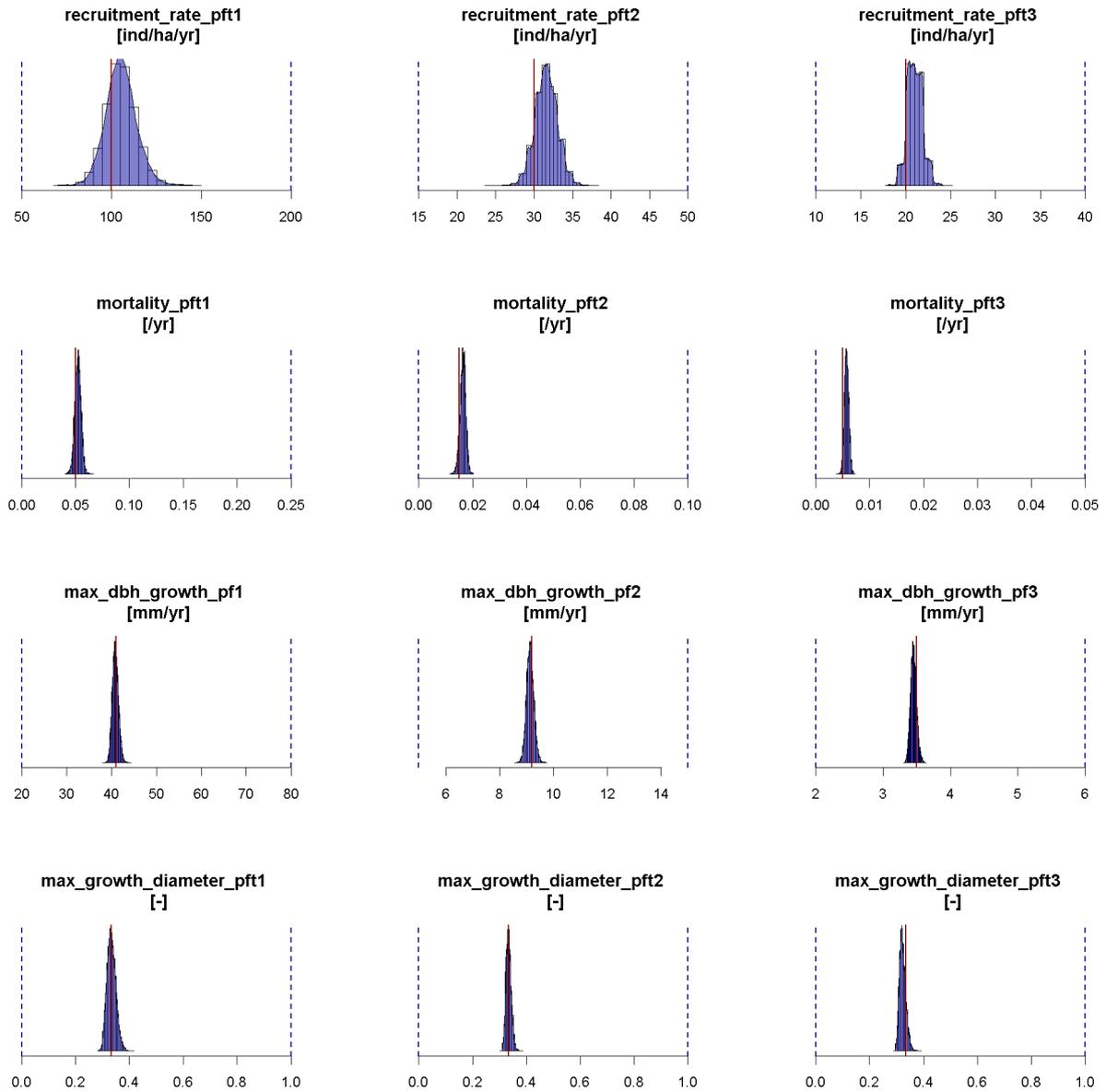


Fig. 2. Marginal posterior densities for case V1 in Table 3. Results from 3 chains; sample size per chain: ca. 1.3 million; Gelman-Rubin multivariate potential scale reduction factor: 1.01; runtime: ca. 6 weeks, 6 parallel cores per chain. Model parameters are explained in Table 1

