Biogeosciences

Open Access

EGU

Supplement of

Uncertainty analysis of gross primary production partitioned from net ecosystem exchange measurements

Rahul Raj et al.

Correspondence to: Rahul Raj (r.raj@utwente.nl)

1. Potential scale reduction factor

The Gelman-Rubin potential scale reduction factor (PSRF) diagnostic uses an analysis of variance approach to evaluate convergence (Brooks and Gelman, 1998). Let m be the number of Markov chains and n be the length of each chain. For a parameter β, the simulations can be labelled as $\beta_{ij}(i = 1, 2, ..., n; j = 1, 2, ..., m)$. Within-chain variance (W) is calculated as the mean of the variances of each chain as:

$$W = \frac{1}{m} \sum_{j=1}^{m} s_j^2, \quad \text{where} \quad s_j^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\beta_{ij} - \overline{\beta}_{.j})^2, \quad \overline{\beta}_{.j} = \frac{1}{n} \sum_{i=1}^{n} \beta_{ij}. \tag{S1}$$

Between-chain variance (B) is based on $n \times m$ samples pooled together:

$$B = \frac{n}{m-1} \sum_{j=1}^{m} (\overline{\beta}_{.j} - \overline{\beta}_{..})^2, \quad \text{where} \quad \overline{\beta}_{..} = \frac{1}{m} \sum_{j=1}^{m} \overline{\beta}_{.j}. \tag{S2}$$

The variance of the stationary distribution is estimated as a weighted average of W and B:

$$\widehat{\text{Var}}(\beta) = \frac{n-1}{n} W + \frac{1}{n} B. \tag{S3}$$

The potential scale reduction factor (PSRF) is estimated as:

$$\widehat{R} = \sqrt{\frac{\widehat{\text{Var}}(\beta)}{W}}. \tag{S4}$$

If the PSRF is close to 1, this can be concluded that each of m Markov chains of n simulated observations of β is close to the target distribution. For practical purpose the convergence is assumed if PSRF<1.05 (Lunn et al., 2000). We estimated PSRF for each non-rectangular hyperbola (NRH) parameter and precision of likelihood (τ_e) using CODA package (Plummer et al., 2015) for the R Programming Language and Environment (R Core Team, 2014). NRH parameters are degree of curvature (θ) , quantum yield (α), photosynthetic capacity at light saturation ($A_{\max}$), ecosystem respiration at reference temperature (r_0),and temperature sensitive parameter (k_T).

The PSRF is estimated with uncertainty as Markov chain lengths are finite, therefore CODA package reports the lower (50%) and the upper (97.5%) quantiles of PSRF for each NRH parameter.

2. Gelman-Rubin-Brooks plot to identify burn-in period

Gelman-Rubin-Brooks (GRB) plot shows the evolution of PSRF as the number of iterations increases(Brooks and Gelman, 1998). We identified the appropriate length of the burn-in for both informative and non-informative prior distributions using the following steps:

1. We generated three Markov chains for each 10-day block. Trace plots of each β_i were generated by WinBUGS and were examined to evaluate whether the Markov chains converged approximately.

2. The Gelman-Rubin-Brooks (GRB) plot was then generated for each NRH parameter and τ_e to diagnose the convergence. The GRB plot was generated using he CODA package. Markov chains were divided into bins. The first bin contained samples from 1:50 iterations, the second bin from 1:50+n ($n = 10$ in CODA), the third contains sample from 1:(50+2n) and so on. The lower (50%) and the upper (97.5%) quantiles of Gelman-Rubin potential scale reduction factor (PSRF) or shrink factor (Eq. S4) is calculated for each bin and for each NRH parameter and τ_e that is plotted against bins to generate the GRB plot. A convergence is diagnosed if both quantiles approach 1 ($1 \leq$ PSRF ≤ 1.05). We analysed all 154 GRB plots (22 10-day blocks $\times$ 7 estimated parameters) for both informative and non-informative priors.

3. For each GRB plot, we identified the number of iterations after which the lower and upper quantiles of PSRF become close to 1. The number of iterations varied from 500 to 6000 when informative priors were used and from 2000 to 15000 when non-informative priors were used. We adopted the upper limits 6000 and 15000 as burn-in periods to ensure that Markov chains converged to stationary distributions for any NRH parameter and τ_e and for any 10-day block.

