Supplement of Hydrol. Earth Syst. Sci., 25, 2663–2683, 2021
https://doi.org/10.5194/hess-25-2663-2021-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

A Bayesian approach to understanding the key factors influencing temporal variability in stream water quality – a case study in the Great Barrier Reef catchments

Shuci Liu et al.

Correspondence to: Shuci Liu (shucil@student.unimelb.edu.au)

Hierarchical prior specification and Bayesian inference of key drivers

Bayesian inference required specification of prior distributions for each model parameter. A minimally-informative uniform prior (denote as $U(\cdot)$) between 0 and 10 was assigned to the global standard deviation (σ, Eq. A1) (Gelman, 2006). The prior of I_n assumes that each indicator comes from an independent Bernoulli distribution, with a probability of 0.5 (Eq. A2) (Raftery et al., 1997). This vague prior results in each model structure having an equal prior model probability.

$$\sigma \sim U(0,10) \quad \text{A1}$$

$$I_n \sim Bernoulli(0.5) \quad \text{A2}$$

We used a hierarchical conditional prior specification for predictor coefficients, allowing the site-specific parameter values that describe the effects of each temporal predictors ($\beta_{1,j}, \beta_{2,j\ldots}, \beta_{n,j}$) to be exchangeable between sites (Liu et al., 2008; O'Hara and Sillanpää, 2009; Webb and King, 2009). The prior of $\beta_{n,j}$ was conditioned on I_n, resulting in a mixture distribution with 'slab and spike' prior, which was defined as follows,

$$\beta_{n,j} | I_n \sim I_n N(0,\tau_n) + (1-I_n) N(0,\tau_{n,tune}) \quad \text{A3}$$

where $\beta_{n,j} | (I_n = 1)$ is the slab part of the mixture distribution. The $\beta_{n,j} | (I_n = 1)$ was estimated by including a higher-level distribution. The prior of $\beta_{n,j} | (I_n = 1)$ followed a normal distribution with random effect (Eq. A4), with the τ_n drawn from a common prior distribution, defined as a hyperparameter (i.e., uniform distribution between 0 to 20, Eq. A5) (Gelman, 2006; Kruschke, 2014).

$$\beta_{n,j} | (I_n = 1) \sim N(0,\tau_n) \quad \text{A4}$$

$$\tau_n \sim U(0,20) \quad \text{A5}$$

For the spike component, a data-dependent prior was specified for $\beta_{n,j} \mid (I_n = 0)$, drawing from a *pseudo-prior* (Eq. A6), that is, *a priori* distribution with no effect on the posterior distribution, but facilitating the mixing of the Gibbs sampler.

$$\beta_{n,j} \mid (I_n = 0) \sim N(0, \tau_{n,tune}) \quad \text{A6}$$

We estimated $\tau_{n,tune}$ from the standard deviations of the posterior of the $\beta_{n,j}$ in a global model structure (i.e., modelling structure using all predictors), as suggested by Carlin and Chib (1995) and Linden and Roloff (2015). The prior of $\beta_{n,j} \mid (I_n = 0)$ was near the posterior estimates to facilitate mixing in the MCMC (Hooten and Hobbs, 2015).

The posterior inclusion probability (PIP - $P(I_n = 1 \mid \boldsymbol{y})$, Eq. A7) of each predictor was used to compare the relative importance of individual predictors (i.e., how often the n^{th} predictor was 'in' the model).

$$P(I_n = 1 \mid \boldsymbol{y}) = \frac{1}{T} \sum_{t=1}^{T} I(I_n^{(t)} = 1) \quad \text{A7}$$

where T is the total number of iterations of Markov chains. The different combination of I_n at each MCMC sampling represents a specific model structure. According to Bayes' theorem, the posterior model probability (PMP - $P(M_k \mid \boldsymbol{y})$) can be estimated as,

$$P(M_k \mid \boldsymbol{y}) = \frac{[\boldsymbol{y} \mid M_k] P(M_k)}{\sum_{x=1}^{L} [\boldsymbol{y} \mid M_x] P(M_x)} \quad \text{A8}$$

where L is the total number of possible models, and $P(M_k)$ is the prior probability of model M_k, among a group of models M_x, $x = 1, \ldots, X$. This posterior model probability can be obtained by assessing the frequency of a particular combination of I_n during the MCMC sampling.

Reference

Carlin, B. P. and Chib, S.: Bayesian model choice via Markov chain Monte Carlo methods, Journal of the Royal Statistical Society: Series B (Methodological), 57, 473-484, 1995.

Gelman, A.: Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper), Bayesian analysis, 1, 515-534, 2006.

Hooten, M. B. and Hobbs, N. T.: A guide to Bayesian model selection for ecologists, Ecological Monographs, 85, 3-28, 2015.

Kruschke, J.: Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, Academic Press, 2014.

Linden, D. W. and Roloff, G. J.: Improving inferences from short-term ecological studies with Bayesian hierarchical modeling: white-headed woodpeckers in managed forests, Ecology and evolution, 5, 3378-3388, 2015.

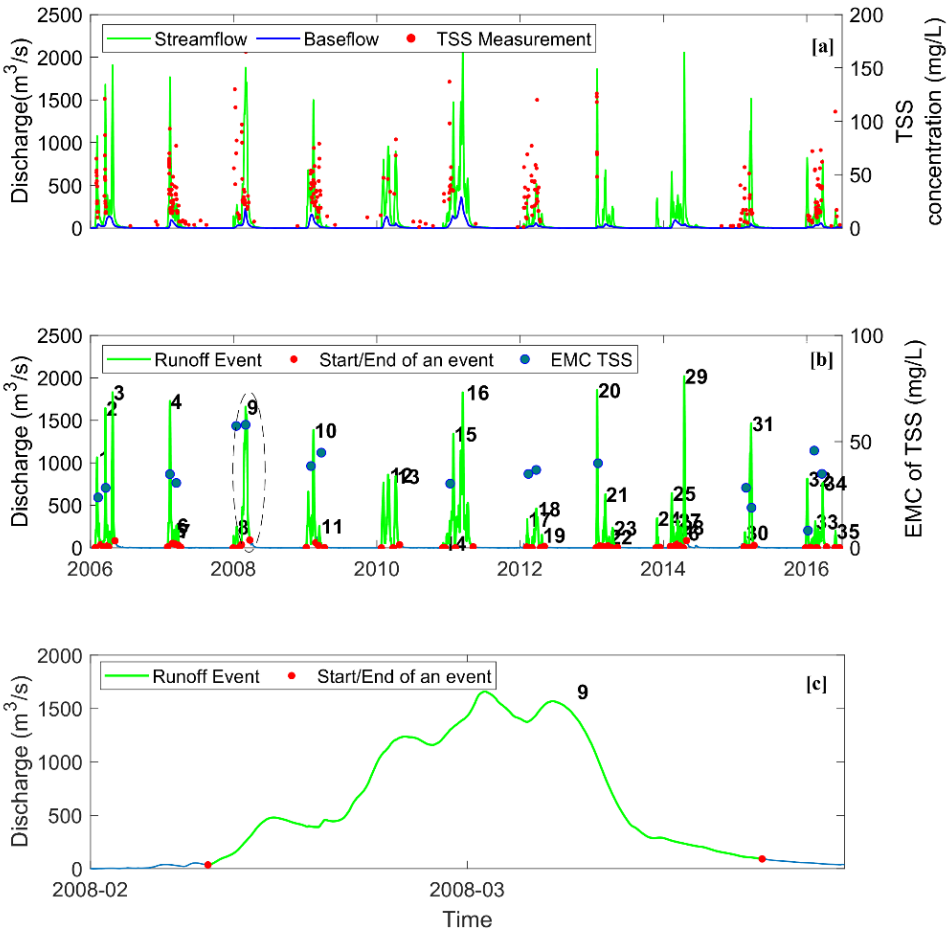
Liu, Y., Guo, H., Mao, G., and Yang, P.: A bayesian hierarchical model for urban air quality prediction under uncertainty, Atmospheric Environment, 42, 8464-8469, 2008.

O'Hara, R. B. and Sillanpää, M. J.: A review of Bayesian variable selection methods: what, how and which, Bayesian analysis, 4, 85-117, 2009.

Raftery, A. E., Madigan, D., and Hoeting, J. A.: Bayesian model averaging for linear regression models, Journal of the American Statistical Association, 92, 179-191, 1997.

Webb, A. and King, E. L.: A Bayesian hierarchical trend analysis finds strong evidence for large-scale temporal declines in stream ecological condition around Melbourne, Australia, Ecography, 32, 215-225, 2009.

Figure



60

Figure S1: Delineation of runoff events and estimation of EMCs, based on the hydrograph for 105107A Normanby River at Kalpowar Crossing in the GBR catchments: [a] baseflow separation from continuous streamflow observations; [b] event identification and development of EMC, and 35 runoff events are identified with red dots representing either the start or end of a runoff event; and [c] A zoom in event #9 in 2008.

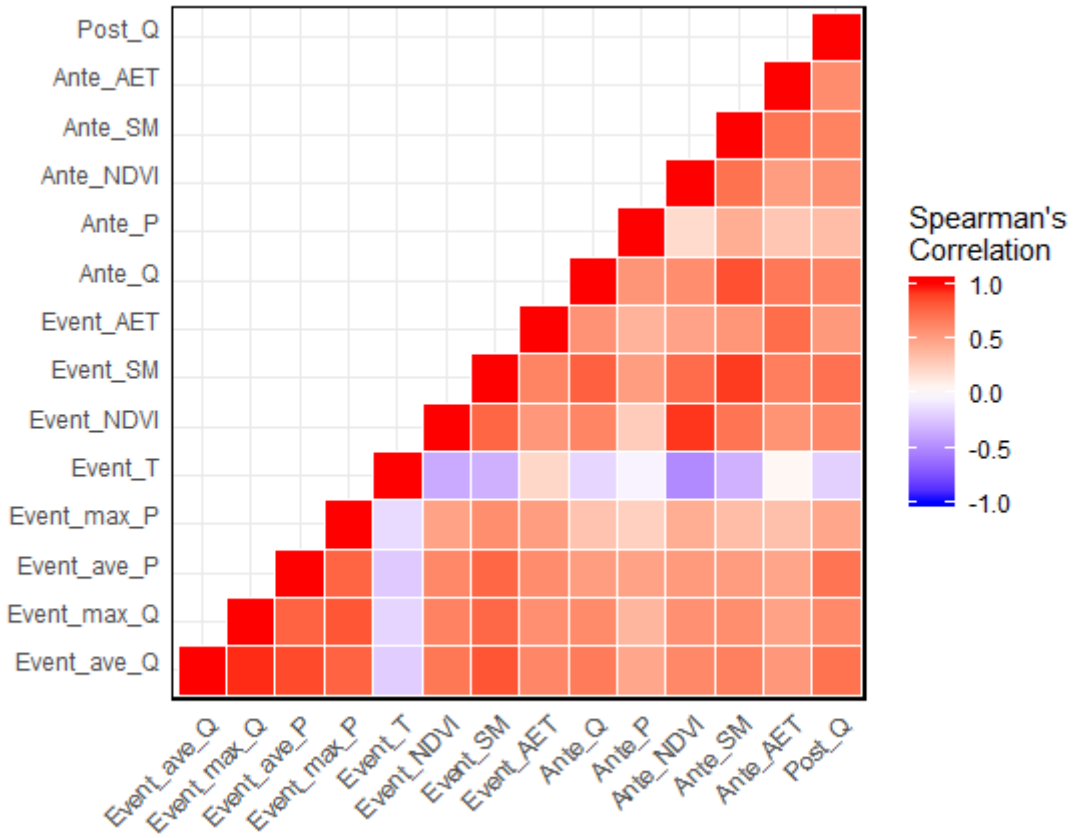


Figure S2: Spearman's Rank correlation between 14 candidate predictors.