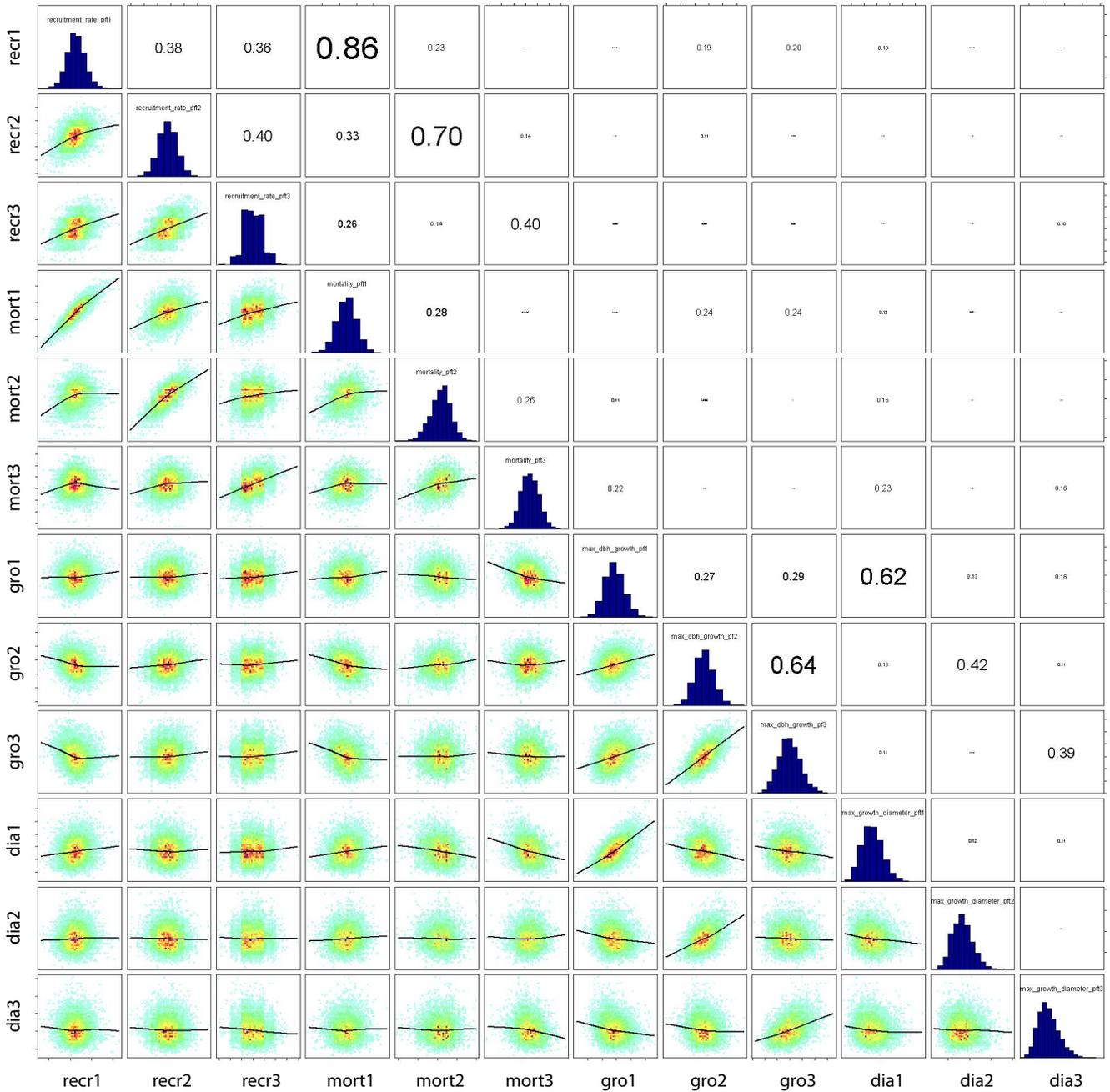


Fig. 3. Pair correlation density plot of the posterior distribution for case V1 in Table 3. As in Fig. 3b, main text, the histograms on the diagonal show the marginal distributions for the posterior parameter estimates that were also displayed in the previous figure. The panels in the lower triangle show pairwise correlations between the parameters. The diagonal shows the marginal densities for the parameters. The numbers in the upper triangle show Spearman's rank correlation coefficients for the correlations in the lower triangle. Parameter abbreviations are explained in Table 1.

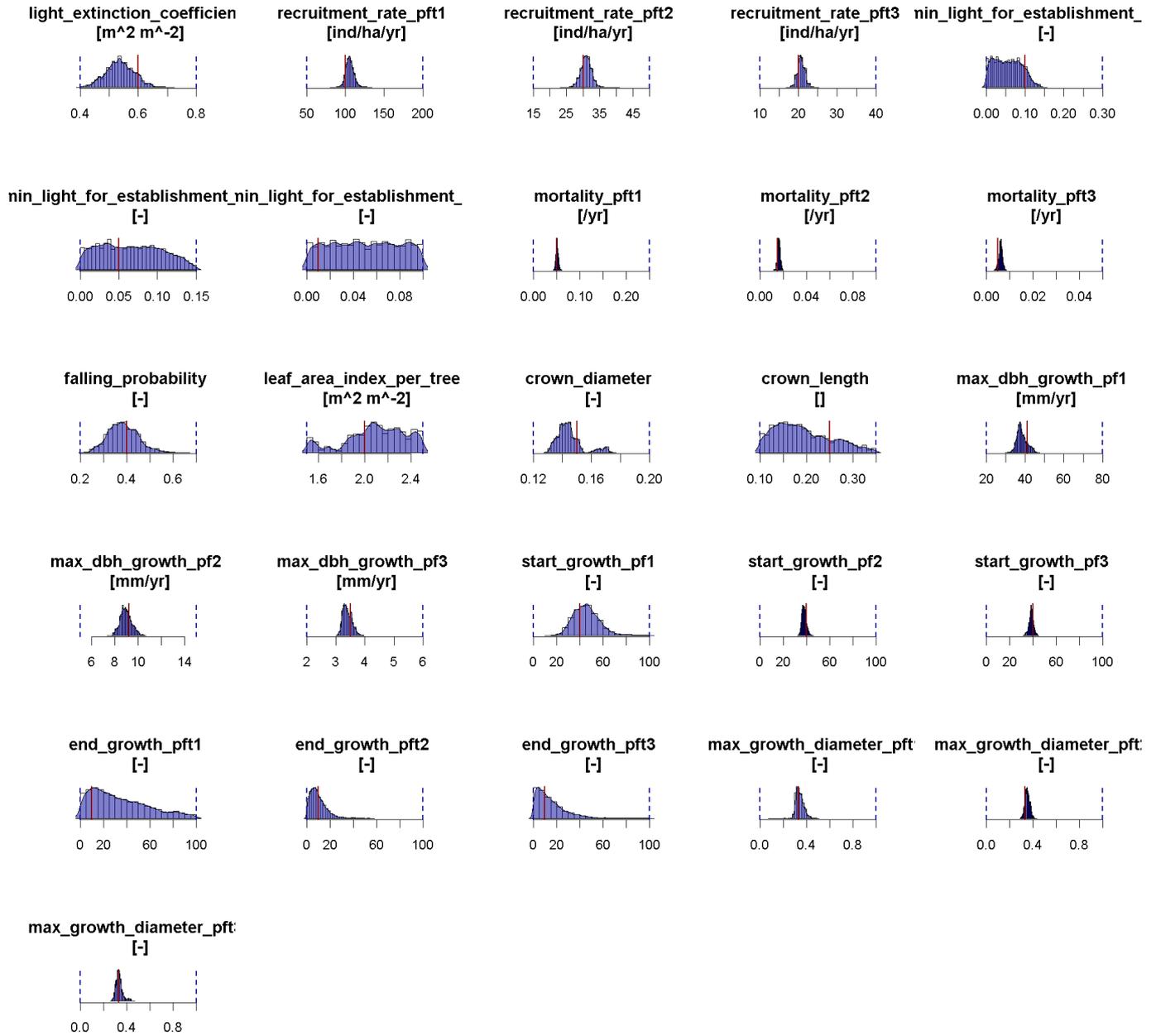


Fig. 4. Marginal posterior densities for case V2 in Table 3. Results from 3 chains; sample size per chain: ca. 1.4 million; Gelman-Rubin multivariate potential scale reduction factor: 1.17; runtime: ca. 6 weeks, 6 parallel cores per chain. Model parameters are explained in Table 1.