We presented below the GRB plot of each NRH parameter only for that 10 days block where we identified the maximum number of iterations for the convergence. Figure S1 shows the GRB plots in the case of informative prior distributions of the NRH parameters. It can be observed that 50th (median) and 97.5^{th} percentile of PSRF become close to 1 ($1 \leq$ PSRF ≤ 1.05) after 6000 iterations for the NRH parameters α and θ ("theta" and "alfa" respectively in the figure). This shows the convergence of Markov chains after 6000 iterations. For other NRH parameters, Markov chains converge before 6000.
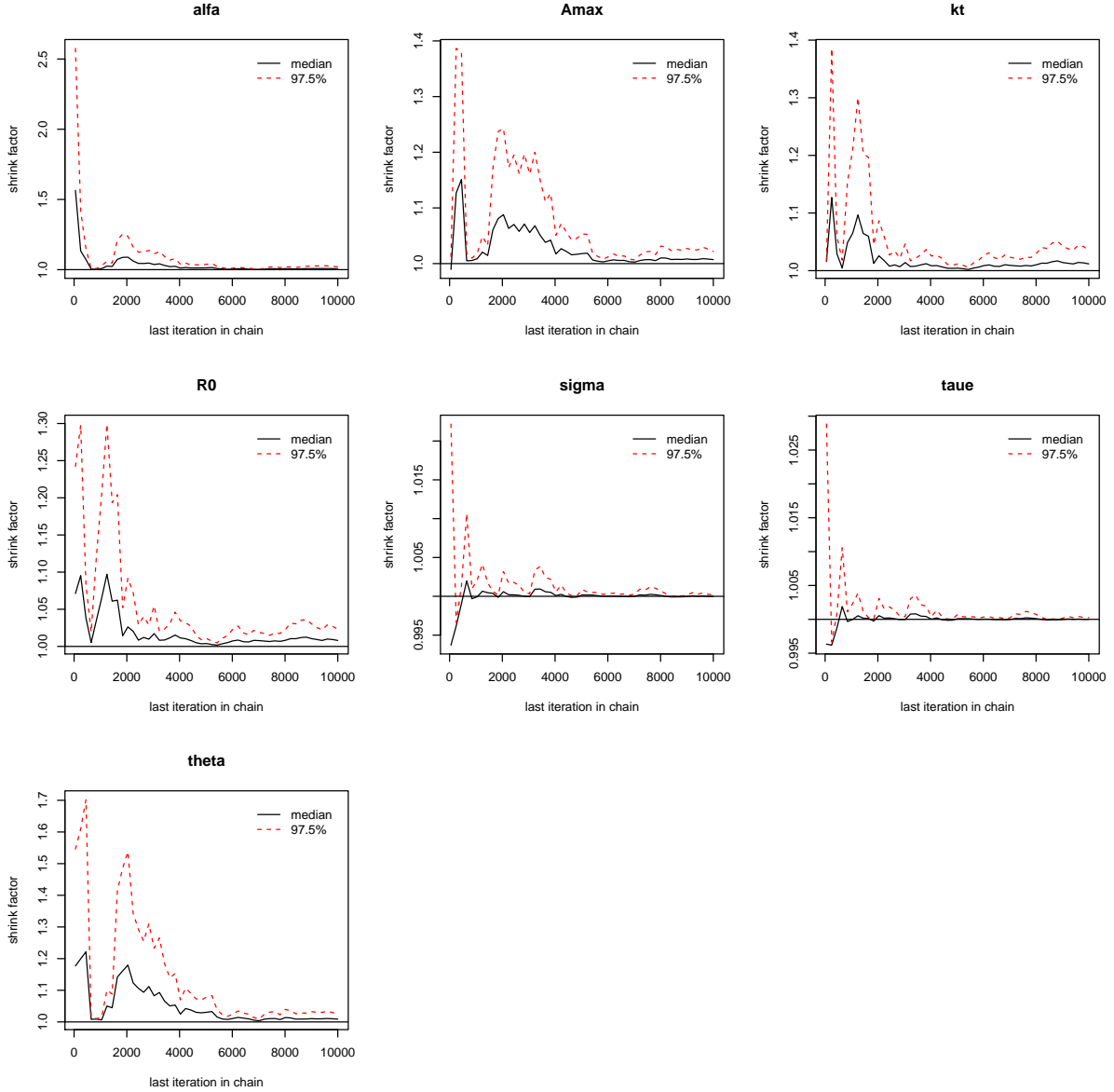
Figure S1: Gelman-Rubin-Brooks (GRB) plot of each NRH parameter for 8th September to 17th September 2009 (Julian days 251 to 260) for the choice of informative prior distributions. "alfa","Amax","kt","R0",and "theta" correspond to α, θ, A_{max}, r_0, k_{T}, and τ_{e} respectively. "sigma" and "taue" correspond to standard deviation (σ) and precision (τ_{e}) of the normal distribution of likelihood. Note that $\tau_e = 1/\sigma^2$. For information about the NRH parameters see Table 1 of the main paper.