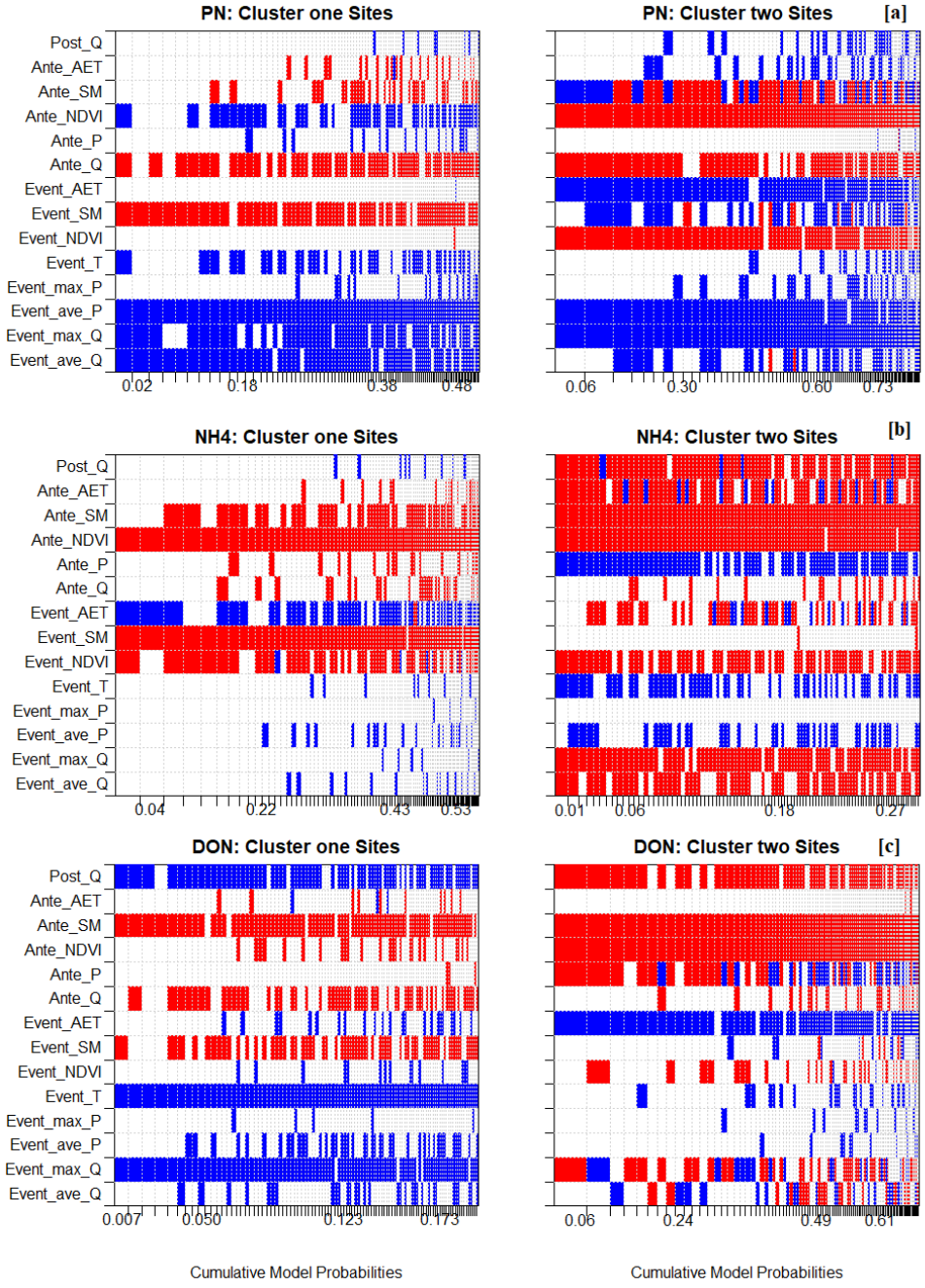


Figure S3: Comparison of BMA model coefficient and cumulative model probability (top 100 models) between two clusters for: (a) PN, (b) NH_4 and (c) DON. Left - cluster one sites and Right – cluster two sites. Each column in the heatmap represents the one specific model (ranked from highest model probability) and the width of the column is normalised by the posterior model probability. The colour indicates the direction of the coefficients, red – negative and blue – positive. Note: the coefficient value was averaged across the posterior median value of site-specific coefficient within each cluster (effect size, $\theta_{n,j}$, in Eq. (6)).

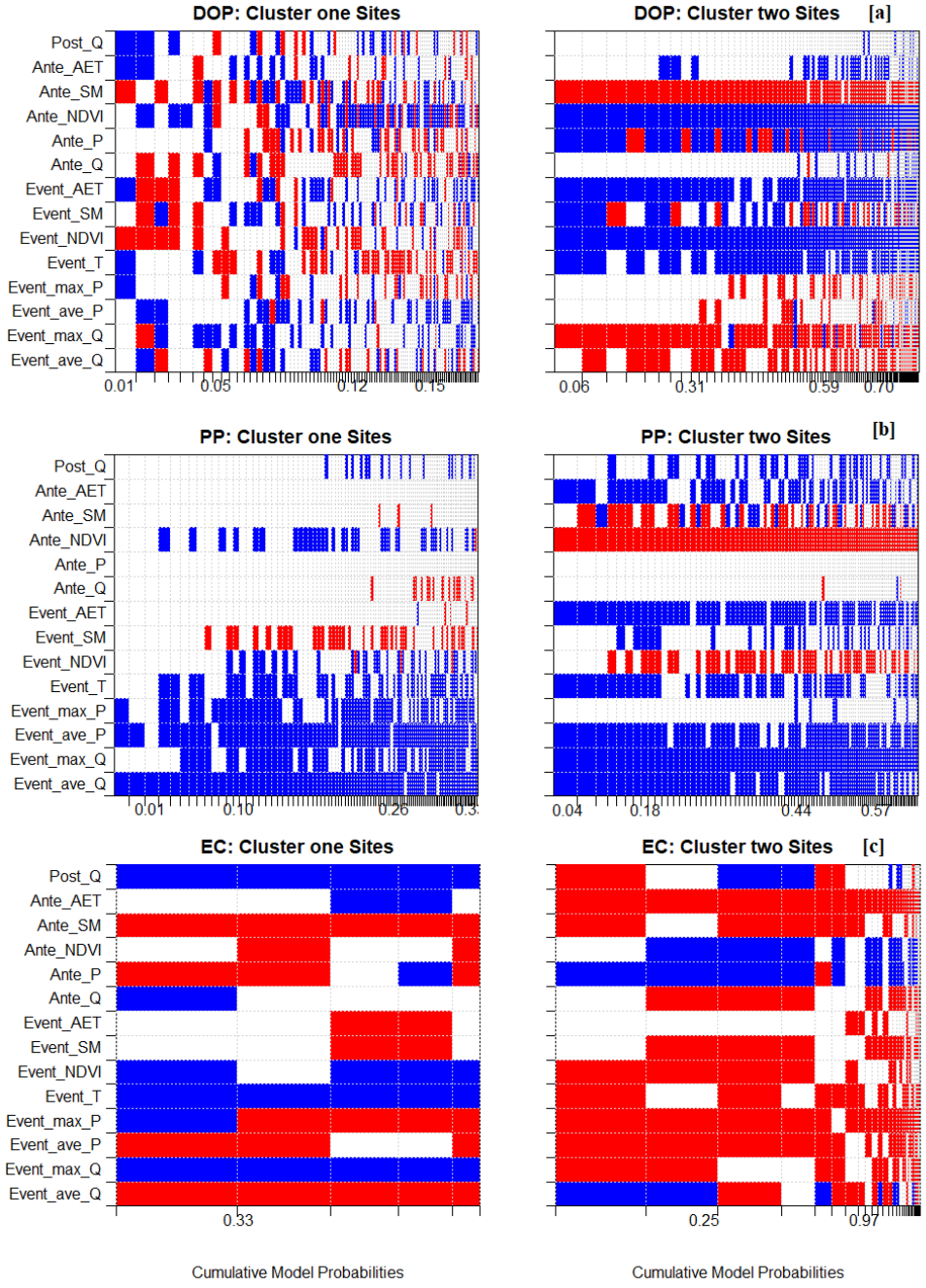
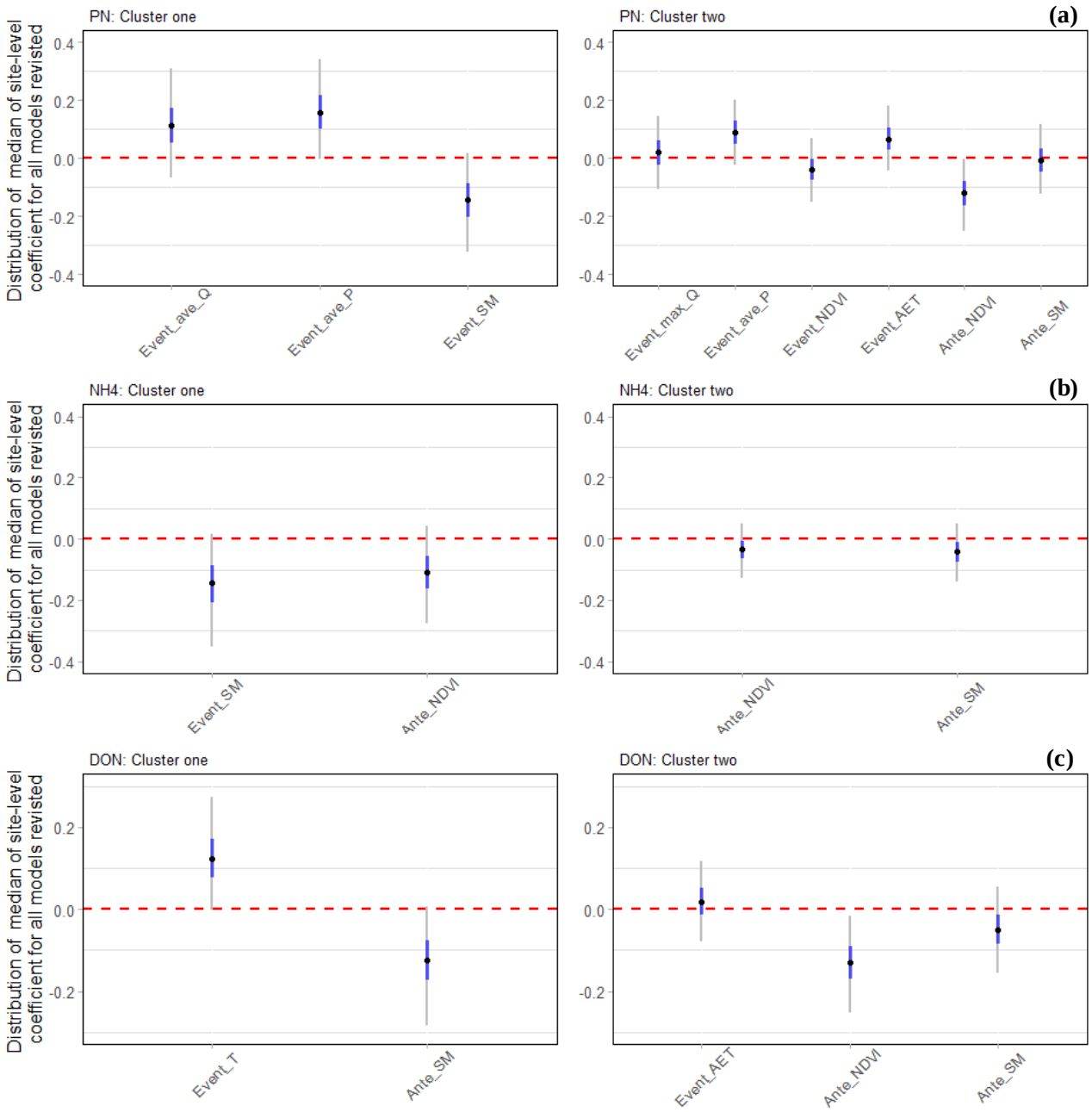


Figure S4: Comparison of BMA model coefficient and cumulative model probability (top 100 models) between two clusters for: (a) DOP, (b) PP and (c) EC. Left - cluster one sites and Right – cluster two sites. The order of predictors on the y-axis was ranked based on the posterior inclusion probability. Each column in the heatmap represents the one specific model (ranked from highest model probability) and the width of the column is normalised by the posterior model probability. The colour indicates the direction of the coefficients, red – negative and blue – positive. Note: the coefficient value was averaged across the posterior median value of site-specific coefficient within each cluster (effect size, $\theta_{n,j}$, in Eq. (6)).



