Supplement of Hydrol. Earth Syst. Sci., 25, 2789–2804, 2021
https://doi.org/10.5194/hess-25-2789-2021-supplement

Hydrology and
Earth System
Sciences
Open Access

Supplement of

Assessing interannual variability in nitrogen sourcing and retention through hybrid Bayesian watershed modeling

Jonathan W. Miller et al.

Correspondence to: Jonathan W. Miller (jwmille7@ncsu.edu)

FIGURES

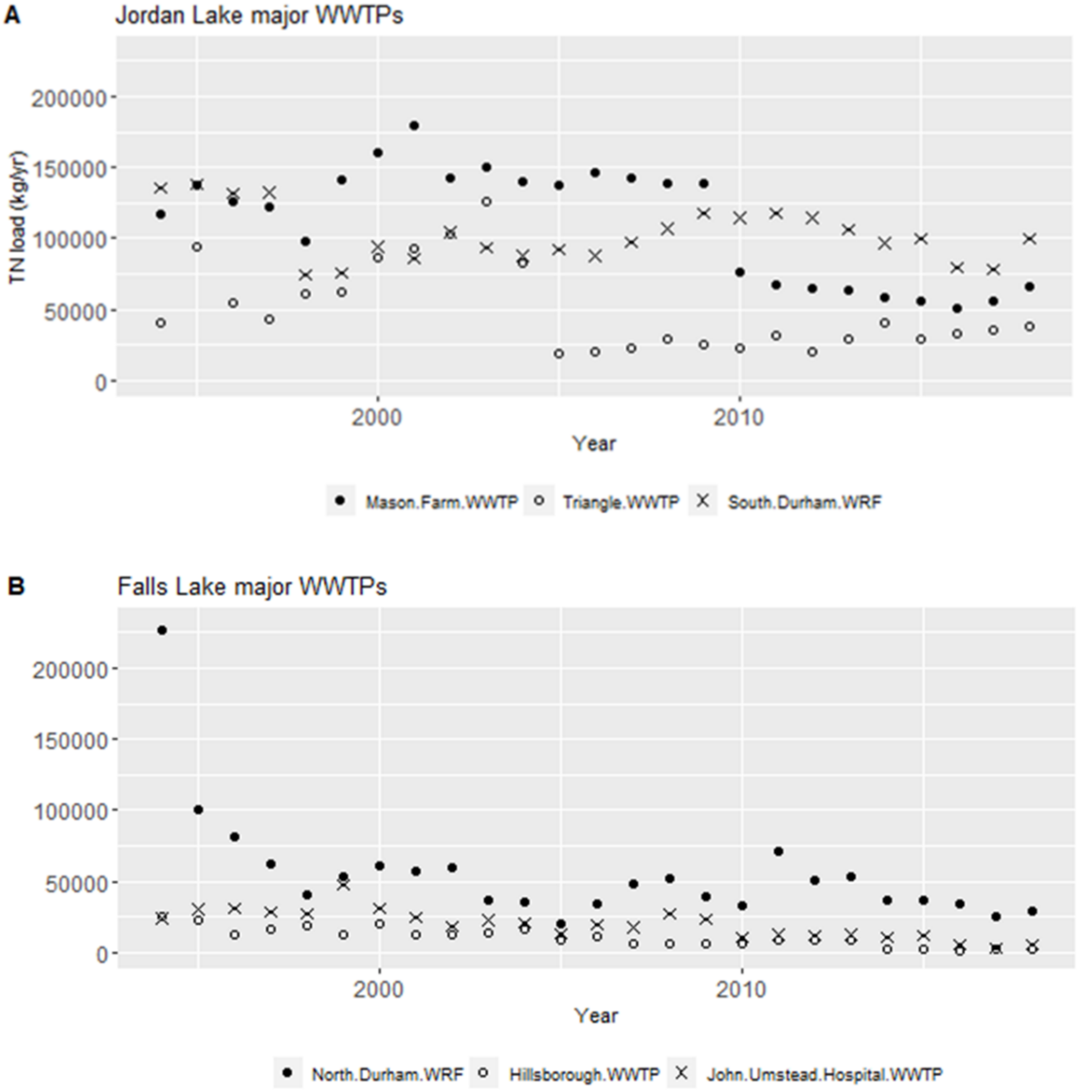

Figure S1: Yearly total nitrogen (TN) loadings from major wastewater treatment plants (WWTPs) located closest to (A) Jordan and (B) Falls Lake.