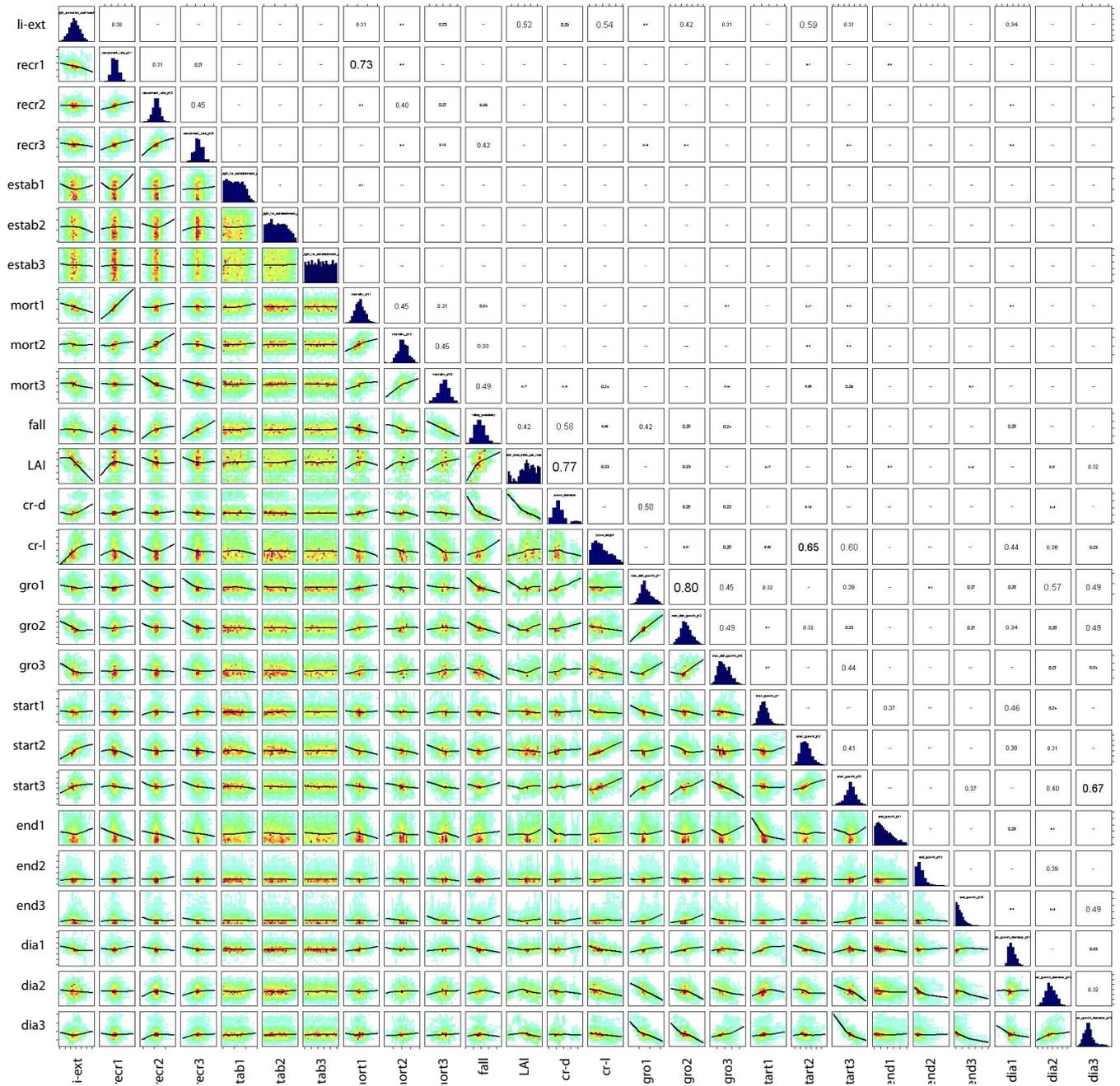


Fig. 5. Pair correlation density plot of the posterior distribution for case V2 in Table 3. As in Fig. 3b, main text, the histograms on the diagonal show the marginal distributions for the posterior parameter estimates that were also displayed in the previous figure. The panels in the lower triangle show pairwise correlations between the parameters. The diagonal shows the marginal densities for the parameters. The numbers in the upper triangle show Spearman's rank correlation coefficients for the correlations in the lower triangle. Parameter abbreviations are explained in Table 1.