90 **Figure S5: Distribution of median of site-level coefficients for all plausible models in BMA. (a) PN, (b) NH_4 and (c) DON. Only predictors with PIP > 0.8 are included. For each specific model structure, the coefficient value of a predictor was the median of site-specific coefficient across all sites (effect size, $\theta_{n,j,}$ in Eq. (6). The distribution of this value thus represents the probability of the model (PMP), as well as variability in the same predictor across different sites. Note: black dots indicate the median; grey vertical**

lines indicate 95% CI and blue coloured vertical lines indicates 50% CI. The definition of abbreviation of each predictor can be found in Table 3.

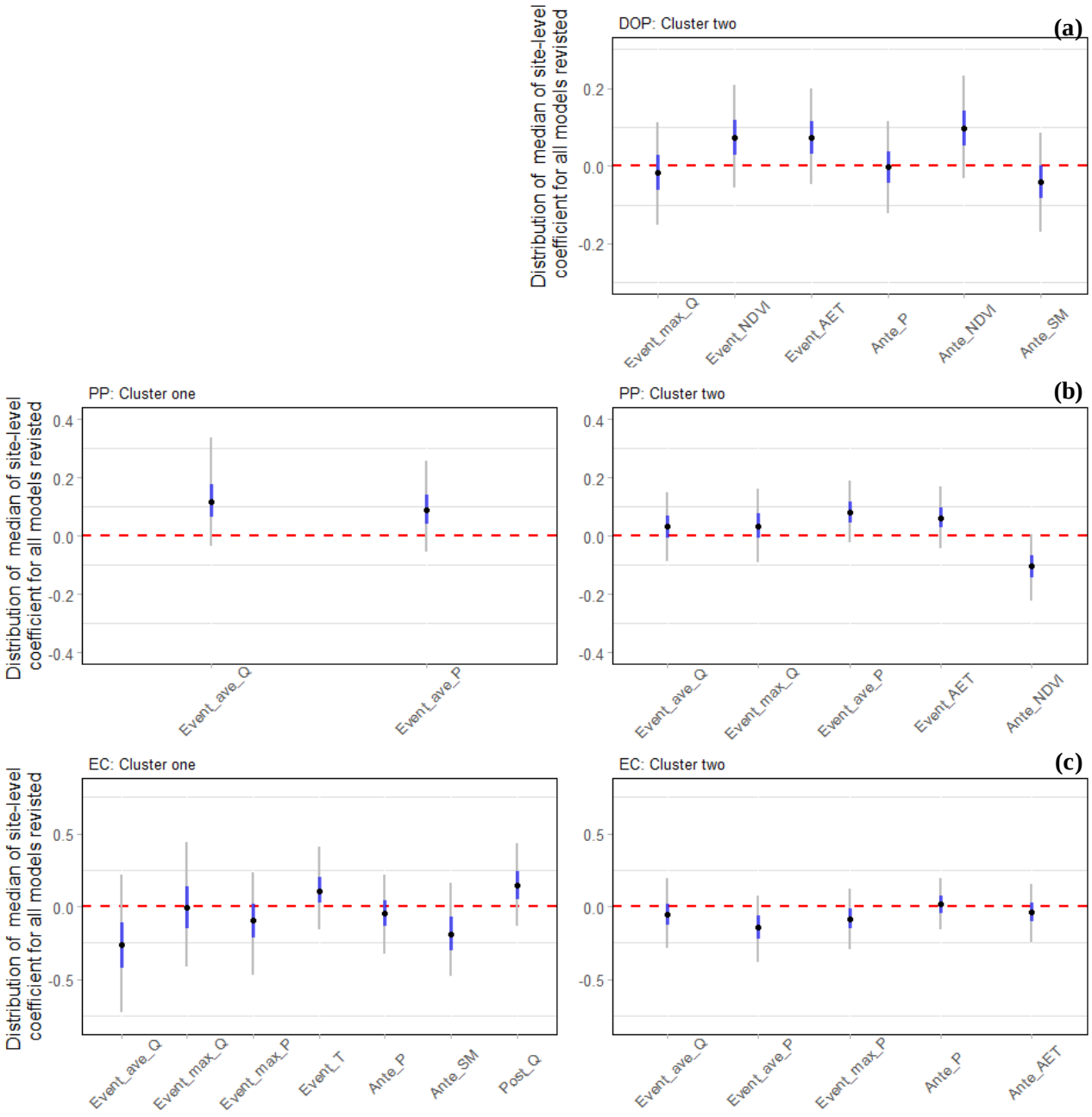


Figure S6: Distribution of median of site-level coefficients for all plausible models in BMA. (a) DOP, (b) PP and (c) EC. Only predictors with PIP > 0.8 are included. For each specific model structure, the coefficient value of a predictor was the median of site-

specific coefficient across all sites (effect size, $\theta_{n,j,}$ in Equation 6). The distribution of this value thus represents the probability of the model (PMP), as well as variability in the same predictor across different sites. Note: black dots indicate the median; grey vertical lines indicate 95% CI and blue coloured vertical lines indicates 50% CI. The definition of abbreviation of each predictor can be found in Table 3.

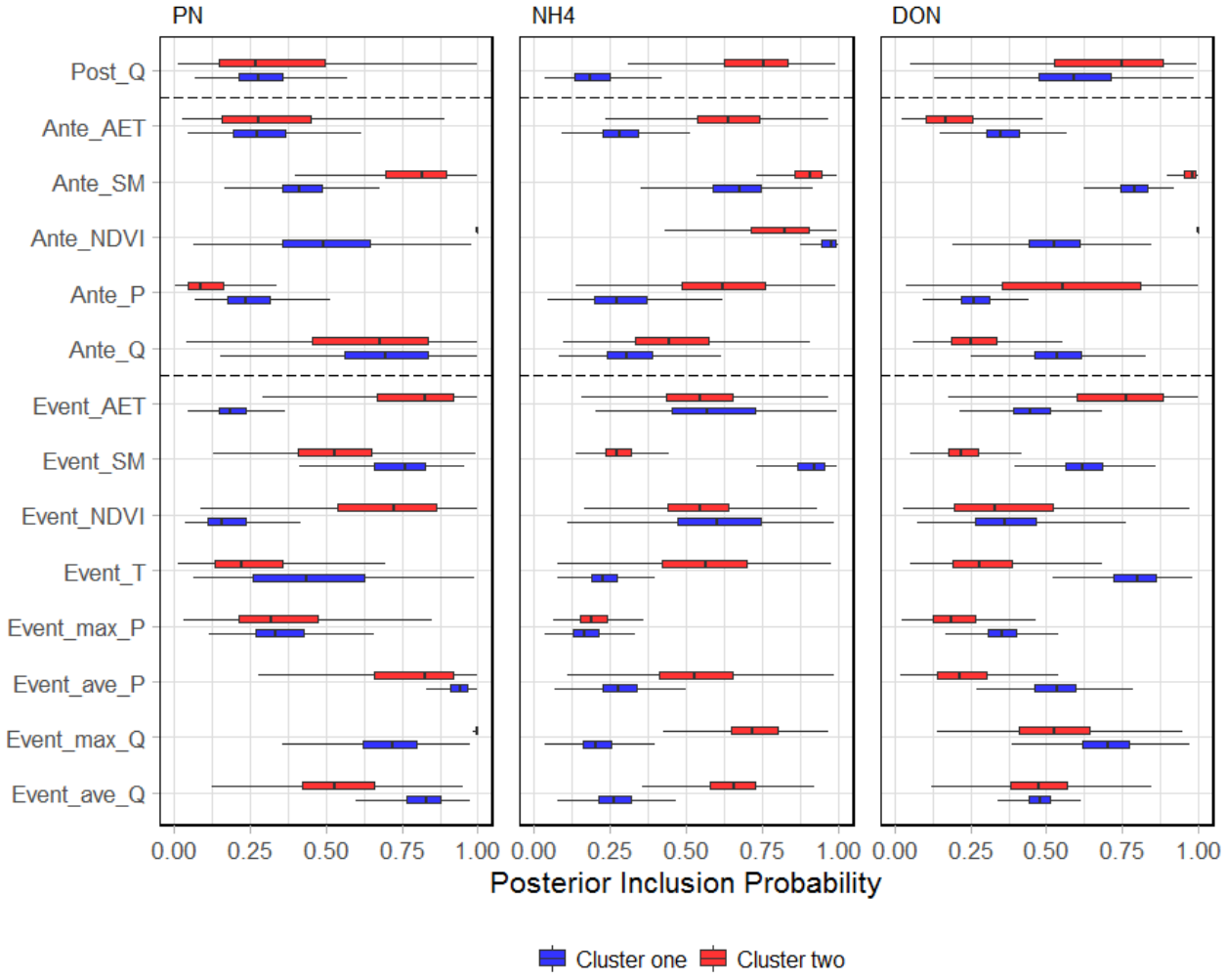


Figure S7: The comparisons of distribution of posterior inclusion probability of individual predictors derived from 1,000 subsampled BMA runs. The interpretation of boxplot is the same as Fig. 7. *Note*: colour represents different clusters: blue - Cluster and red - Cluster two. The definition of abbreviation of each predictor can be found in Table 3.

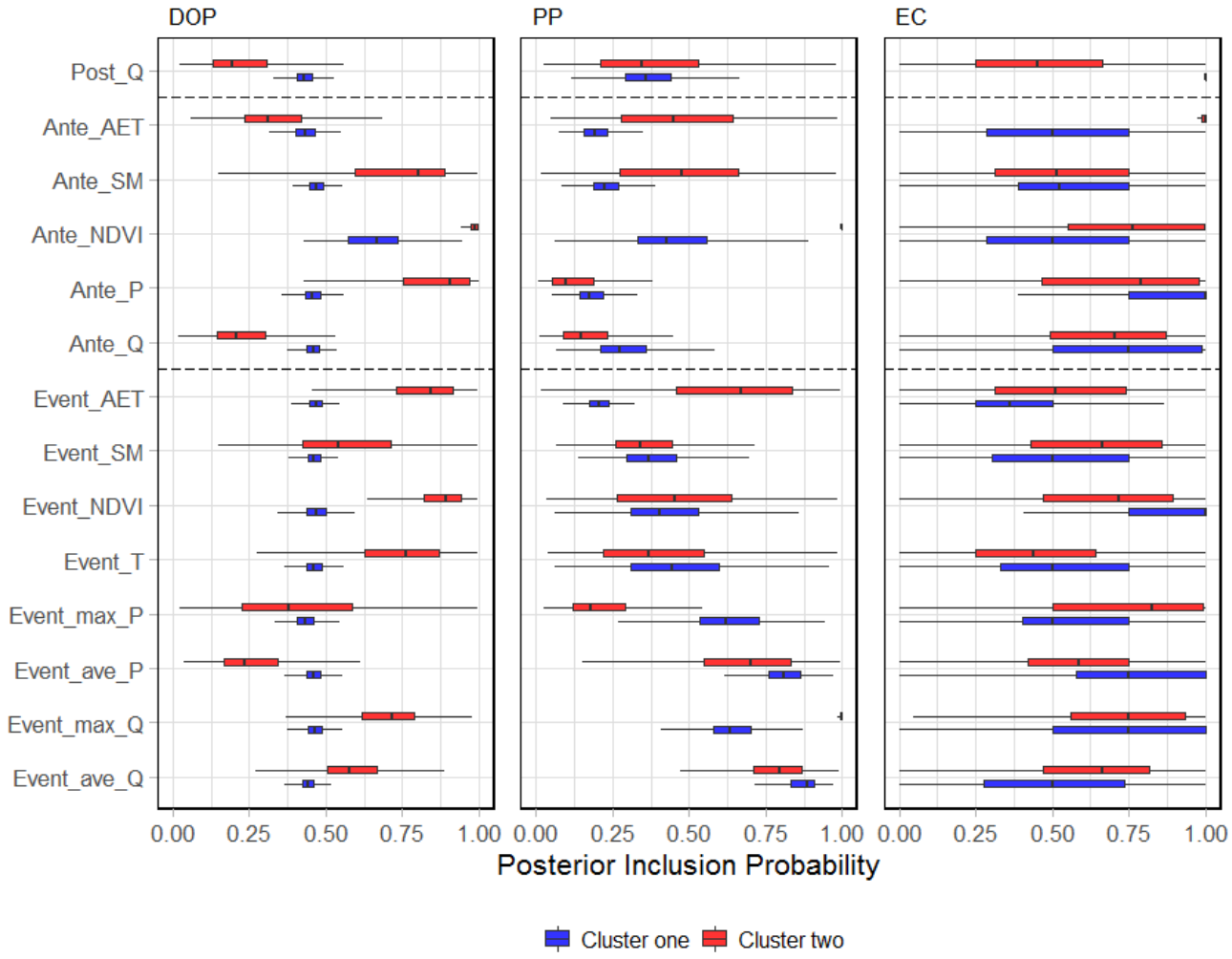


Figure S8: The comparisons of distribution of posterior inclusion probability of individual predictors derived from 1,000 subsampling BMA. The interpretation of boxplot is the same as Fig. 7. *Note*: colour represents different clusters: blue - Cluster and red - Cluster two.