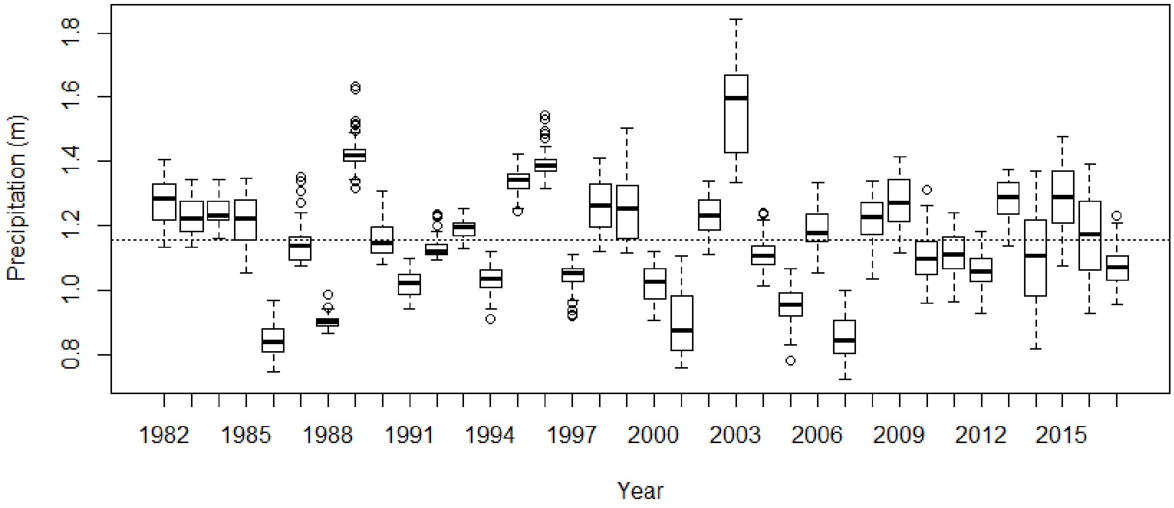

Figure S2: Box and whisker plots of yearly subwatershed precipitation. The horizontal dashed line represents mean yearly precipitation for the Jordan and Falls Lake watersheds.

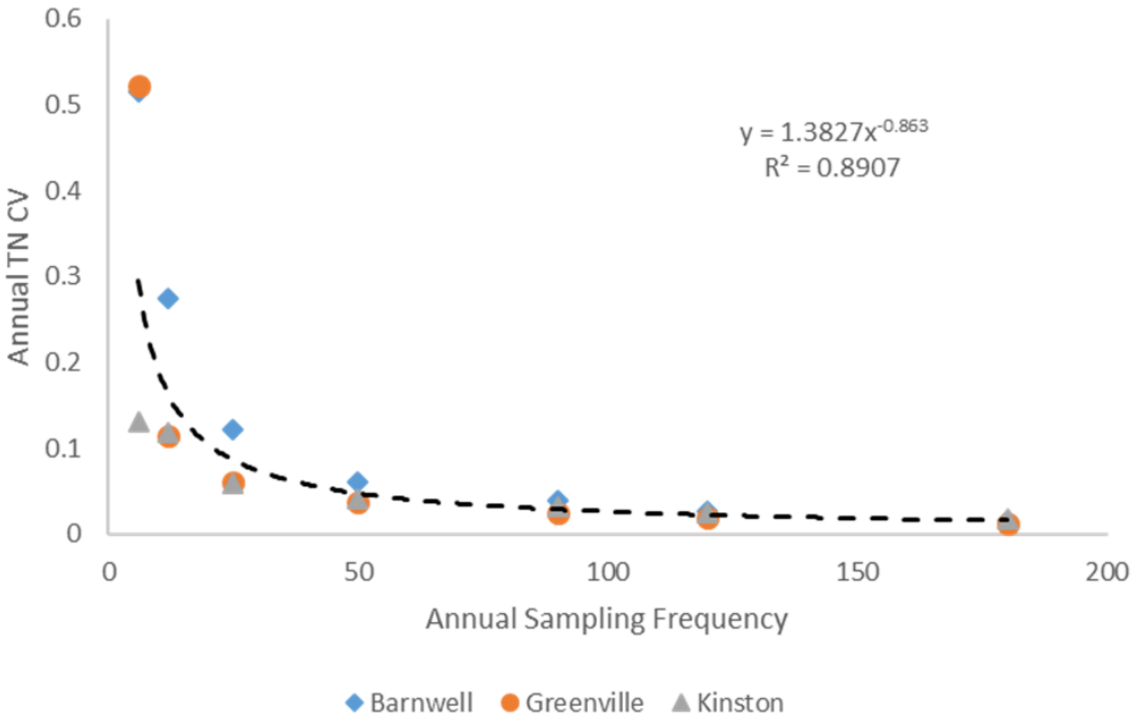

Figure S3: Power relationship (dashed line) quantifying the uncertainty (i.e., coefficient of variation (CV)) of total nitrogen (TN) yearly loadings based on the number of water quality samples available. Yearly loadings are based on Weighted Regression on Time Discharge and Season (WRTDS) estimates.