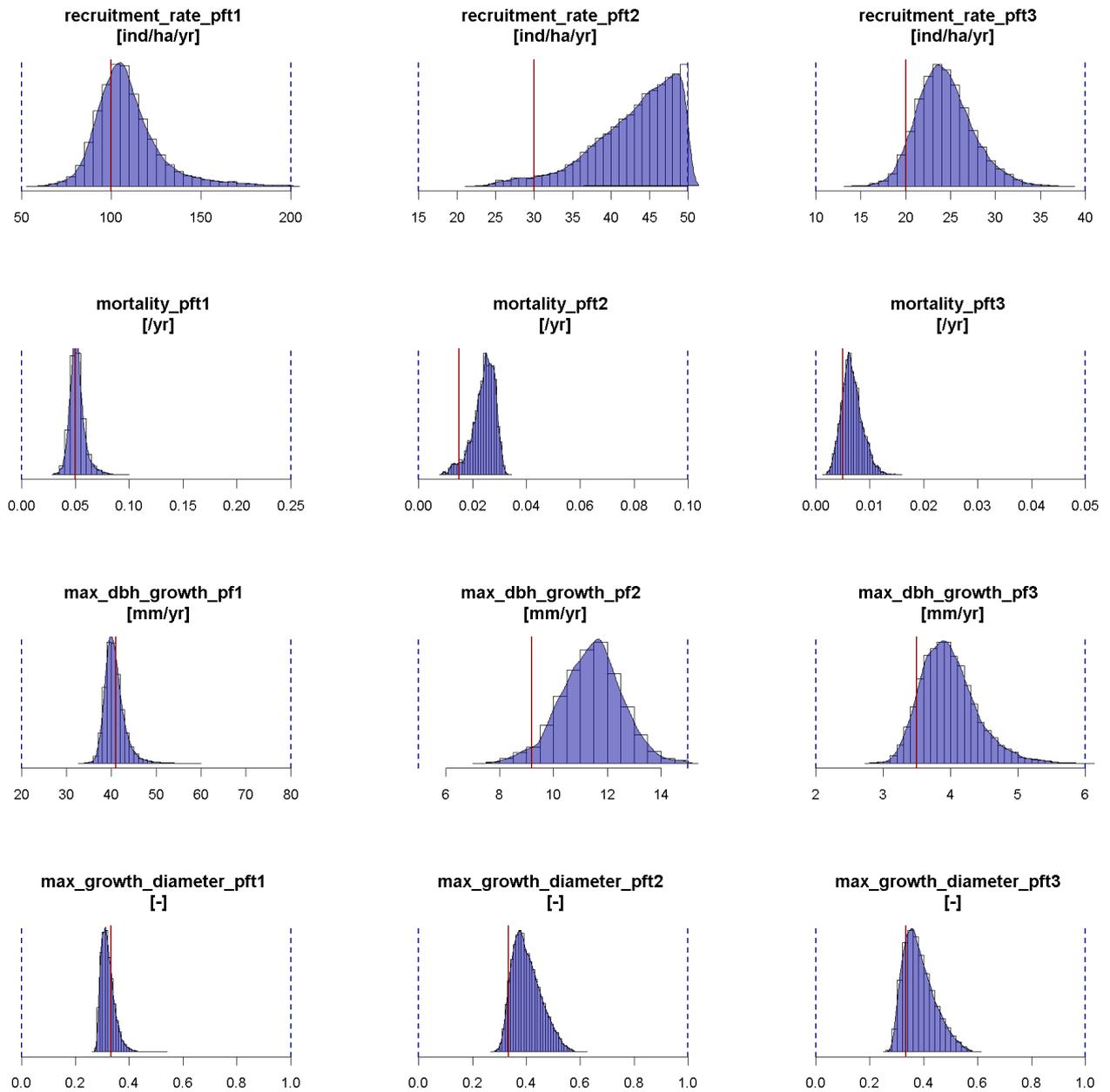


Fig. 6. Marginal posterior densities for case V3 in Table 3. Results from 2 chains; sample size per chain: ca. 1.4 million; Gelman-Rubin multivariate potential scale reduction factor: 1.02; runtime: ca. 6 weeks, 6 parallel cores per chain. Model parameters are explained in Table 1.

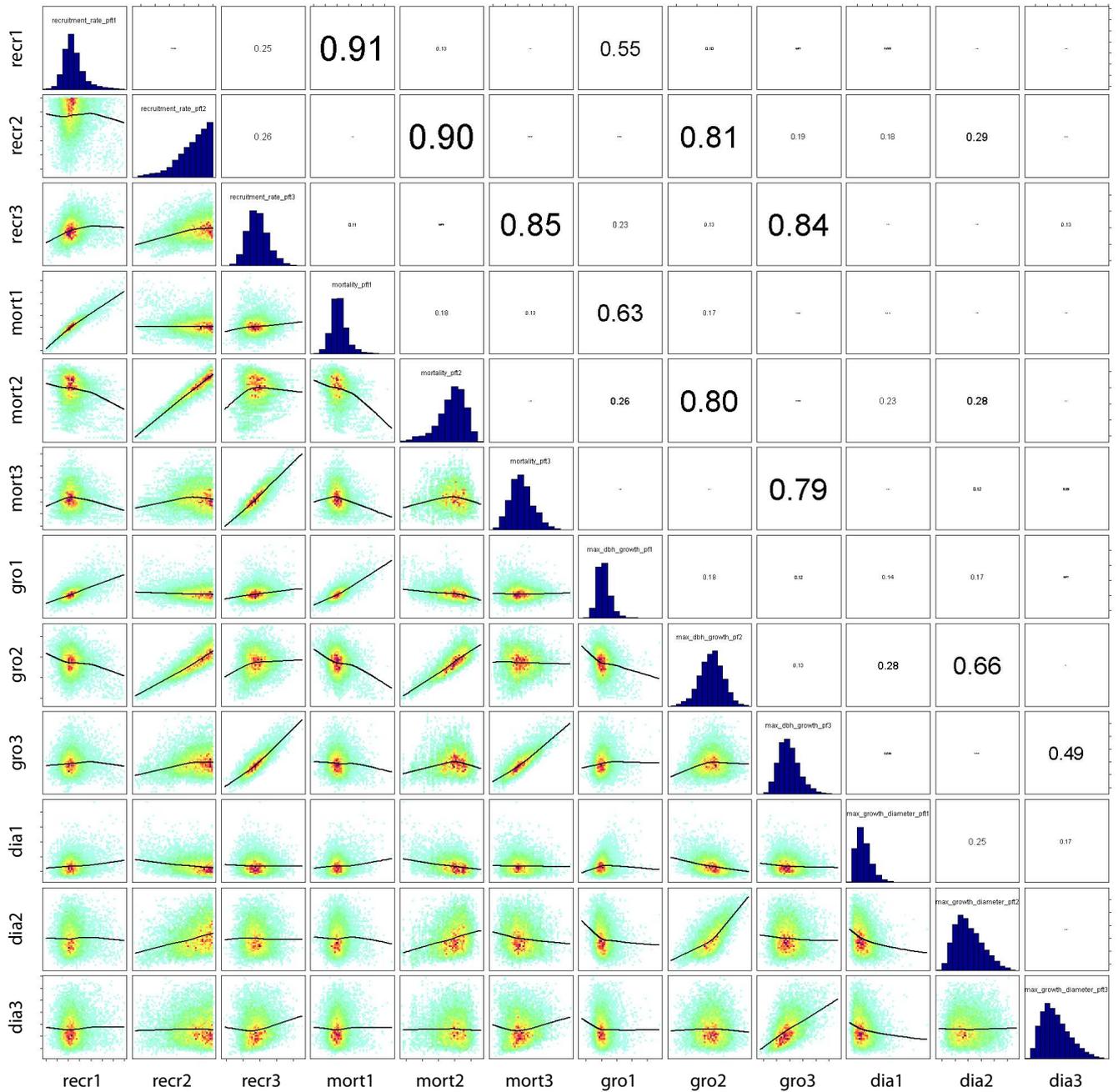


Fig. 7. Pair correlation density plot of the posterior distribution for case V3 in Table 3. As in Fig. 3b, main text, the histograms on the diagonal show the marginal distributions for the posterior parameter estimates that were also displayed in the previous figure. The panels in the lower triangle show pairwise correlations between the parameters. The diagonal shows the marginal densities for the parameters. The numbers in the upper triangle show Spearman’s rank correlation coefficients for the correlations in the lower triangle. Parameter abbreviations are explained in Table 1.