Figure S2 shows the GRB plots in the case of non-informative prior distributions of the NRH parameters. It can be observed that 50th (median) and 97.5^{th} percentile of PSRF become close to 1 ($1 \leq$ PSRF ≤ 1.05) after 15000 iterations for the NRH parameter θ ("theta" in the figure). This shows the convergence of Markov chains after 15000 iterations. For other NRH parameters, Markov chains converge before 15000.
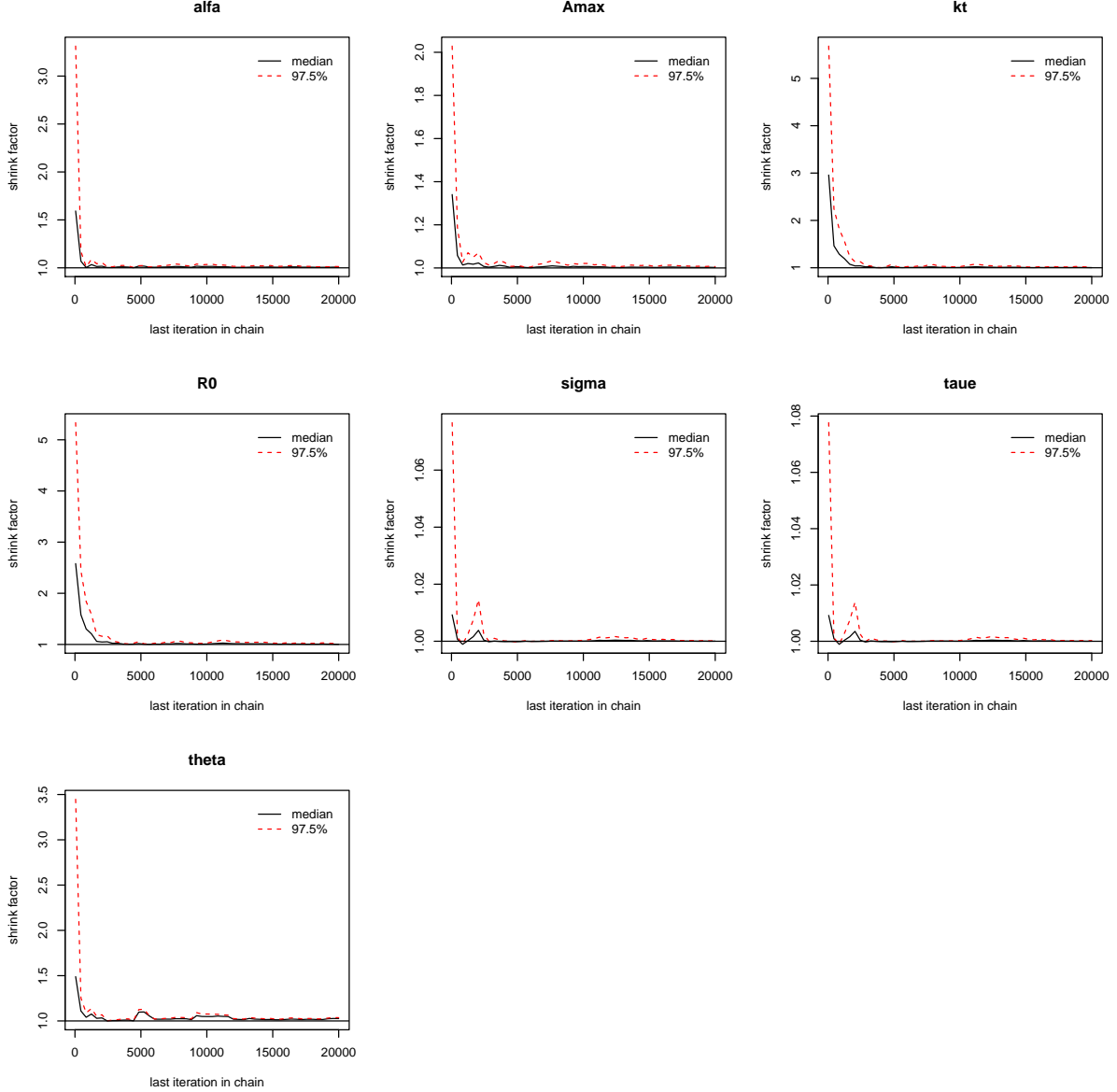