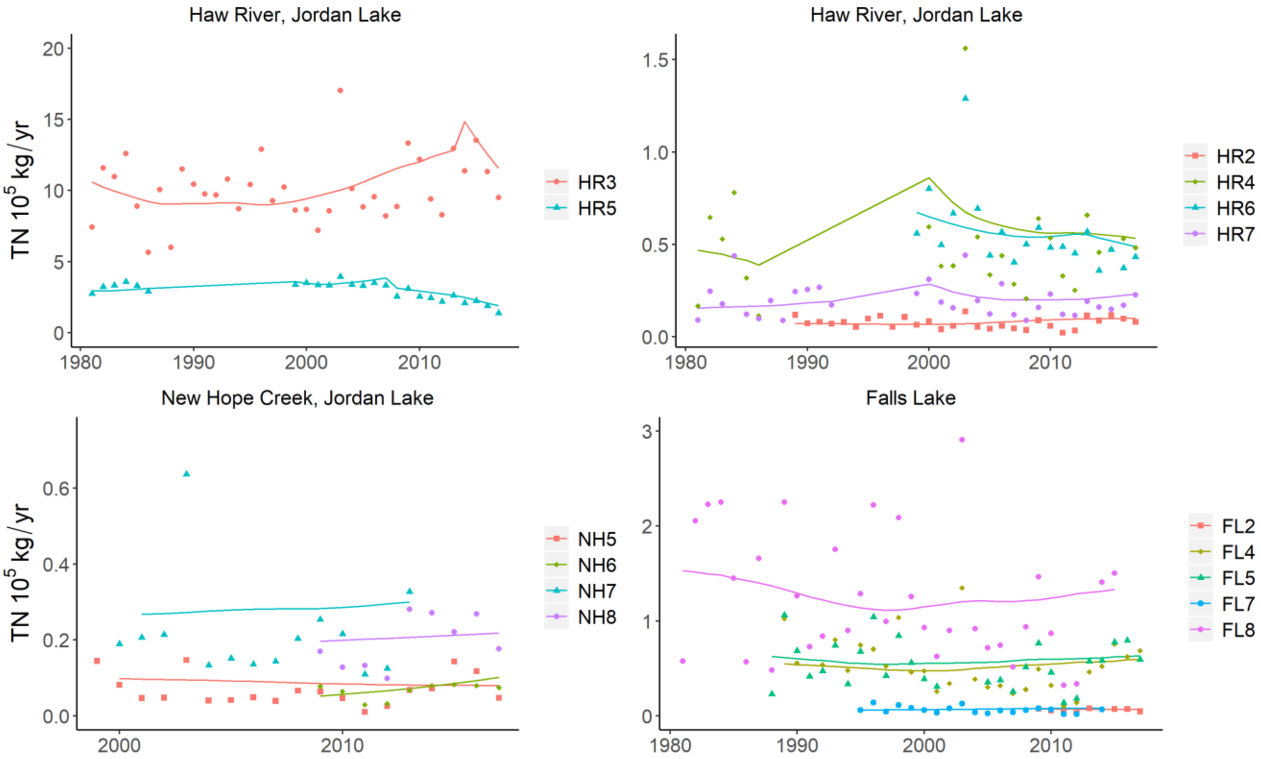

Figure S4: Weighted-regression on Time, Discharge and Season (WRTDS) annual TN loading estimates (points) and flow-normalized estimates (lines) for the Haw River (HR) and New Hope Creek (NH) basins of Jordan Lake (JL) and Falls Lake (FL). These LMSs are not shown in Figure 3.