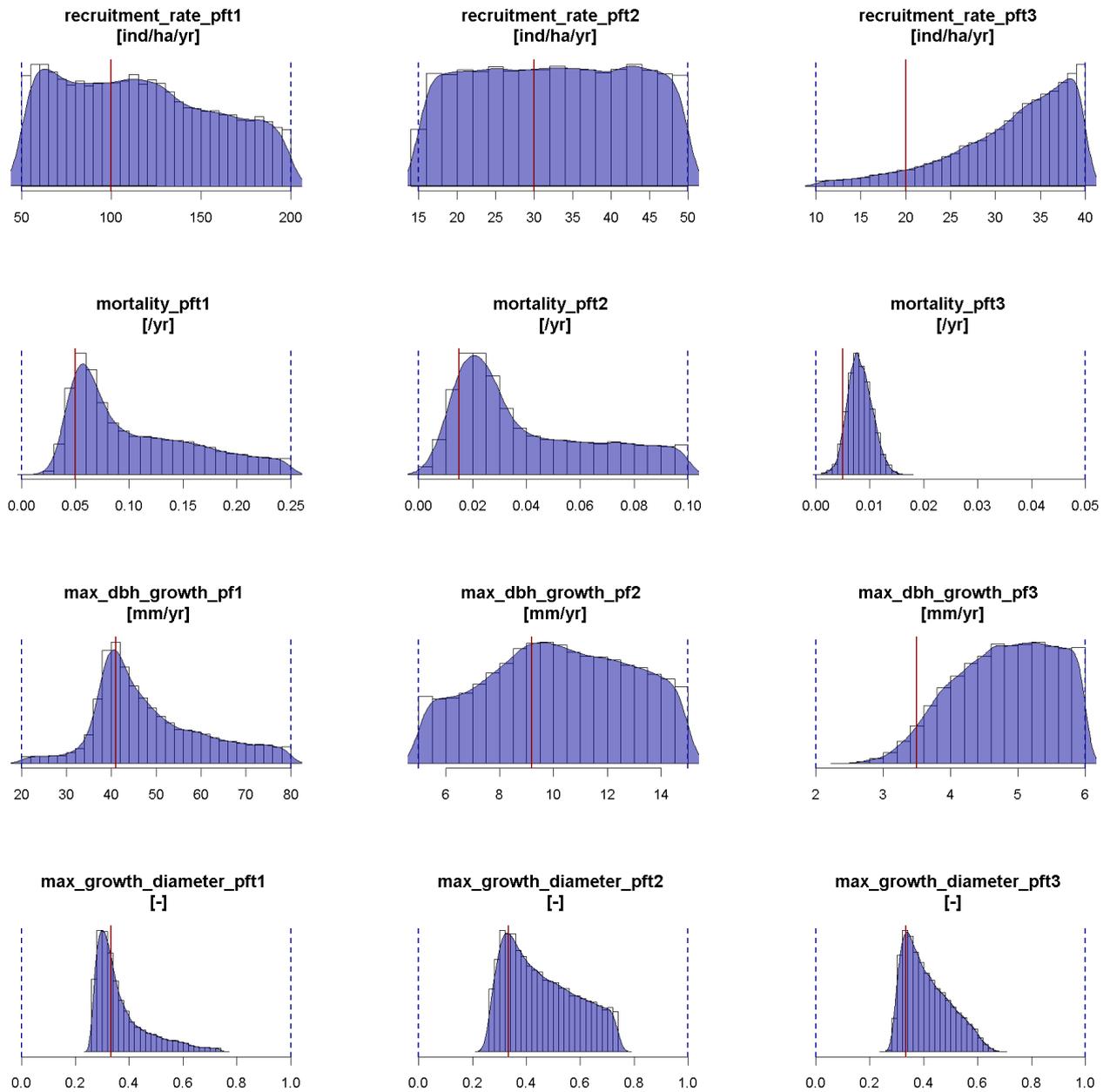


Fig. 8. Marginal posterior densities for case V4 in Table 3. Results from 2 chains; sample size per chain: ca. 0.7 million; Gelman-Rubin multivariate potential scale reduction factor: 1.04; runtime: runtime: ca. 6 weeks, 6 parallel cores per chain. Model parameters are explained in Table 1.