115

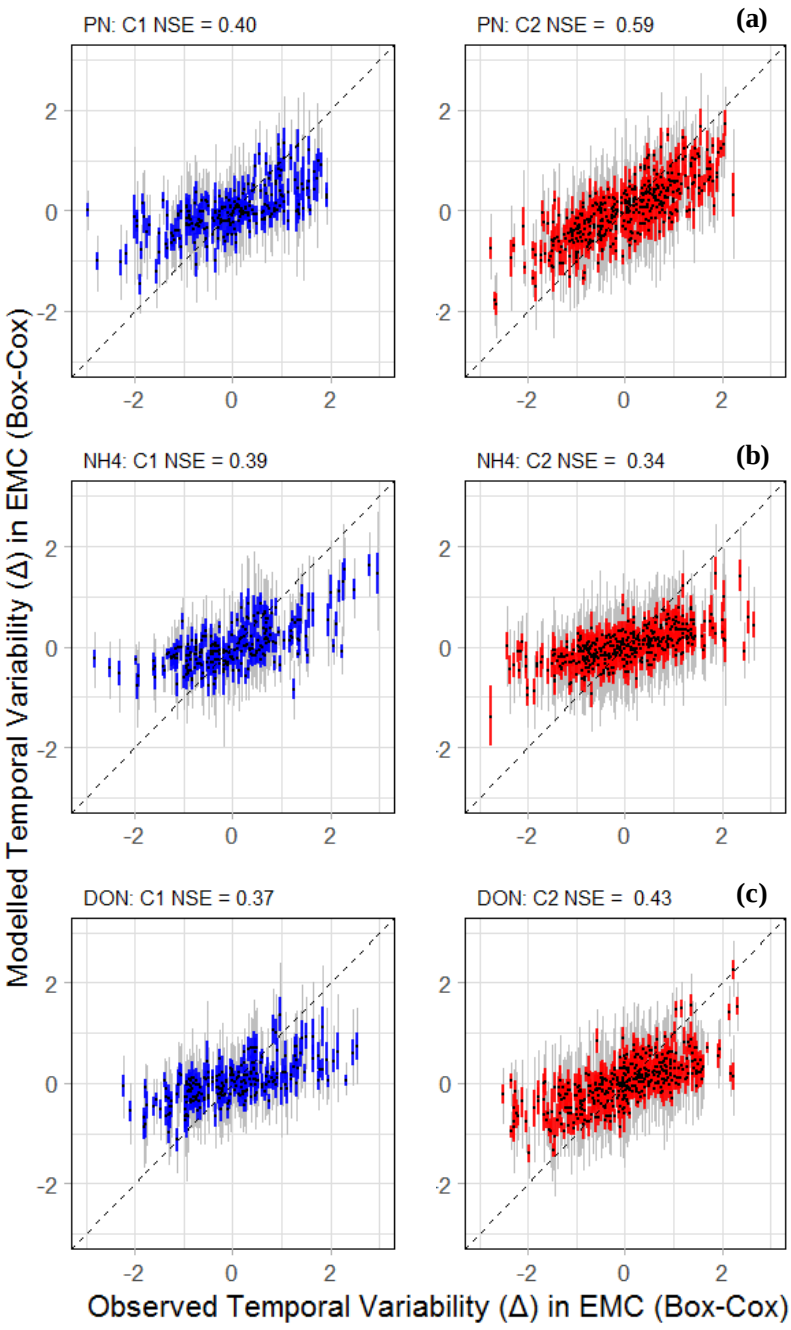


Figure S9: Performance of the BMA models of the temporal variability of nine constituents across 32 sites, represented by prediction intervals from BMA and observed Box-Cox EMC across two clusters of sites for: (a) PN; (b) NH_4 and (c) DON. The NSE values are calculated based on predictions within group- (cluster) level. *Note*: black dots are the prediction median; grey vertical lines are the 95% CI and coloured vertical lines indicates 50% CI: blue - Cluster and red - Cluster two. The dashed black lines is the 1:1 relationship.

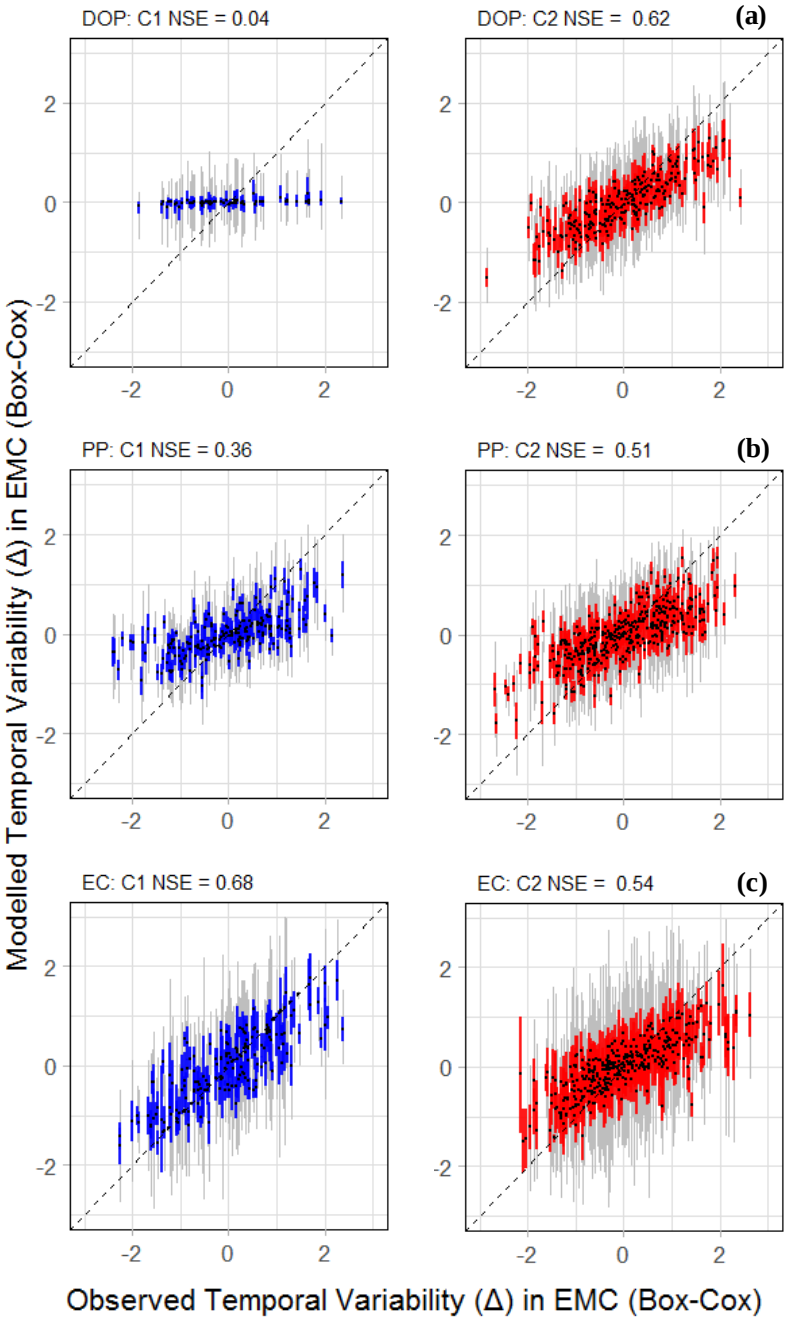


Figure S10: Performance of the BMA models of the temporal variability of nine constituents across 32 sites, represented by prediction intervals from BMA and observed Box-Cox EMC across two clusters of sites for: (a) DOP; (b) PP and (c) EC. The NSE values are calculated based on predictions within group- (cluster) level. *Note*: black dots are the prediction median; grey vertical lines are the 95% CI and coloured vertical lines indicates 50% CI: blue - Cluster and red - Cluster two. The dashed black lines is the 1:1 relationship.