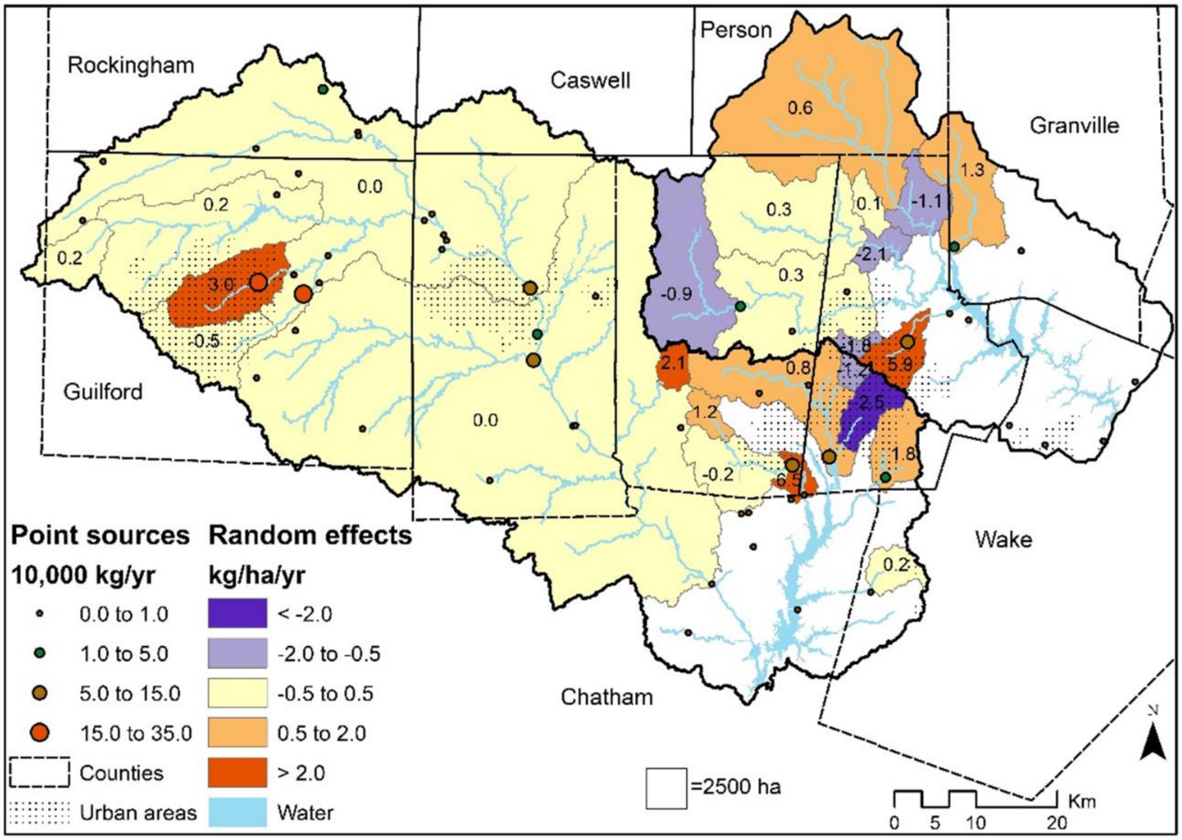

Figure S5: Area-normalized watershed-level random effects for each incremental watershed (i.e., load monitoring site (LMS) watershed). Point sources loads are shown separately as dots.