Figure S2: Gelman-Rubin-Brooks (GRB) plot of each non-rectangular hyperbola (NRH) parameter for 21st May to 30th September 2009 (Julian days 141 to 150) for the choice of non-informative prior distributions. "alfa","Amax","kt","R0",and "theta" correspond to α, θ, A_{max}, r_0, k_T, and τ_e respectively. "sigma" and "taue" correspond to standard deviation (σ) and precision (τ_e) of the normal distribution of likelihood. Note that $\tau_e = 1/\sigma^2$. For information about the NRH parameters see Table 1 of the main paper.

3. Trace plots for the NRH parameters after burn-in periods
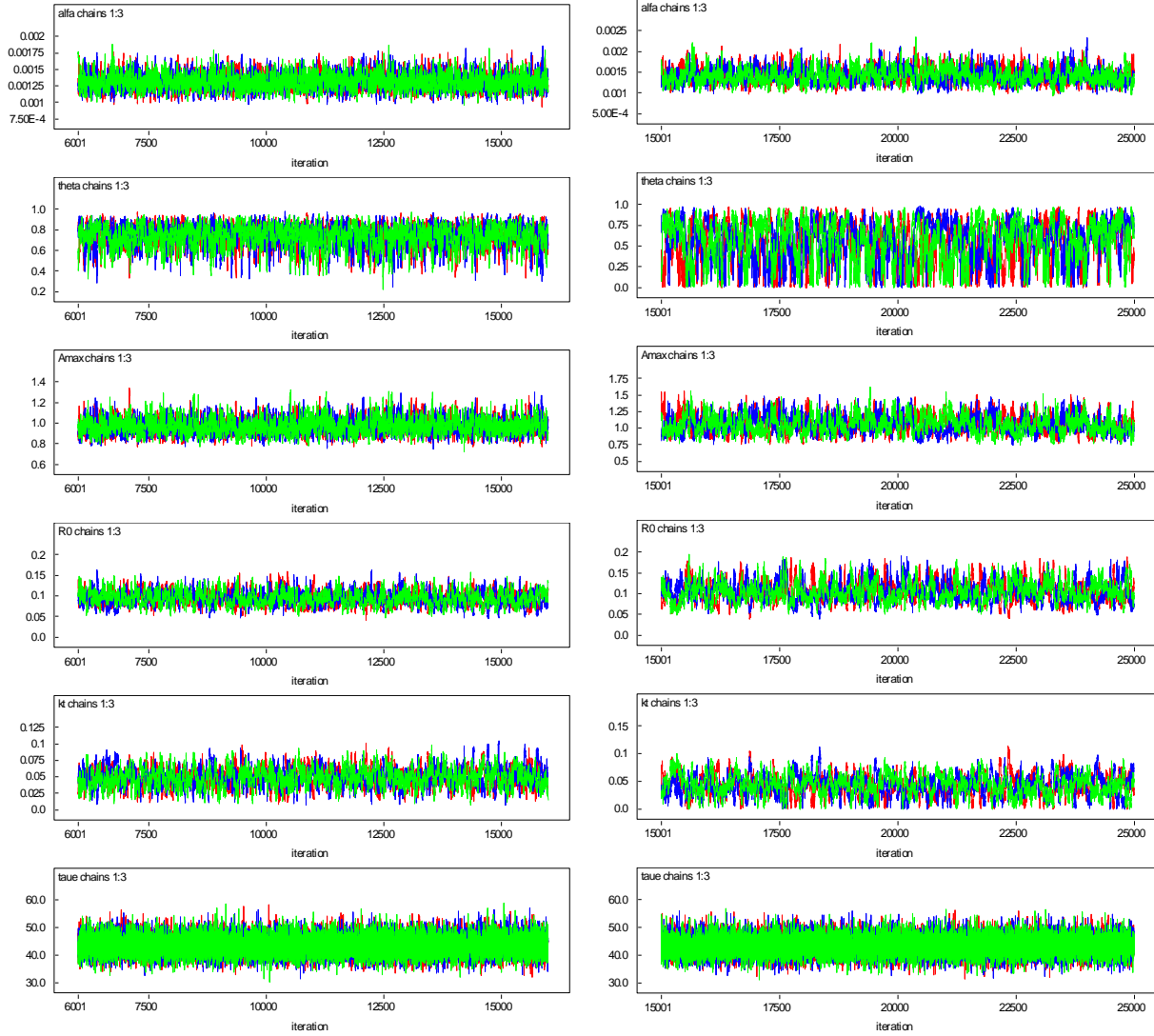
(a) Trace plots for informative prior distributions