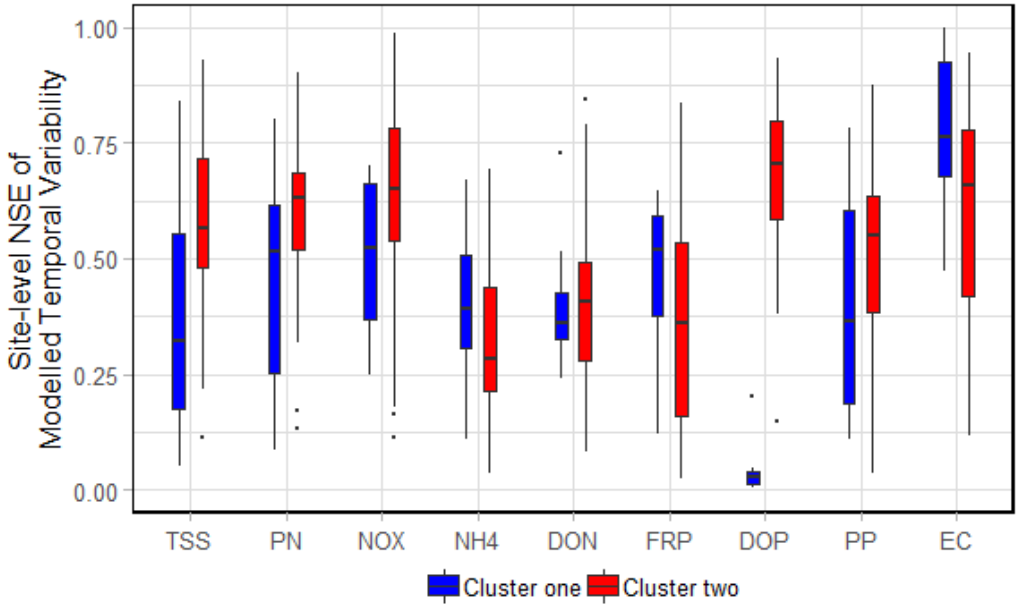
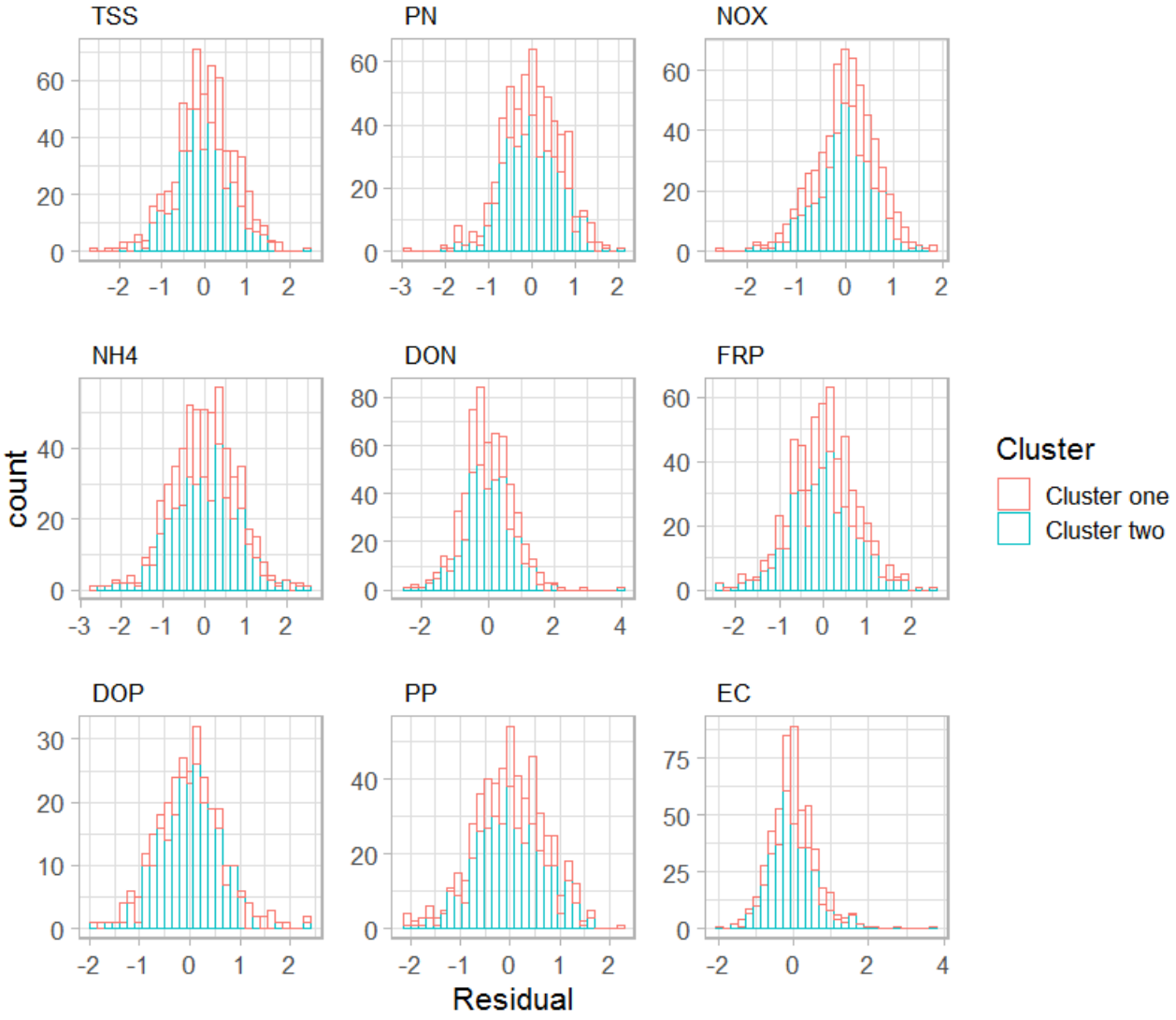
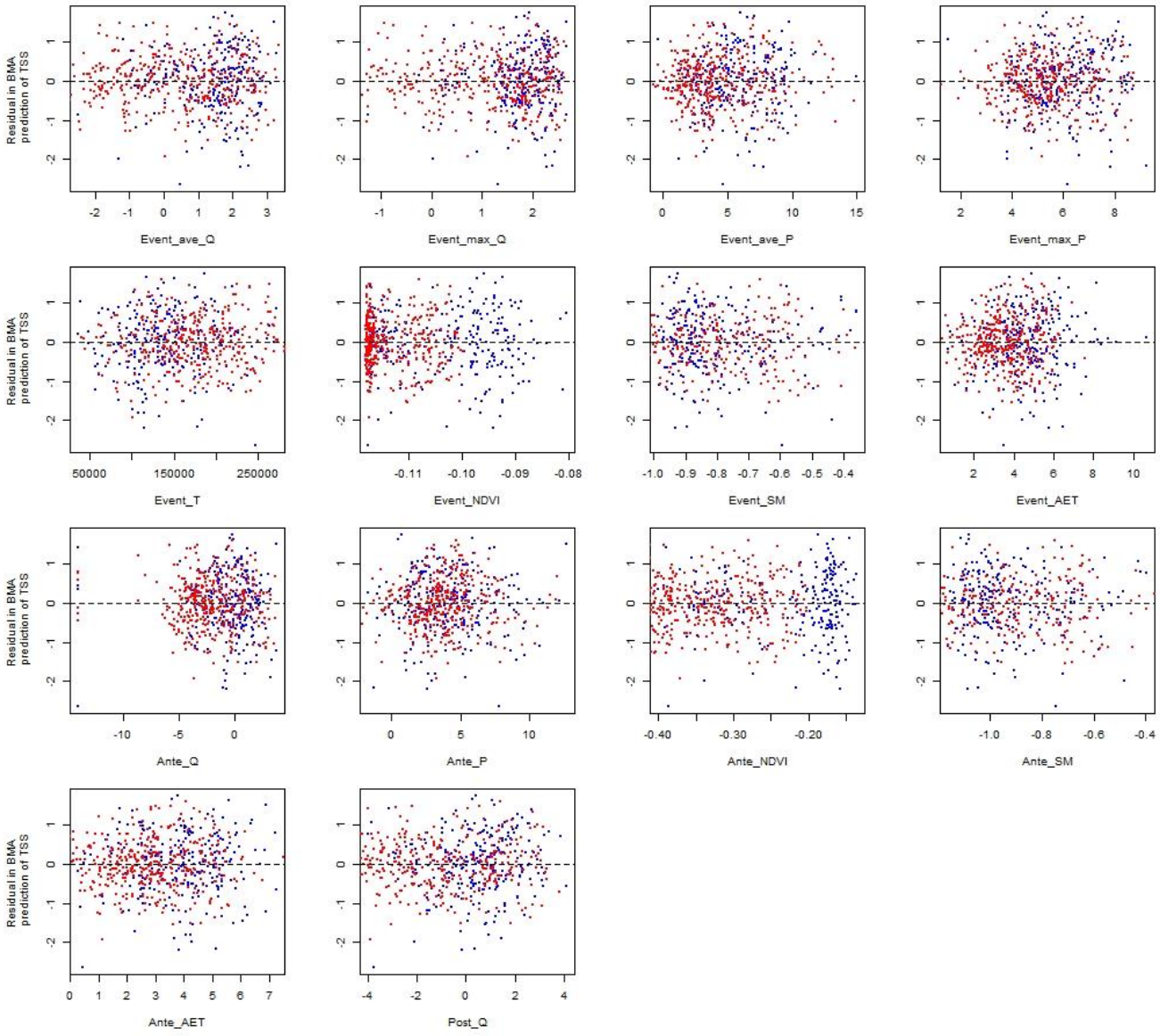


Figure S11: Distribution of site-level NSE for modelled the temporal variability of two clusters of sites. The interpretation of boxplot is the same as Fig. 7. NSE values were calculated based on site-level predictions of event median EMC; blue is Cluster 1("wet"); and red is Cluster 2 ("dry") (i.e., each boxplot is comprised of respective number of sites in each cluster, one for each catchment).



135

Figure S12: Histograms showing distribution of residuals of nine constituents from BMA predictions. Red – Cluster one; Blue – Cluster two.



140

Figure S13: Relationship between residual in median of BMA prediction of TSS and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.

145

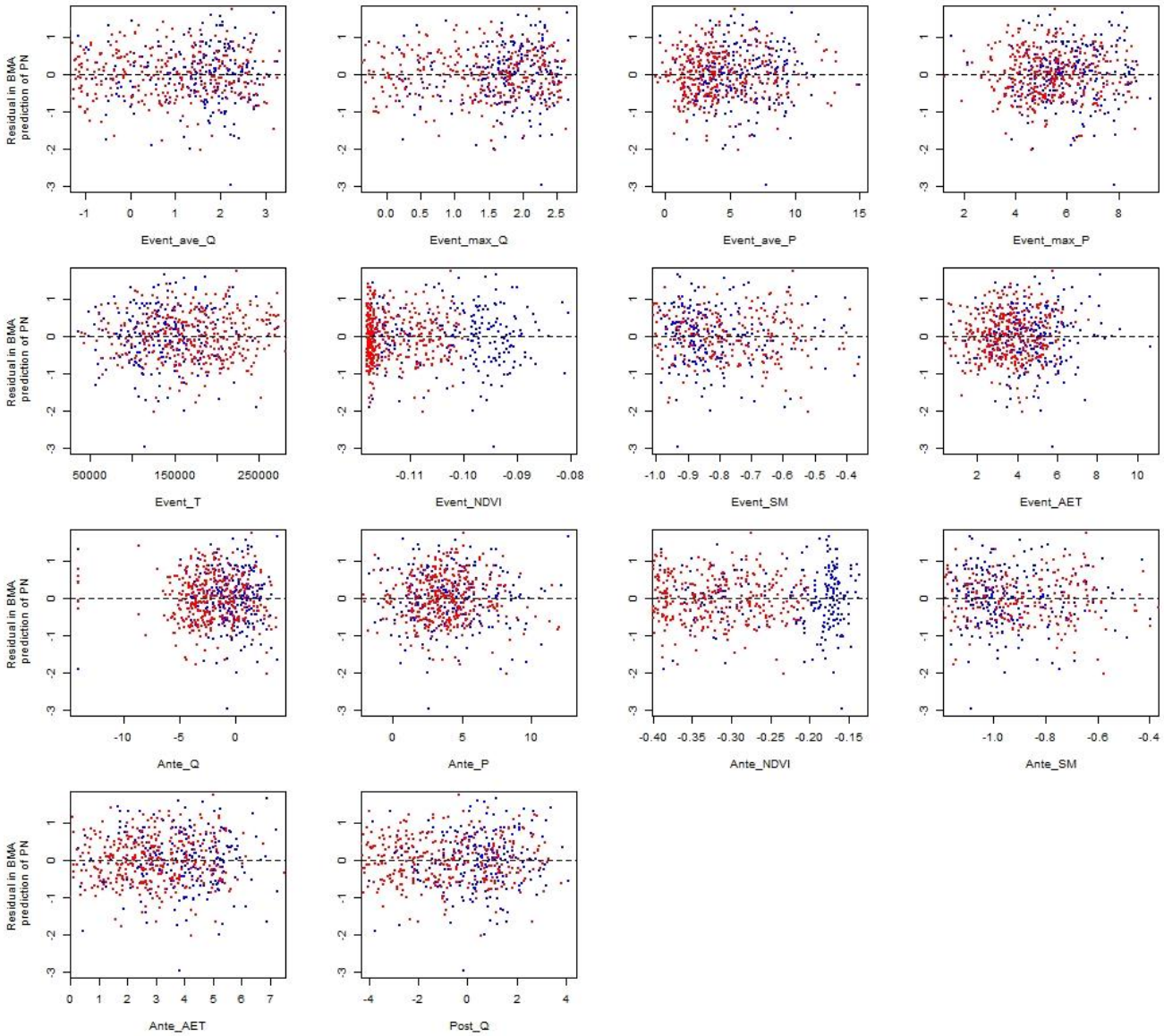


Figure S14: Relationship between residual in median of BMA prediction of PN and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.