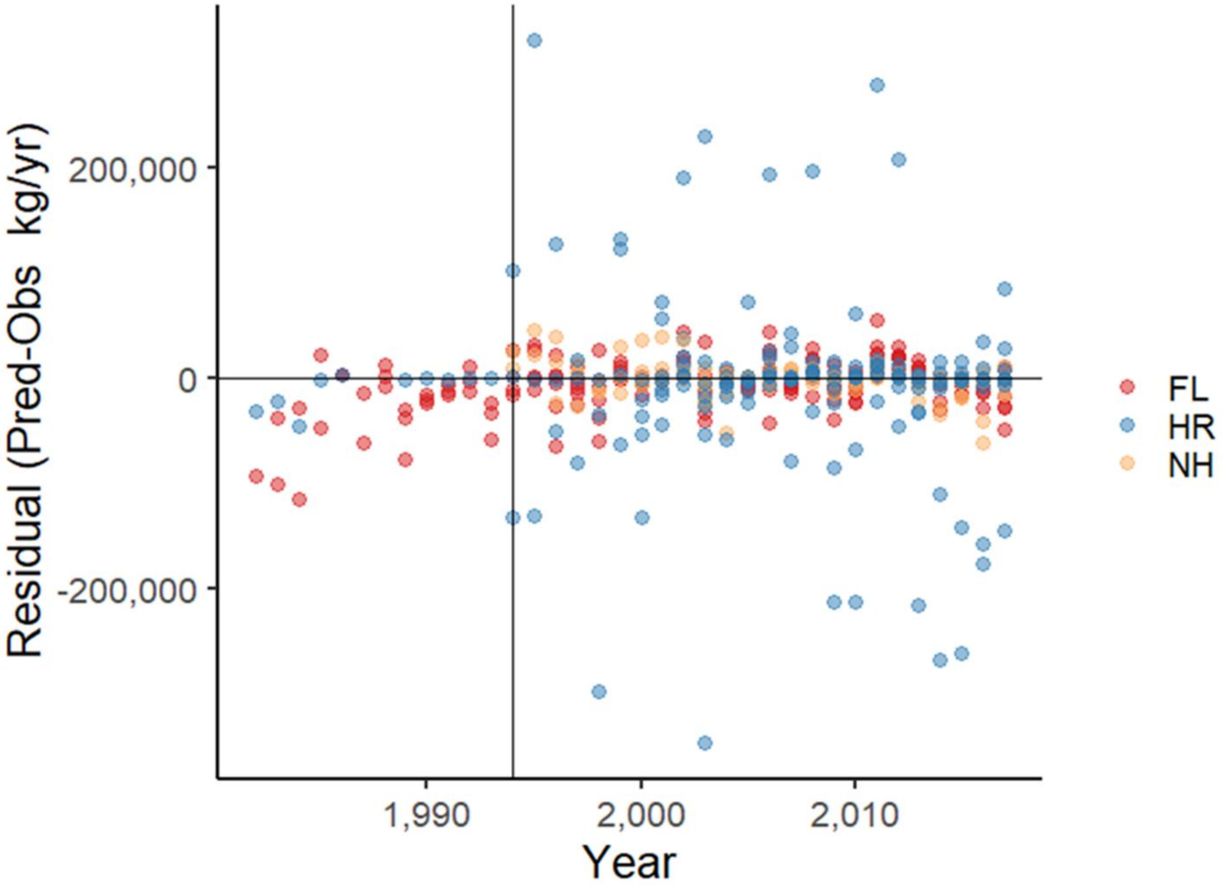

Figure S6: Model residuals plotted by year for the Haw River (HR) and New Hope (NH) watersheds of Jordan Lake and Falls Lake (FL). Load monitoring sites with major point source dischargers were not included before 1994 (solid vertical line).