(b) Trace plots for non-informative prior distributions

Figure S3: Trace plots of three Markov chains of 10000 post burn-in iterations for each NRH parameter and precision of τ_e for a 10-day block (1st May to 10th May 2009, Julian days 121 to 130). alfa, theta, Amax, R0, kt, and taue correspond to α, θ, A_{max}, r_0, k_T and τ_e respectively. For information about the NRH parameters see Table 1 of the main paper.

4. PSRF for the NRH parameters after burn-in period

Table S1: 50 and 97.5 percentile of potential scale reduction factor (PSRF) calculated for quantum yield (α), degree of curvature (θ) , photosynthetic capacity at light saturation (A_{max}), ecosystem respiration at reference temperature (r_0),and temperature sensitive parameter (k_{T}) and precision of likelihood (τ_{e}) after burn-in period for the choice of informative and non-informative prior distributions for a 10-day block (1$^{\mathrm{st}}$ May to 10$^{\mathrm{th}}$ May 2009, Julian days 121 to 130).

	Informative prior distributions		Non-informative prior distributions	
	50%	97.5%	50%	97.5%
α	1.00	1.00	1.00	1.01
θ	1.00	1.00	1.00	1.01
A_{max}	1.00	1.00	1.00	1.00
r_0	1.00	1.00	1.00	1.00
k_{T}	1.00	1.00	1.00	1.00
τ_{e}	1.00	1.00	1.00	1.00