150

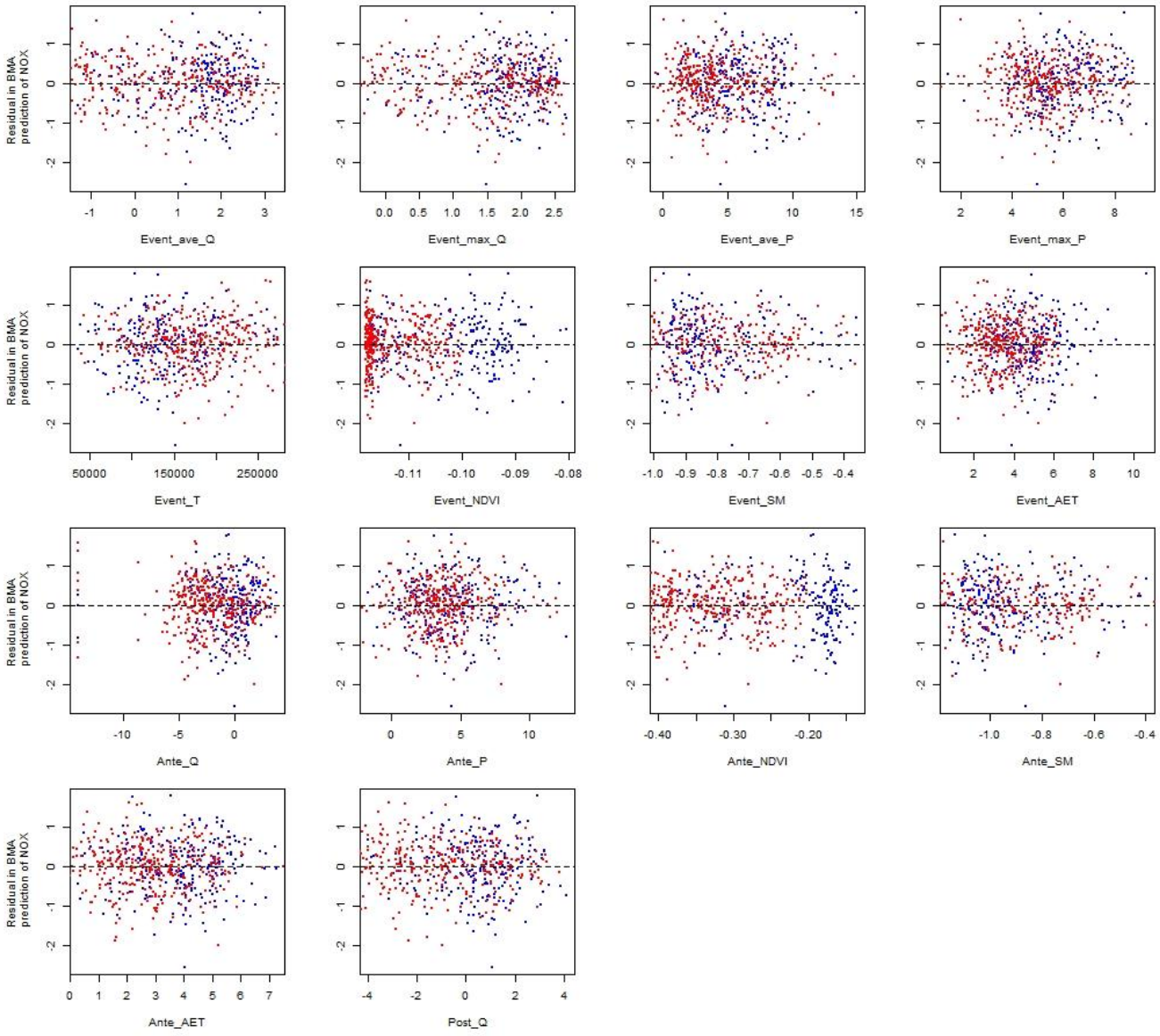


Figure S15: Relationship between residual in median of BMA prediction of NO_X and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.

155

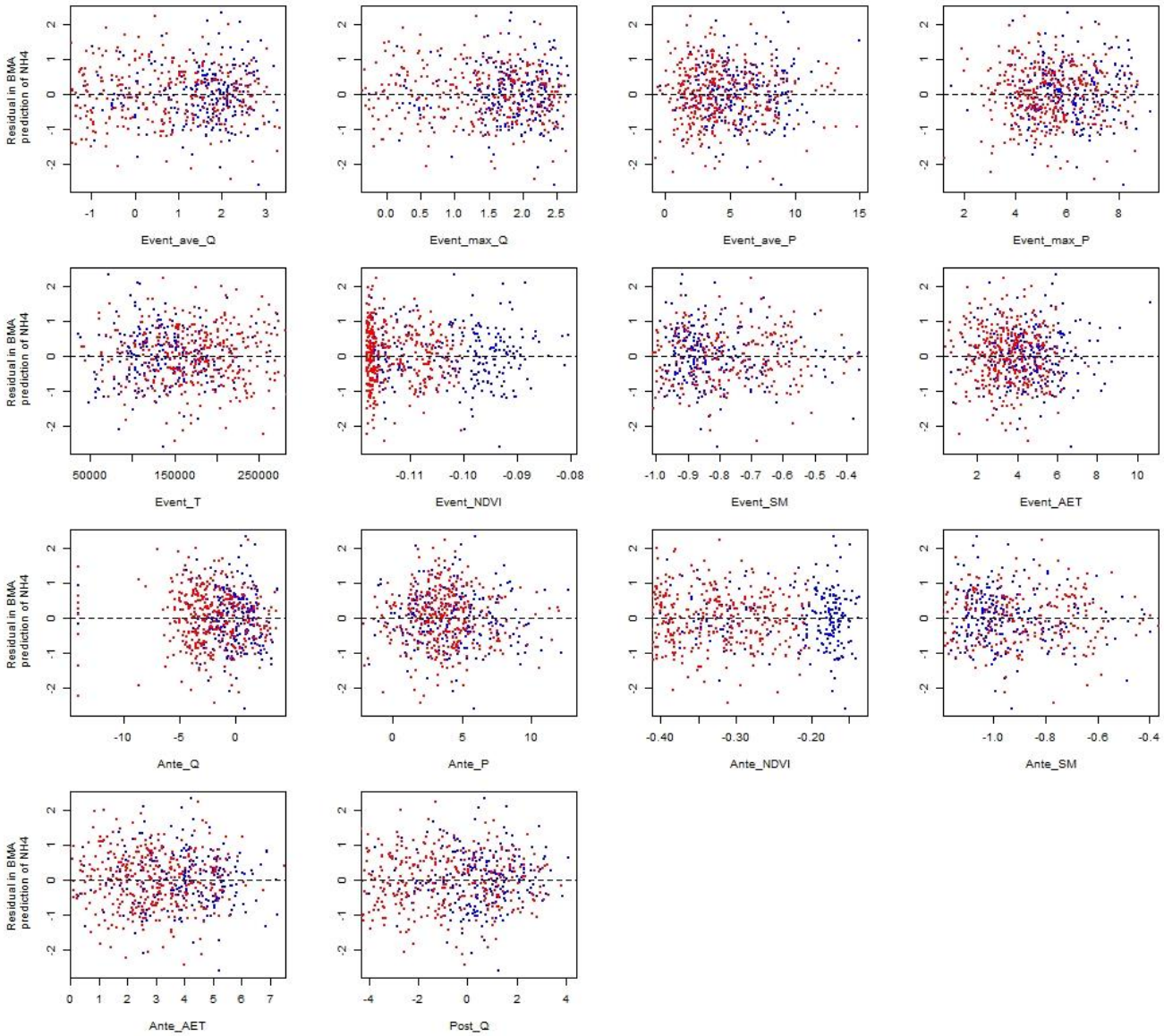


Figure S16: Relationship between residual in median of BMA prediction of NH_4 and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.

160

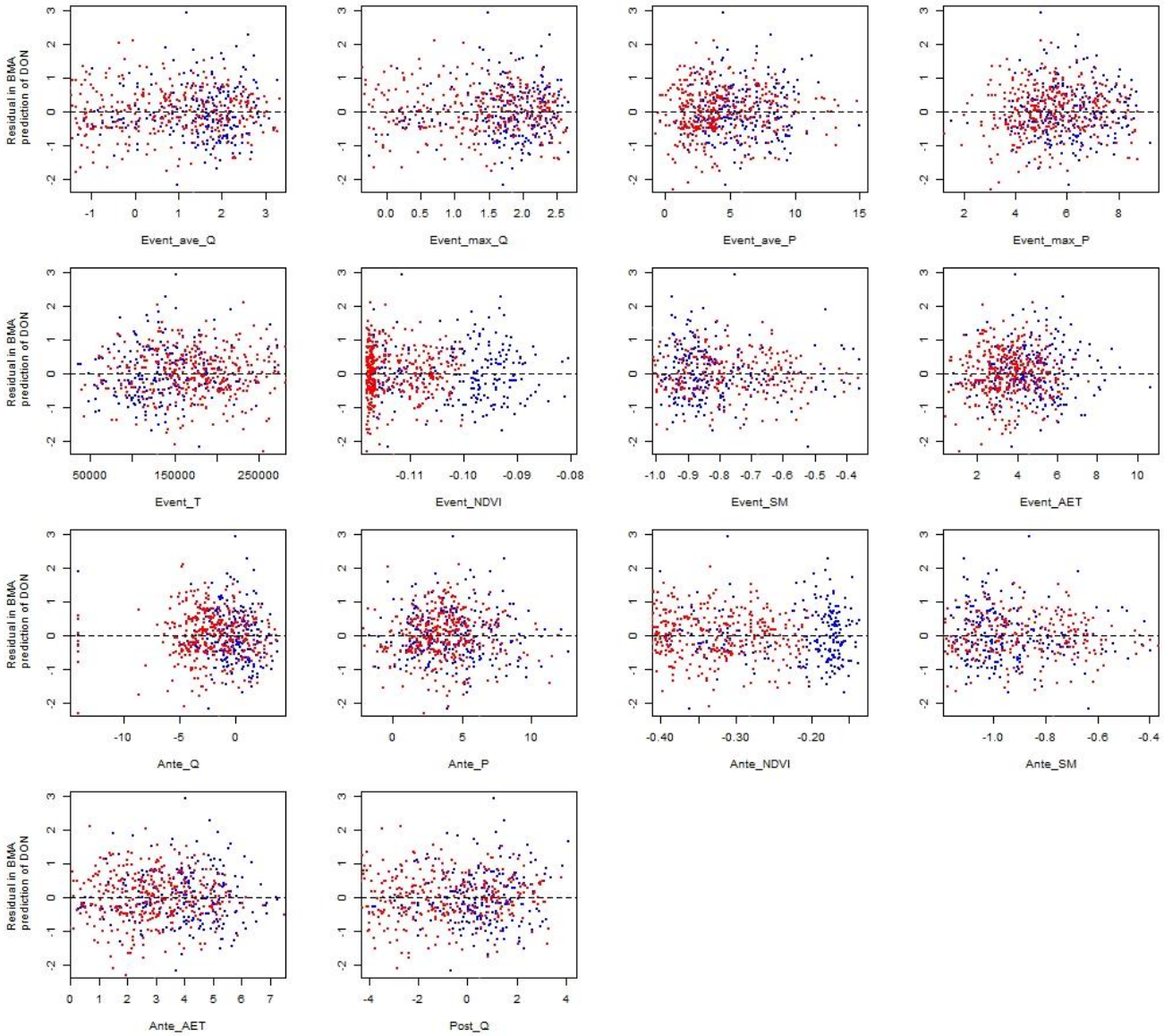
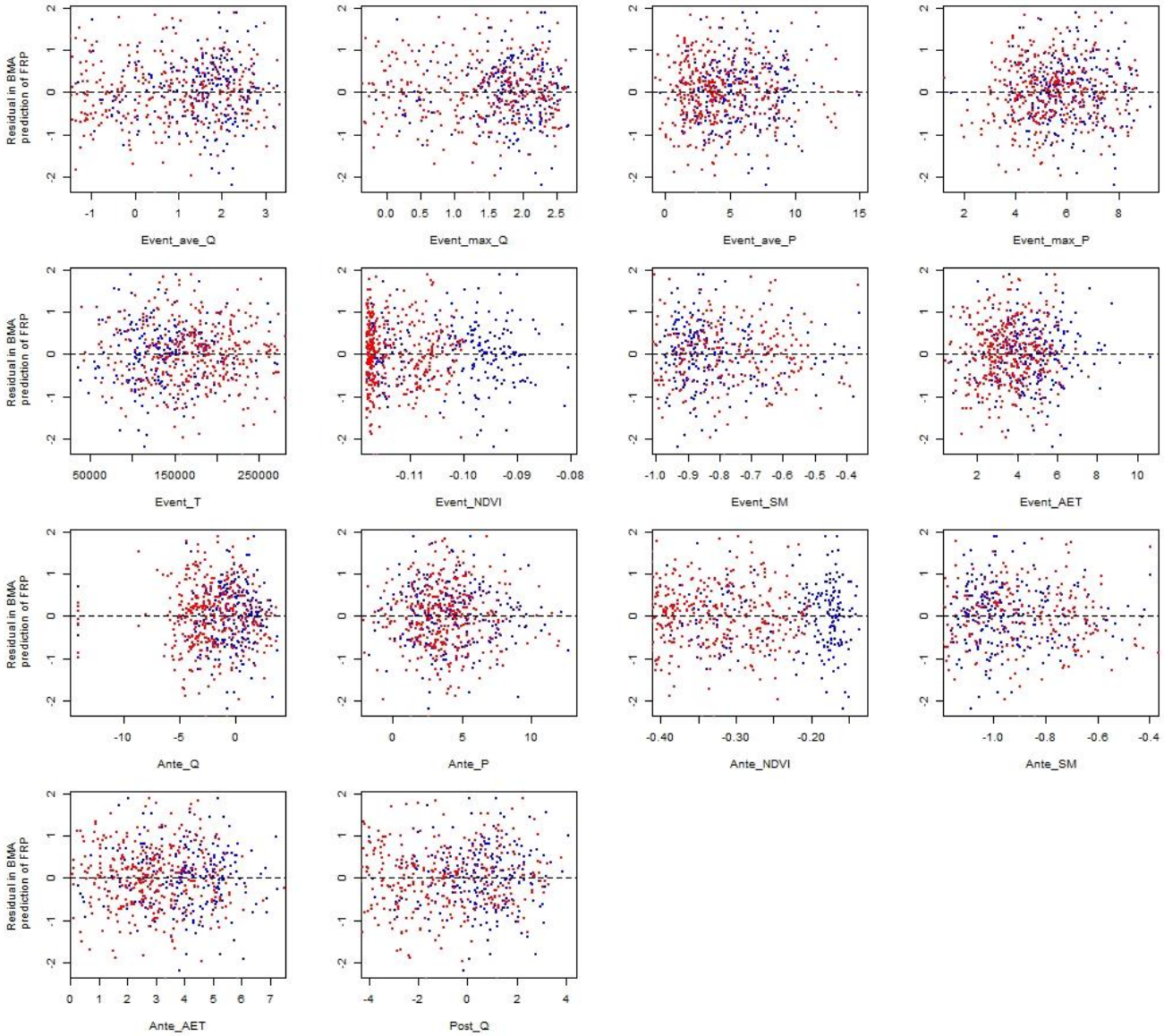