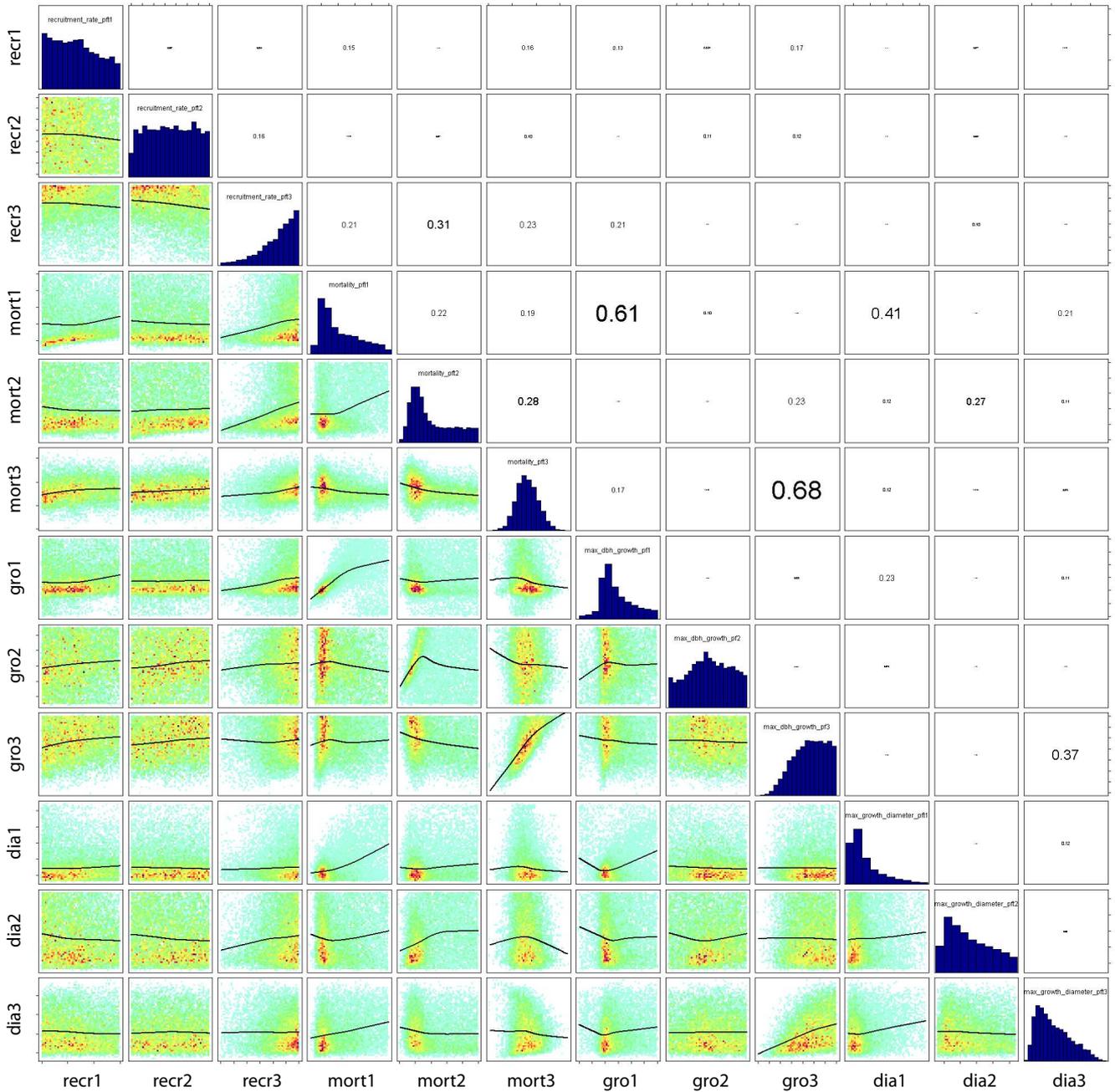


Fig. 9. Pair correlation density plot of the posterior distribution for case V4 in Table 3. As in Fig. 3b, main text, the histograms on the diagonal show the marginal distributions for the posterior parameter estimates that were also displayed in the previous figure. The panels in the lower triangle show pairwise correlations between the parameters. The diagonal shows the marginal densities for the parameters. The numbers in the upper triangle show Spearman's rank correlation coefficients for the correlations in the lower triangle. Parameter abbreviations are explained in Table 1.

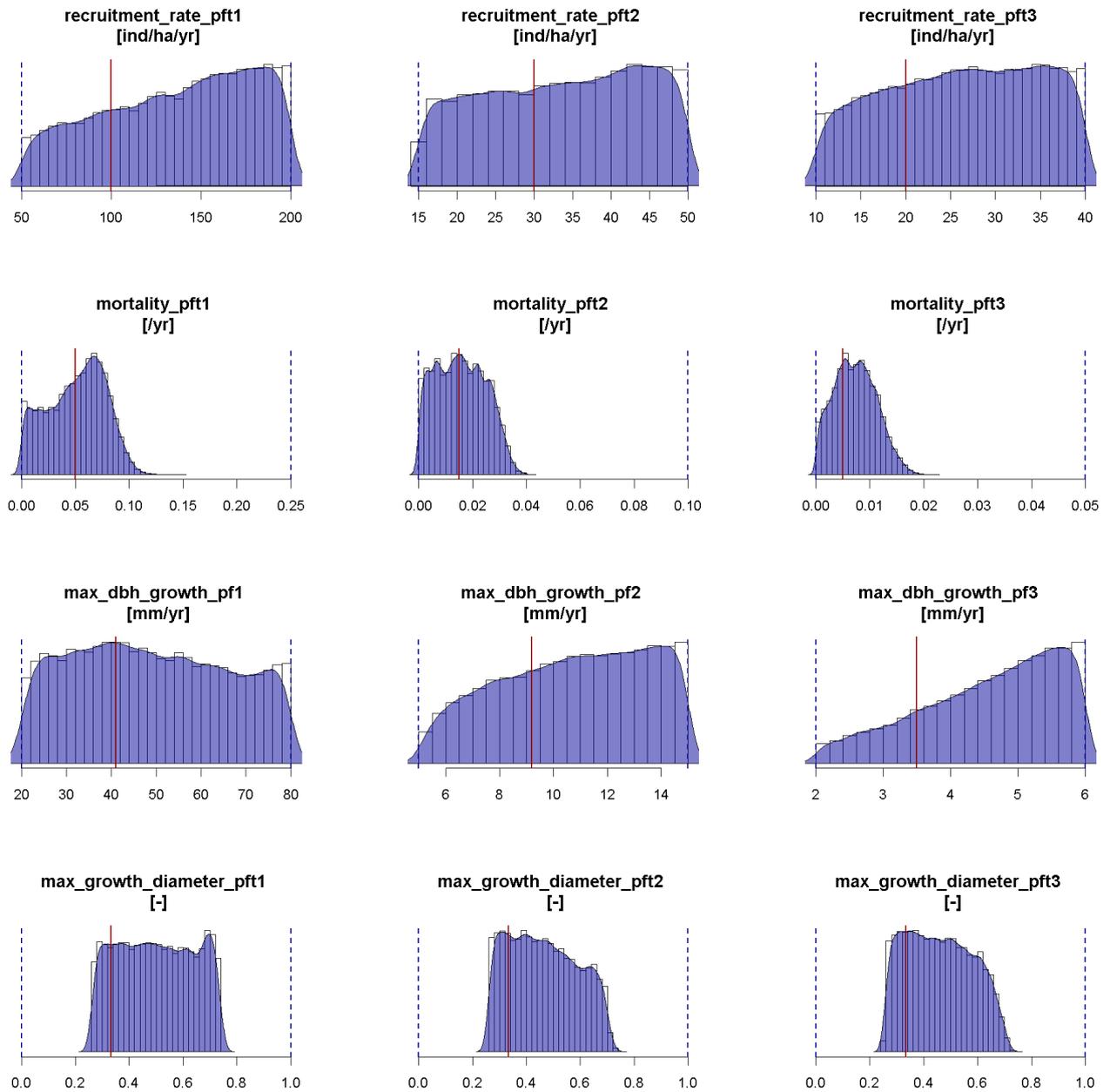


Fig. 10. Marginal posterior densities for case V5 in Table 3. Results from 2 chains; sample size per chain: ca. 1 million; Gelman-Rubin multivariate potential scale reduction factor: 1.08; runtime: ca. 6 weeks, 6 parallel cores per chain. Model parameters are explained in Table 1.