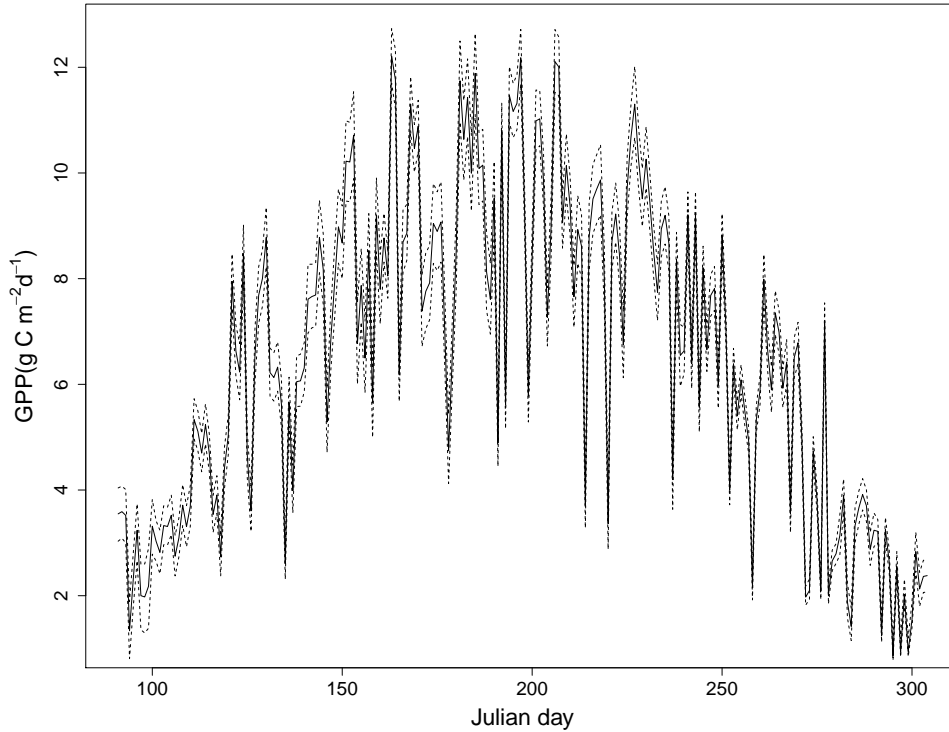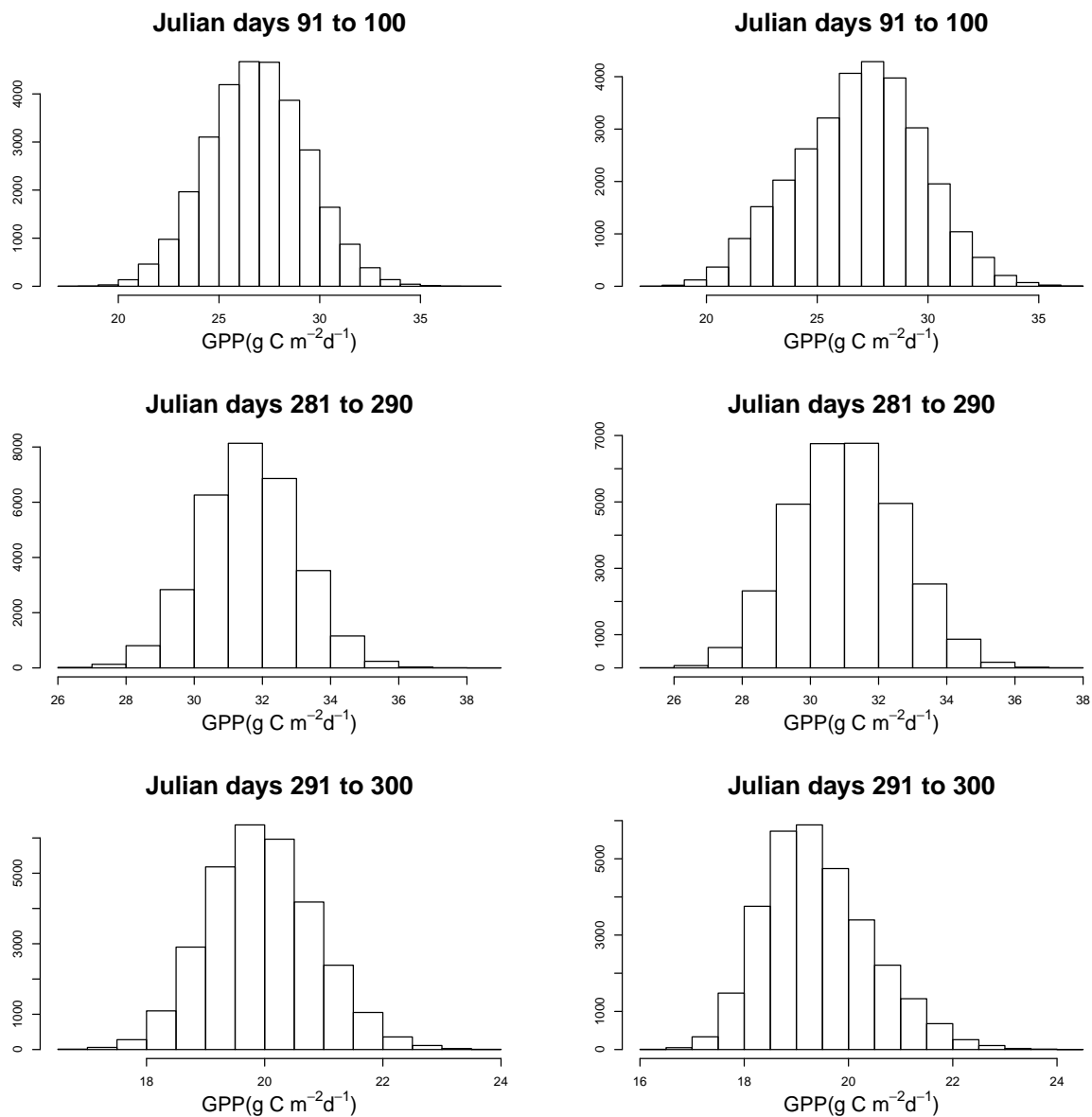## 5. Variation in daily gross primary production (GPP) during growing season