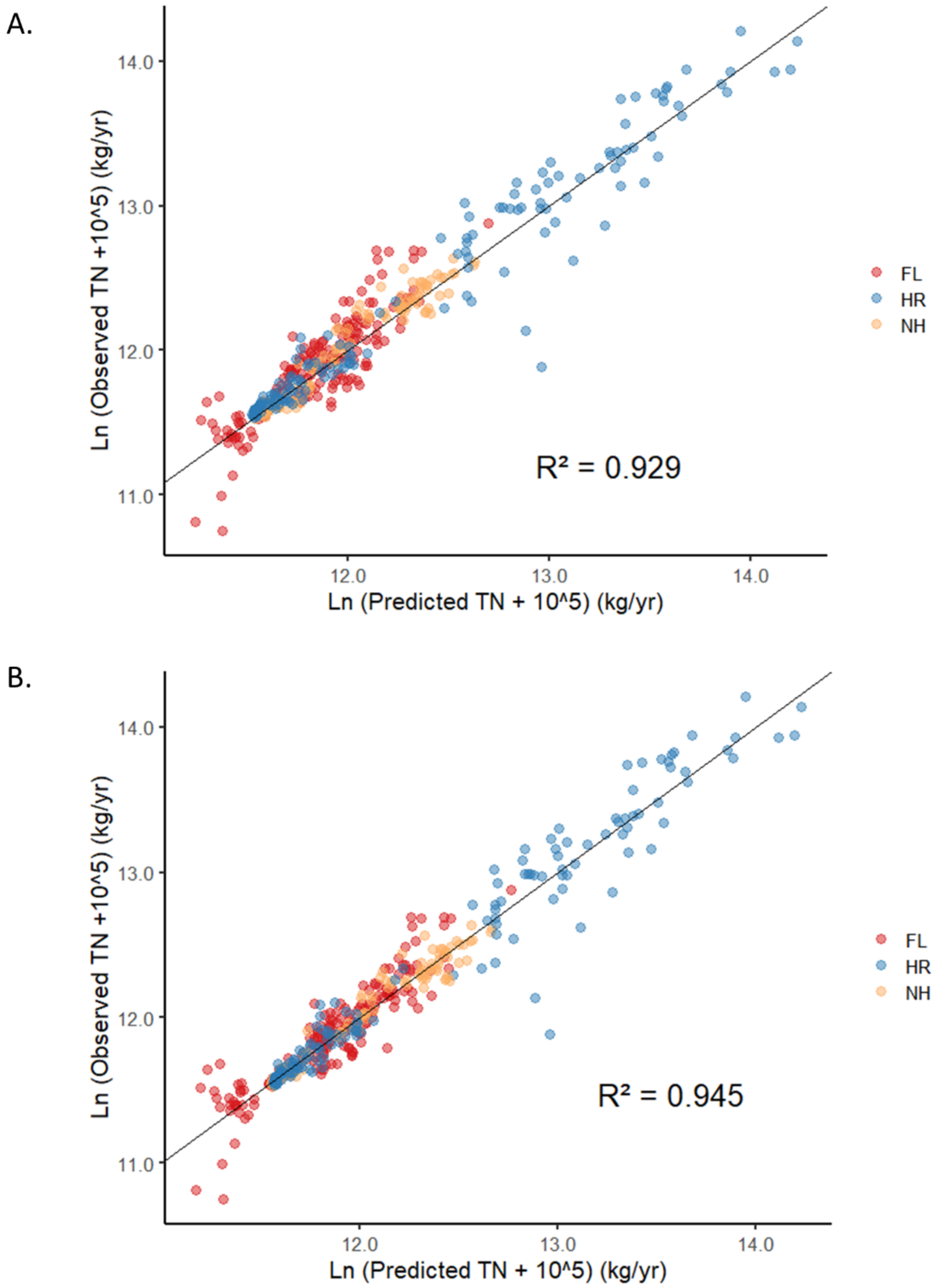

Figure S7: Observed vs. Predicted plots without (A) and with (B) watershed-level random effects for Haw River (HR) and New Hope (NH) watersheds of Jordan Lake and Falls Lake (FL).

TABLES

Table S1: Point source total nitrogen (TN) loads, flows , residence times (RT), and hydraulic loading rates (HL) to downstream loading monitoring sites (LMSs) for major and minor wastewater treatment plants (WWTPs). Flows are shown in million gallons per day (MGD). Data comes from the North Carolina Department of Environmental Quality (NC DEQ).

NC DEQ ID#	Name	TN load (kg) (2017)	Discharge (MGD) (2017)	Downstream LMS	Mean RT (days) to LS	Mean HL (m/yr) to LS
NC0026824	John Umstead Hospital WWTP	3396	1.610	FL10	0.05	
NC0023841	North Durham WRF	25,765	8.280	FL1	0.24	
NC0026433	Hillsborough WWTP	2016	0.919	FL3	1.10	
NC0037869	Arbor Hills Mobile Home Park	18	0.001	FL3	0.55	312.0
NC0056731	Grande Oak Subdivision WWTP	41	0.002	FL3	0.20	107.4
NC0023876	Southside WWTP	43,342	5.872	HR1	1.26	
NC0021211	Graham WWTP	0	0.000	HR1	1.30	
NC0021474	Mebane WWTP	6,177	1.269	HR1	1.76	96.0
NC0035866	Bynum WWTP	122	0.006	HR1	0.02	
NC0042285	Trails WWTP	485	0.020	HR1	1.19	
NC0022675	Birmingham Place	0	0.000	HR1	2.94	
NC0022098	Cedar Valley WWTP	126	0.006	HR1	3.29	
NC0038164	Nathanael Greene Elem. School WWTP	0	0.000	HR1	2.64	
NC0045128	Sylvan Elementary School	172	0.001	HR1	1.78	
NC0045152	Jordan Elementary School	130	0.002	HR1	1.08	
NC0042528	B Everett Jordan 1927 LLC	123	0.008	HR1	1.04	
NC0025241	Mason Farm WWTP	55,655	5.878	NH1	0.20	
NC0056413	Carolina Meadows WWTP	2115	0.148	NH1	0.01	
NC0047597	South Durham WRF	78,164	8.023	NH2	0.55	
NC0042803	Birchwood Mobile Home Park	102	0.007	NH2	1.20	
NC0074446	Hilltop Mobile Home Park	335	0.010	NH2	1.54	

NC0026051	Triangle WWTP	35,557	4.502	NH3	0.14
NC0023868	Eastside WWTP	55,515	3.883	HR3	0.04
NC0024881	Reidsville WWTP	48,902	2.280	HR3	1.50
NC0066966	Quarterstone Farm WWTP	1,438	0.038	HR3	1.28
NC0046019	The Summit WWTP	0	0.000	HR3	1.76
NC0046809	Pentecostal Holiness Church	19	0.000	HR3	1.59
NC0060259	Willow Oak Mobile Home Park	164	0.008	HR3	1.14
NC0077968	Horners Mobile Home Park	127	0.003	HR3	0.52
NC0045161	Altamahaw/Ossipee Elementary School	283	0.003	HR3	0.50
NC0045144	Western Alamance High School	478	0.006	HR3	0.40
NC0031607	Western Alamance Middle School	0	0.000	HR3	0.45
NC0055271	Shields Mobile Home Park	0	0.000	HR3	0.41
NC0065412	Pleasant Ridge WWTP	40	0.008	HR3	1.14
NC0073571	Countryside Manor WWTP	300	0.008	HR3	2.78
NC0046043	Oak Ridge Military Academy	255	0.004	HR3	2.74
NC0022691	Autumn Forest Manuf. Home Community	160	0.018	HR4	0.44
NC0047384	T.Z. Osborne WWTP	582,529	27.680	HR3	1.55
NC0038172	McLeansville Middle School WWTP	0	0.000	HR3	1.55
NC0029726	Guilford Correctional Center WWTP	0	0.000	HR3	1.53
NC0024325	North Buffalo Creek WWTP	161,171	6.619	HR5	0.29