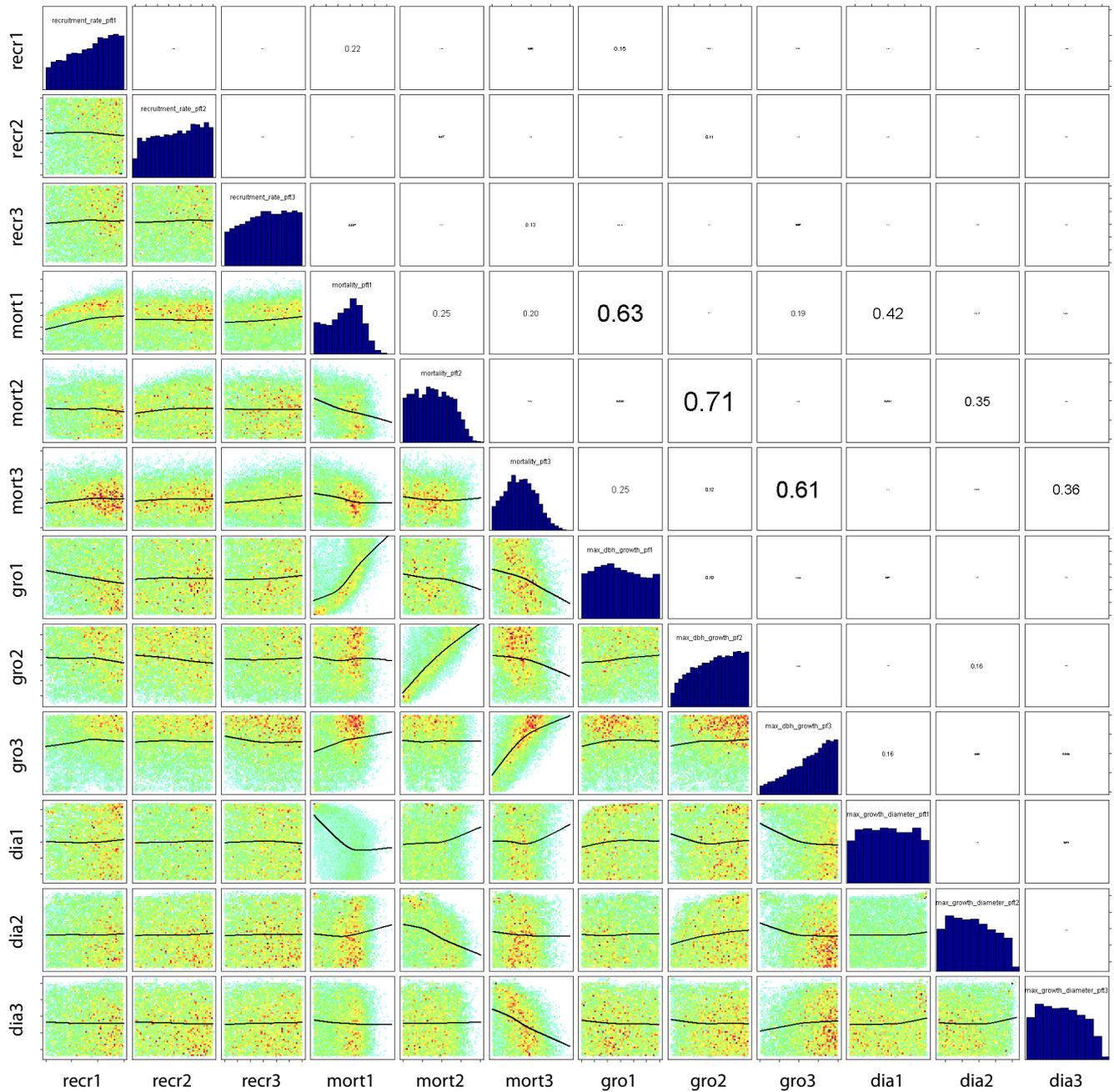


Fig. 11. Pair correlation density plot of the posterior distribution for case V5 in Table 3. As in Fig. 3b, main text, the histograms on the diagonal show the marginal distributions for the posterior parameter estimates that were also displayed in the previous figure. The panels in the lower triangle show pairwise correlations between the parameters. The diagonal shows the marginal densities for the parameters. The numbers in the upper triangle show Spearman's rank correlation coefficients for the correlations in the lower triangle. Parameter abbreviations are explained in Table 1.