Figure S4: Median (solida line) and 95% credible intervals (dashed lines) of daily GPP distributions during the growing season of 2009 (1$^{\mathrm{st}}$ April to 31$^{\mathrm{st}}$ October 2009, Julian days 91 to 304) for the choice of non-informative prior distribution.

6. Histogram of 10 days sum of daily GPP

(a) For informative prior distributions   (b) For non-informative prior distributions

Figure S5: Distributions of sum of daily GPP for each of three 10-day blocks 91-100, 281-290, and 291-300.

7. Model file for WinBUGS

Model file specification for informative prior distributions of non-rectangular hyperbola (NRH) model parameters:

```
model
{
for(i in 1:n)
{
## normal distribution on likelihood with precision ''taue'' (τₑ).
y[i] ~ dnorm(mu[i], taue)
## likelihood
mu[i] <- (0.5/theta)*(alfa*x1[i]+Amax-sqrt(pow((alfa*x1[i]+Amax),2)-4*alfa*
Amax*theta*x1[i]))-R0*exp(kt*x2[i])
}
## gamma distribution on ''taue''
taue ~ dgamma(0.001,0.001)
## calculation of standard deviation ''sigma'' from precision ''taue''
sigma <- 1/sqrt(taue)
## informative priors of NRH parameters
alfa ~ dnorm(0.0022,2295684.114)
theta ~ dbeta(10,3)
Amax ~ dgamma(4,2.5)
kt ~ dgamma(4,120)
R0 ~ dbeta(2,64)
}
```

Model file specification for non-informative prior distributions of NRH model parameters:

```
model
{
for(i in 1:n)
{
y[i] ~ dnorm(mu[i], taue)
mu[i] <- -(0.5/theta)*(alfa*x1[i]+Amax-sqrt(pow((alfa*x1[i]+Amax),2)-4*alfa*
Amax*theta*x1[i]))-R0*exp(kt*x2[i])
}
taue ~ dgamma(0.001,0.001)
sigma <- 1/sqrt(taue)
## non-informative priors of NRH parameters with constraints
alfa ~ dnorm(0,0.001)I(0,)
theta ~ dnorm(0,0.001)I(0,1)
Amax ~ dnorm(0,0.001)I(0,)
kt ~ dnorm(0,0.001)I(0,)
R0 ~ dnorm(0,0.001)I(0,)
}
```

Where n is number of data points, y is half-hourly NEE data, x1 is half-hourly photosynthetic photon flux density (PPFD), x2 is half-hourly air temperature (T_a). alfa, theta, Amax, kt, and R0 correspond to NRH parameter α (quantum yield), θ (degree of curvature), A_{max} (photosynthetic capacity), r_0 (ecosystem respiration at reference temperature), k_T (temperature sensitive parameter) respectively.

References

Brooks, S. P. and Gelman, A.: General Methods for Monitoring Convergence of Iterative Simulations, Journal of Computational and Graphical Statistics, 7, 434–455, 1998.

Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D.: WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility, Statistics and Computing, 10, 325–337, 2000.

Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D., Bates, D., and Almond, R.: Package coda, URL `http://cran.r-project.org/web/packages/coda/coda.pdf`, accessed January 25, 2015, 2015.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org/`, 2014.