Table S2: Load monitoring sites (LMSs) where Weighted Regressions on Time, Discharge, and Season (WRTDS) estimates were calculated using a break (i.e. wall) in their record if a large sustained change (> 25% up or down) in an upstream point source load occurred.

		Years of record		# of samples		Point source	
LMS	Name	Pre	Post	Pre	Post	Permit	WWTP
NH1	Morgan Creek, JL	1994-2009	2010-2017	360	228	NC0025241	Mason Farm
NH3	Northeast Creek	1995-2004	2005-2017	108	322	NC0026051	Triangle
HR3	Haw River, Burlington	1994-2013	2014-2017	223	56	NC0047384	T.Z. Osborne
HR5	N. Buffalo Creek	1999-2007	2008-2017	148	241	NC0024325	North Buffalo Creek
FL3	Eno River, Durham	1994-2006	2007-2017	158	217	NC0026433	Hillsborough

Table S3: Prior distributions and hierarchical distributions for watershed model parameters. Note, EC is export coefficient, DC is delivery coefficient, PIC is precipitation impact coefficient, and SD is standard deviation. Priors are normal distributions (N) with mean and standard deviation or uniform distributions (U) with upper and lower bounds.

Parameter	Description	Units	Prior	Source
β_{ag}	Agriculture EC	kg/ha/yr	N(9,7)	Dodd et al. 1992
β_{ur1}	Urban type 1 EC	kg/ha/yr	N(8,3)	Dodd et al. 1992
β_{ur2}	Urban type 2 EC	kg/ha/yr	N(8,3)	Dodd et al. 1992
β_{und}	Undeveloped EC	kg/ha/yr	N(2,2)	Dodd et al. 1992
β_{ch}	Chicken EC	kg/an/yr	N(0.001,0.0003)	Strickling and Obenour 2018
β_{h}	Hog EC	kg/an/yr	N(0.04,0.02)	Strickling and Obenour 2018
β_{cw}	Cow EC	kg/an/yr	U(0,5)	-
β_{ps}	Point source DC	-	N(1,0.10)	-
γ_{ag}	Agriculture PIC	-	N(μ_γ, σ_γ)	-
γ_{ur1}	Urban type 1 PIC	-	N(μ_γ, σ_γ)	-
γ_{ur2}	Urban type 2 PIC	-	N(μ_γ, σ_γ)	-
γ_{und}	Undeveloped PIC	-	N(μ_γ, σ_γ)	-
γ_{ch}	Chicken PIC	-	N(μ_γ, σ_γ)	-
γ_{sw}	Swine PIC	-	N(μ_γ, σ_γ)	-
γ_{cw}	Cow PIC	-	N(μ_γ, σ_γ)	-
γ_{ret}	Retention rate PIC	-	N(0,1)	-
κ	Stream decay rate	d^{-1}	N(0.14,0.05)	Hoos and McHahon 2009, Garcia et al. 2011
ω	Reservoir loss rate	m/yr	N(11,2)	Hoos and McHahon 2009, Garcia et al. 2011
σ_ε	Model residual SD	kg/yr	U(0,1x10^6)	-
σ_{LMS}	LMS random effect SD	kg/yr	U(0,1x10^6)	-
μ_γ	PIC mean	-	N(1,1)	-
σ_γ	PIC SD	-	N(0,1)	-

Table S4: Percent of nutrient sources that contributes to downstream reservoirs for normal precipitation years (i.e., middle third of years based on precipitation). In parenthesis are nutrient source percentages during low (lower 33%) and high (upper 67%) precipitation years. Results are for the Haw River (HR) and New Hope (NH) watersheds of Jordan Lake and the Falls Lake (FL) watershed.