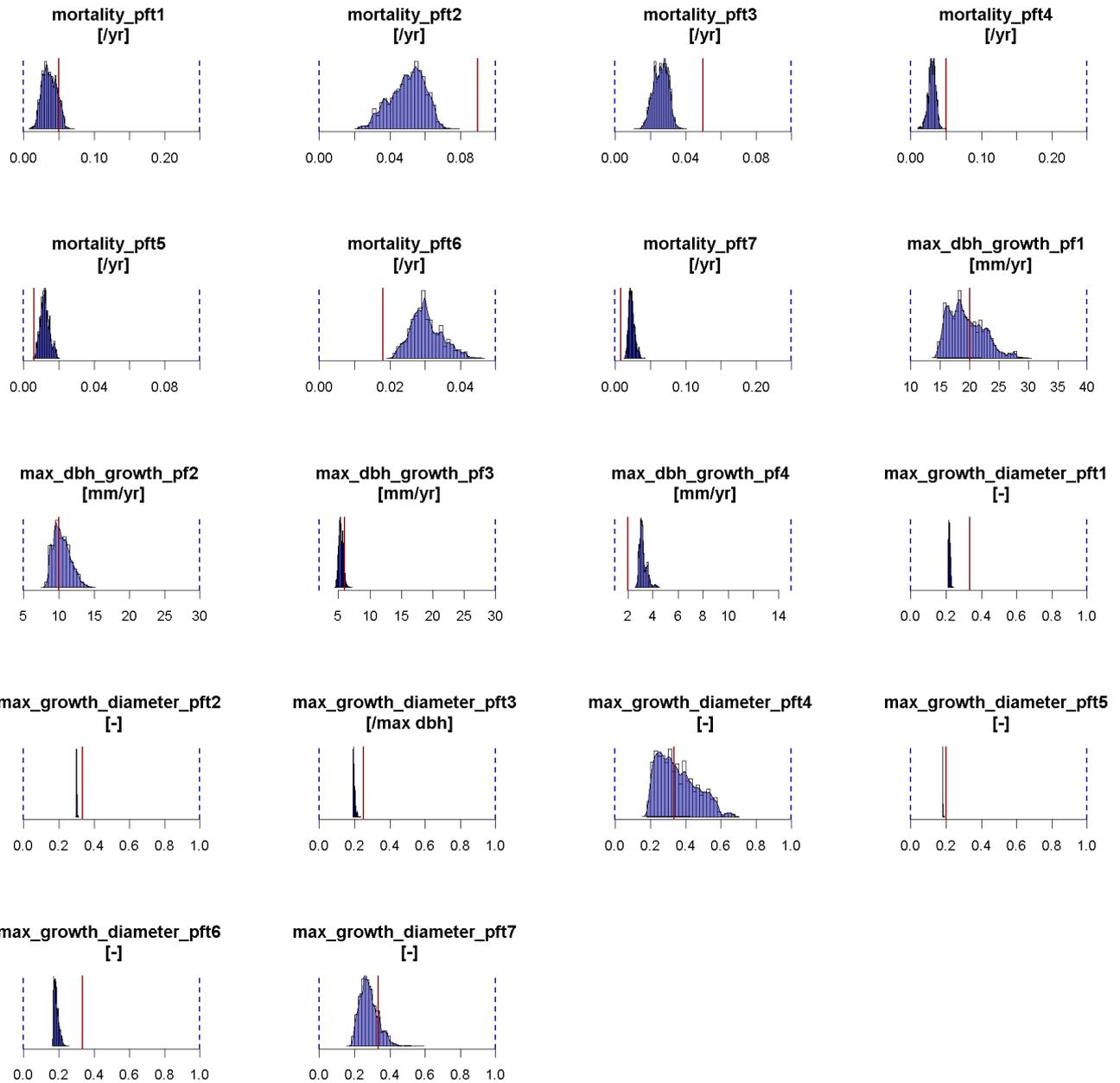


Fig. 12. Marginal posterior densities for case E1 in Table 3. Red bars show the parameter estimates by Dislich et al. (2009). Results from 3 chains; sample size per chain: 0.9 million; Gelman-Rubin multivariate potential scale reduction factor: 1.19; runtime: ca. 5 weeks, 12 parallel cores per chain. Model parameters are explained in Table 2.

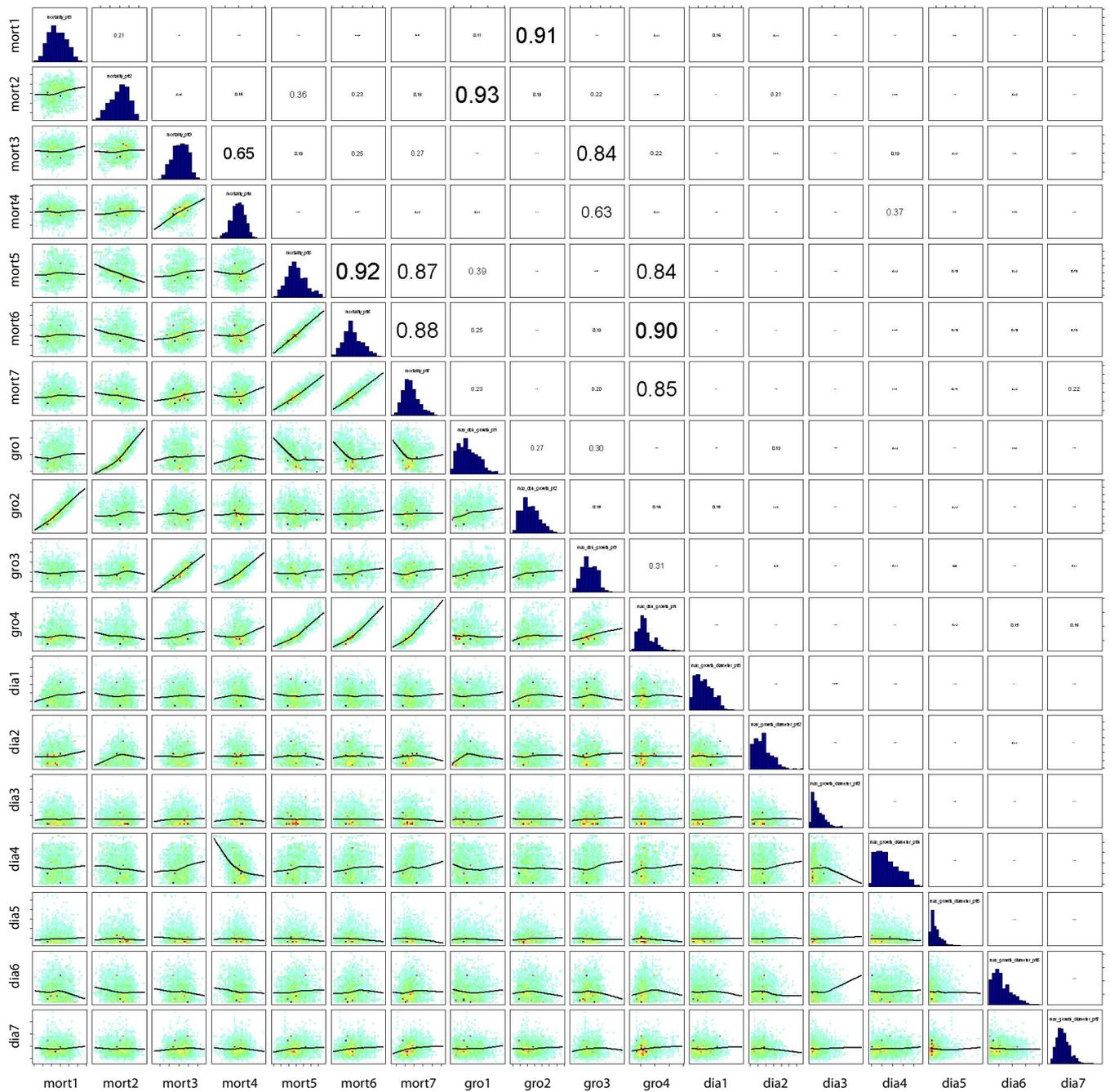


Fig. 13. Pair correlation density plot of the posterior distribution for case E1 in Table 3. As in Fig. 3b, main text, the histograms on the diagonal show the marginal distributions for the posterior parameter estimates that were also displayed in the previous figure. The panels in the lower triangle show pairwise correlations between the parameters. The diagonal shows the marginal densities for the parameters. The numbers in the upper triangle show Spearman's rank correlation coefficients for the correlations in the lower triangle. Parameter abbreviations are explained in Table 2. Note that the 4 growth parameters are assigned to the 7 PFTs as follows: gro1-PFT 2, gro2-PFT 1, gro3-PFTs 3,4, gro4-PFTs 5,6,7, which is reflected by the correlation structure.