Figure S17: Relationship between residual in median of BMA prediction of DON and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.



170 **Figure S18: Relationship between residual in median of BMA prediction of FRP and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.**

175

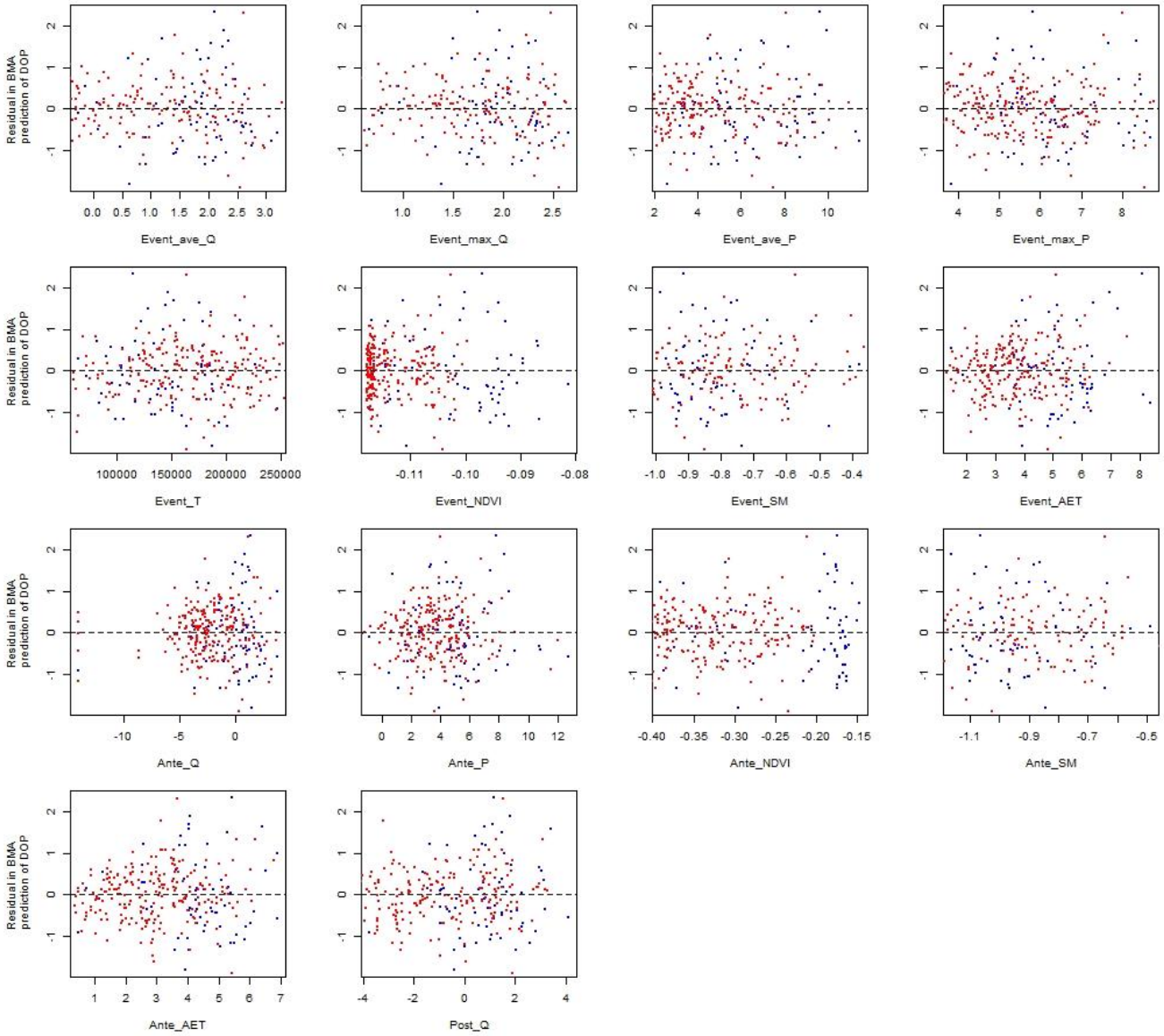


Figure S19: Relationship between residual in median of BMA prediction of DOP and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.

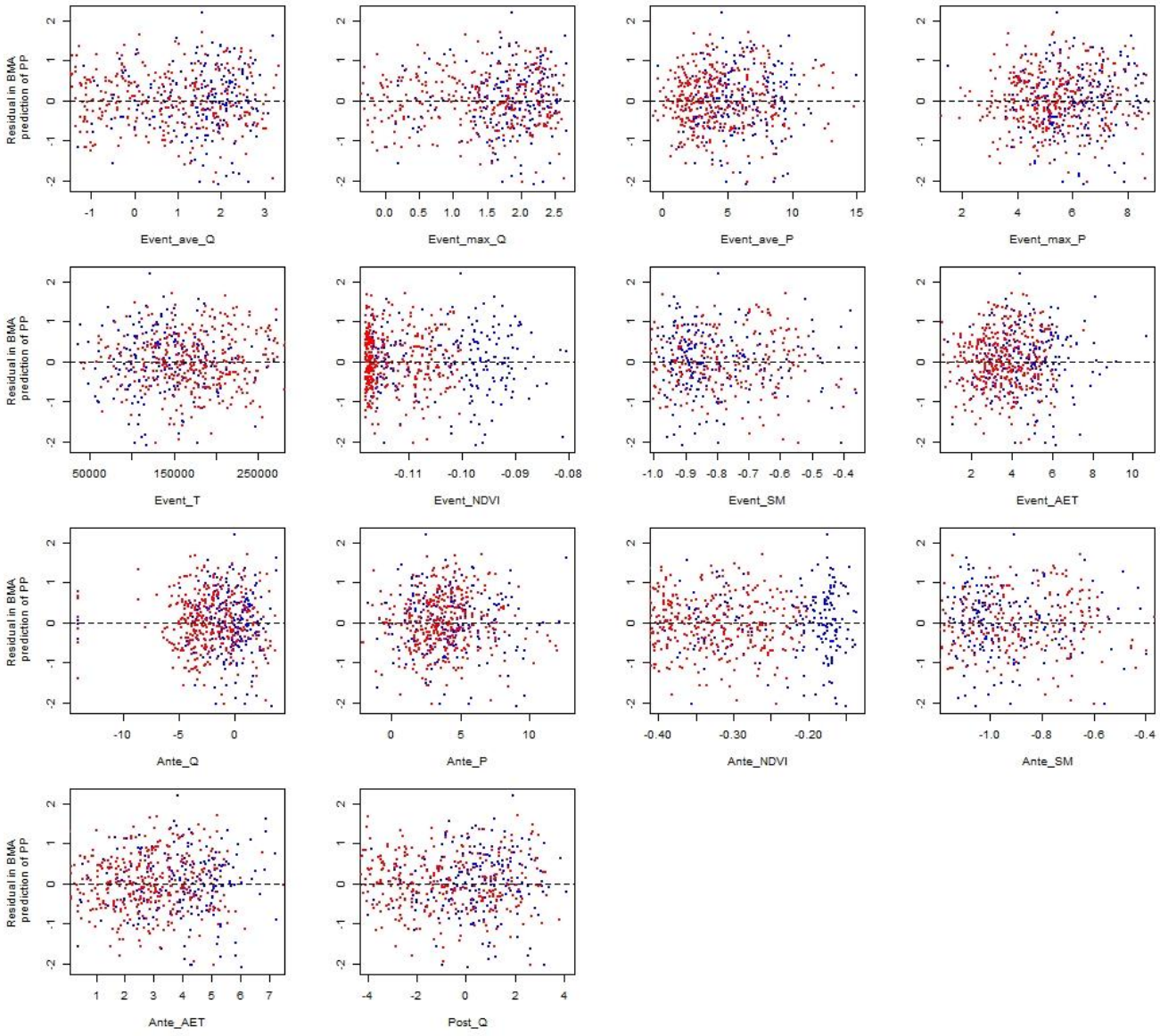


Figure S20: Relationship between residual in median of BMA prediction of PP and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.

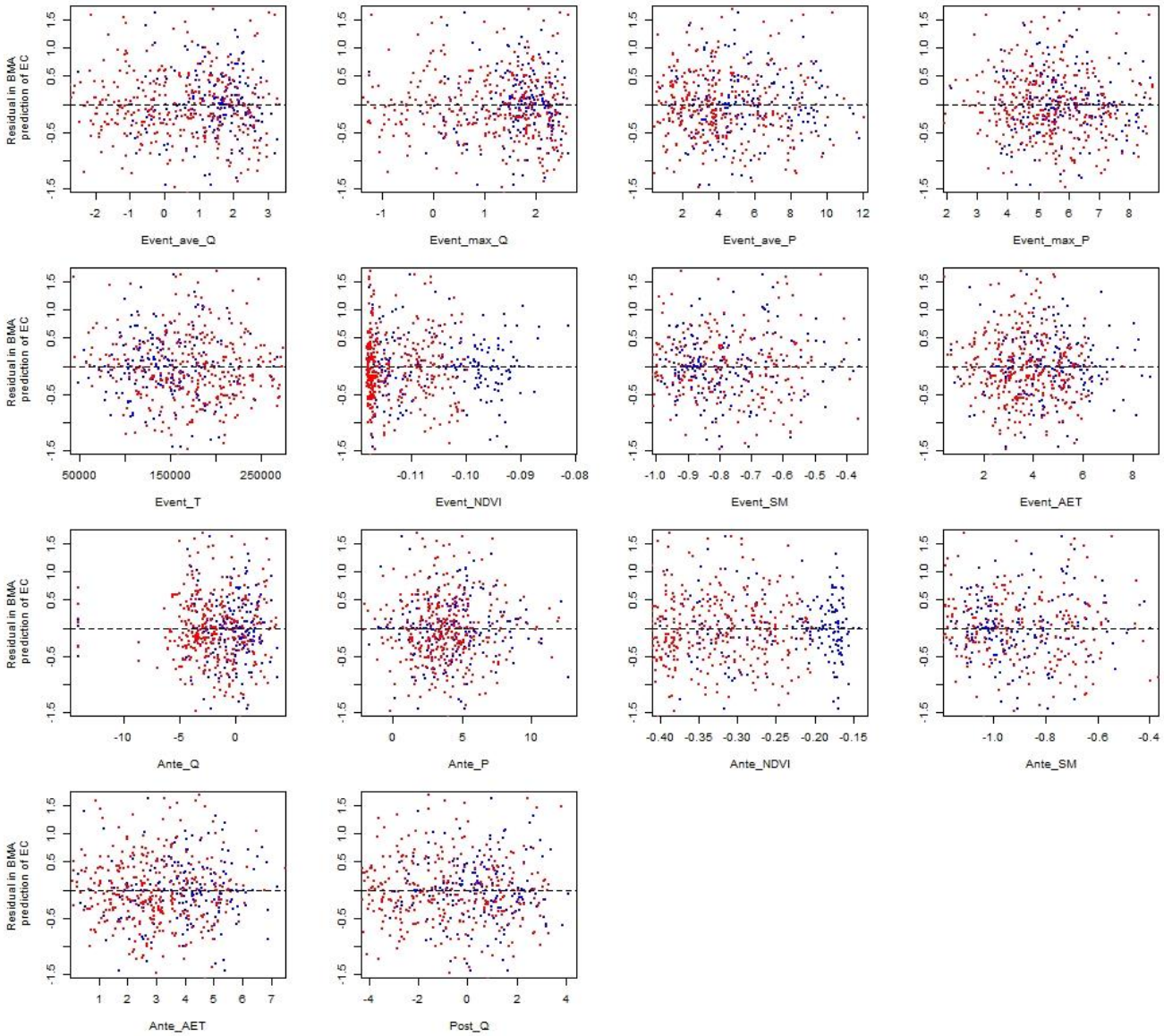


Figure S21: Relationship between residual in median of BMA prediction of EC and 14 candidate covariates in BMA. Note, difference colours indicate two clusters: Red – Cluster one; Blue – Cluster two.

190

195 **Table**

Table S1. Description of 32 sites in the GBR catchments

NRM	Site ID	River and site name	Latitude/°	Longitude/°	Catchment area / km^2
Cape York	105107A	Normanby River at Kalpowar Crossing	-14.9185	144.2100	12934
Wet tropics	110001D	Barron River at Myola	-16.7998	145.6121	1945
Wet tropics	110002A	Barron River at Mareeba	-17.0022	145.4293	836
Wet tropics	110003A	Barron River at Picnic Crossing	-17.2591	145.5386	228
Wet tropics	1110056	Mulgrave River at Deeral	-17.2075	145.9264	785
Wet tropics	1111019	Russell River at East Russell	-17.2672	145.9544	524
Wet tropics	1120049	North Johnstone River at Old Bruce Hwy Bridge (Goondi)	-17.5059	145.9920	959
Wet tropics	112004A	North Johnstone River at Tung Oil	-17.5456	145.9325	925
Wet tropics	112101B	South Johnstone River at Upstream Central Mill	-17.6106	145.9789	400
Wet tropics	113006A	Tully River at Euramo	-17.9936	145.9411	1450
Wet tropics	113015A	Tully River at Tully Gorge National Park	-17.7727	145.6507	482
Wet tropics	116001F	Herbert River at Ingham	-18.6328	146.1427	8581
Burdekin	119101A	Barratta Creek at Northcote	-19.6923	147.1688	753
Burdekin	120001A	Burdekin River at Home Hill	-19.6436	147.3958	129939
Burdekin	120002C	Burdekin River at Sellheim	-20.0078	146.4369	36290
Burdekin	120301B	Belyando River at Gregory Development Rd.	-21.5423	146.8656	35410
Burdekin	120302B	Cape River at Taemas	-20.9996	146.4271	16070
Burdekin	120310A	Suttor River at Bowen Developmental Road	-21.5375	147.0424	10760
Mackay Whitsunday	124001B	O'Connell River at Stafford's Crossing	-20.6526	148.5730	342
Mackay Whitsunday	1240062	O'Connell River at Caravan Park	-20.5664	148.6117	825
Mackay Whitsunday	125013A	Pioneer River at Dumbleton Pump Station	-21.1441	149.0753	1485
Mackay Whitsunday	126001A	Sandy Creek at Homebush	-21.2831	149.0228	326
Fitzroy	1300000	Fitzroy River at Rockhampton	-23.3175	150.4819	139159
Fitzroy	130206A	Theresa Creek at Gregory Highway	-23.4292	148.1514	8485
Fitzroy	130302A	Dawson River at Taroom	-25.6376	149.7901	15850
Fitzroy	130504B	Comet River at Comet Weir	-23.6125	148.5514	16460
Burnett Mary	136002D	Burnett River at Mt Lawless	-25.5447	151.6549	29360
Burnett Mary	136004A	Jones Weir HW	-25.5948	151.2964	21700
Burnett Mary	136014A	Burnett River at Ben Anderson Barrage Head Water	-24.8896	152.2922	32891

Burnett Mary	136094A	Burnett River at Jones Weir (TW)	-25.5948	151.2974	21700
Burnett Mary	136106A	Burnett River at Eidsvold	-25.4023	151.1033	7117
Burnett Mary	138014A	Mary River at Home Park	-25.7683	152.5274	6845

Table S2. Parameters for running the Hydrun toolbox

Filter coefficient	Pass for baseflow separation	Peak discharge threshold	Return ratio	Smooth coefficient	Minimum duration
0.975	3	100	0.01	6	24

200

Table S3. Average number of samples per event for each constituent

TSS	PN	NO_X	NH_4	DON	FRP	DOP	PP	EC
15	14	14	14	14	15	12	14	16

Table S4. Number of EMCs for each constituent

Cluster	TSS	PN	NO_X	NH_4	DON	FRP	DOP	PP	EC
One	225	207	218	217	215	210	66	186	174
Two	381	370	372	370	373	372	231	366	354
% of event monitored	43	41	42	42	42	41	21	39	37

205 **Table S5. Posterior inclusion probability of individual predictor derived from BMA on two clusters of sites.**

Predictor	TSS		PN		NO_X		NH_4		DON		FRP		DOP		PP		EC	
	Cluster one	Cluster two	Cluster one	Cluster two	Cluster one	Cluster two	Cluster one	Cluster two	Cluster one	Cluster two	Cluster one	Cluster two	Cluster one	Cluster two	Cluster one	Cluster two	Cluster one	Cluster two
Event_ave_Q	*0.93*	0.47	*0.87*	0.49	*0.97*	0.56	0.21	0.68	0.47	0.47	0.33	0.38	0.46	0.58	*0.92*	*0.85*	*1.00*	*0.86*
Event_max_Q	0.73	*1.00*	0.76	*1.00*	0.33	0.58	0.14	0.79	0.79	0.60	*0.94*	0.43	0.46	*0.85*	0.66	*0.99*	*1.00*	0.63
Event_ave_P	*0.92*	*0.92*	*0.98*	*0.92*	0.32	0.07	0.25	0.51	0.52	0.16	0.79	*0.96*	0.45	0.24	*0.86*	*0.82*	0.67	*0.90*
Event_max_P	0.24	0.48	0.24	0.29	0.47	0.01	0.10	0.13	0.31	0.17	0.68	0.44	0.41	0.27	0.67	0.15	*1.00*	*0.96*
Event_T	0.07	0.27	0.50	0.21	0.19	*0.98*	0.16	0.58	*0.88*	0.26	*0.86*	*0.90*	0.48	0.78	0.53	0.61	*1.00*	0.64
Event_NDVI	0.03	0.27	0.09	*0.89*	0.77	0.55	0.68	0.62	0.35	0.34	0.38	0.49	0.46	0.97	0.39	0.52	0.75	0.79
Event_SM	0.54	0.21	*0.83*	0.58	*0.99*	0.38	*0.96*	0.21	0.64	0.19	0.59	0.48	0.47	0.66	0.40	0.35	0.33	0.60
Event_AET	0.13	0.07	0.12	*0.90*	0.61	*1.00*	0.68	0.57	0.43	*0.86*	0.57	0.38	0.51	0.87	0.17	*0.81*	0.33	0.10

Ante_Q	0.23	0.18	0.76	0.76	0.15	*0.98*	0.30	0.37	0.56	0.25	0.36	*0.86*	0.47	0.17	0.25	0.12	0.33	0.59
Ante_P	0.20	0.05	0.22	0.06	0.16	0.03	0.25	0.70	0.22	0.75	0.25	*0.88*	0.44	0.98	0.13	0.06	*0.81*	*0.91*
Ante_NDVI	0.23	*1.00*	0.56	*1.00*	*0.86*	*1.00*	*0.99*	*0.89*	0.47	*1.00*	*0.97*	*0.93*	0.67	*1.00*	0.44	*1.00*	0.33	0.61
Ante_SM	0.13	0.74	0.38	*0.90*	0.79	*1.00*	0.63	*0.96*	*0.83*	*0.99*	*0.90*	*0.95*	0.50	*0.89*	0.19	0.59	*1.00*	0.70
Ante_AET	0.09	*0.81*	0.27	0.31	0.31	0.60	0.20	0.72	0.33	0.10	0.42	0.48	0.42	0.30	0.14	0.61	0.33	*1.00*
Post_Q	0.41	0.07	0.21	0.27	0.18	*1.00*	0.16	0.77	0.66	*0.80*	0.17	*0.81*	0.42	0.10	0.32	0.37	*1.00*	0.63

Note: Posterior inclusion probability >= 0.8 in italic.

Table S6. Comparison between BMA performance using rainfall/runoff predictors only and all candidate predictors (full models).

Constituent	NSE for Cluster 1 ("wet")			NSE for Cluster 2 ("dry")		
	Rainfall, runoff only	Full model	% change in NSE	Rainfall, runoff only	Full model	% change in NSE
TSS	0.32	0.35	11	0.42	0.58	38
PN	0.32	0.40	24	0.38	0.59	56
NO_X	0.23	0.49	113	0.32	0.64	101
NH_4	0.00	0.39	/	0.18	0.34	88
DON	0.20	0.37	84	0.20	0.43	117
FRP	0.27	0.45	68	0.26	0.40	56
DOP	0.00	0.04	/	0.22	0.62	181
PP	0.29	0.36	24	0.34	0.51	51
EC	0.41	0.68	66	0.39	0.54	39

Table S7. Performance statistics for nine constituents for the modelled and observed temporal variability, according to Moriasi et al. (2015).

Constituent	Cluster one	Cluster two
TSS	Indicative	Satisfactory
PN	Satisfactory	Satisfactory
NO_X	Satisfactory	Good
NH_4	Satisfactory	Indicative
DON	Satisfactory	Satisfactory
FRP	Satisfactory	Satisfactory
DOP	Indicative	Good
PP	Satisfactory	Satisfactory
EC	Satisfactory	Satisfactory