	% TN		
Nutrient source	HR	NH	FL
Agriculture	17 (11,25)	1 (1,2)	30 (21,37)
Urban, pre-1980	18 (18,17)	6 (6,5)	30 (34,25)
Urban, post-1980	2 (1,1)	2 (1,1)	5 (4,4)
Undeveloped	6 (5,7)	2 (2,2)	19 (16,19)
Livestock	2 (2,2)	0 (0,0)	1 (2,2)
Discharger	33 (40,28)	11 (15,10)	14 (24,13)

Text S1: RStan code used for pre and post-1980 model

```
data {
int nl;                         //3600 subwatersheds separated by year
int nr;                         //483 incremental watershed-years
int nd;                         //985 dischargers associated with "nr"
vector [nr] load;               //WRTDS load at LMS
vector [nr] SD;                 //SD calculated for WRTDS load
vector [nl] chick;              //chickens in subwatershed
vector [nl] cow;                //cow in subwatershed
vector [nl] hog;                //hogs (swine) in subwatershed
int wsd [nr];                   //count variable for LMSs
int wshed_size;                 //# of LMS watersheds
vector [nr] increm_area;        //Incremental area for each loading station
vector [nr] av_prec;            //normalized precipitation
vector [nr] av_prec2;           //scaled precipitation
vector [nr] up_t_load1;         //upstream TN loading for nested wsds
vector [nr] up_t_load2;         //upstream TN loading for nested wsds
vector [nr] up_t_load3;         //upstream TN loading for nested wsds
vector [nr] str_loss_load1;     //stream losses of upstream TN loading
vector [nr] str_loss_load2;     //stream losses of upstream TN loading
vector [nr] str_loss_load3;     //stream losses of upstream TN loading
vector [nr] res_loss_load1;     //reservoir losses of upstream TN loading
vector [nr] res_loss_load2;     //reservoir losses of upstream TN loading
vector [nr] res_loss_load3;     //reservoir losses of upstream TN loading
vector [nd] d_loss_str;         //stream losses of dischargers to LMSs
vector [nd] d_loss_res;         //reservoir losses of dischargers to LMSs
vector [nd] d_vals;             //dischargers TN loadings
vector [nl] l_loss_str;         //stream losses for subwatersheds
vector [nl] l_loss_res;         //reservoir losses for subwatersheds
vector [nl] ag;                 //agriculture in subwatershed
vector [nl] devpre;             //pre-1980 urban in subwatershed
vector [nl] devpost;            //post-1980 urban in subwatershed
vector [nl] wild;               //undeveloped in subwatershed
vector [nl] tot_l;              //total size of subwatershed
int l_start [nr];               //count variables to link subwatersheds to LMSs
int l_end [nr];                 //count variables to link subwatersheds to LMSs
int d_start [nr];               //count variables to link dischargers to LMSs
int d_end [nr];                 //count variables to link dischargers to LMSs
}

transformed data{
}

parameters {
real<lower =0> Be_a;                          //Agriculture export
real<lower =0> Be_d_pre;                      //pre-1980 developed export
real<lower =0> Be_d_post;                     //post-1980 developed export
real<lower =0> Be_w;                          //Undeveloped export
real<lower =0, upper = 1> Be_ch;              //Chickens export
real<lower =0, upper = 1> Be_h;               //Hogs (swine)
real<lower =0, upper = 1> Be_cw;              //Cows
real <lower =0, upper = 1> Sn;                //Stream retention
real <lower =1, upper = 60> Sn2;              //Reservoir retention
real <lower =0, upper = 0.40> PIC_q;          //PIC for stream flow
```

```
vector<lower =0, upper = 10> [7] pic_p;             //PIC for land classes
real <lower=0> Be_dch;                              //Point source discharge coefficient
real<lower=0, upper = 2> sigma_res;                 //Model residual sigma
real <lower=0, upper = 5> sigma_w;                  //Random effect sigma
vector [wshed_size] alpha;                          //# of watersheds
real<lower = 0, upper = 2> sigma_B1;                //PIC sigma
real<lower = 0, upper = 3> Bp_mean;                 //PIC mean
vector [nr] ly;                                     //log unknown true loads
}

transformed parameters {
}

model {
vector [nr] tot;                   // Sum of all loadings from all sources
vector [nr] sigma;                 //sigma for watershed random effects
vector [nr] y_hat;                 //total loadings plus random effects minus waterbody losses
vector [nr] A;                     //To compile Agriculture with PIC
vector [nr] Dpr;                   //To compile pre-1980 urban with PIC
vector [nr] Dpt;                   //To compile post-1980 urban with PIC
vector [nr] W;                     //To compile undeveloped urban with PIC
vector [nr] Dch;                   //To compile dischargers with stream/reservoir losses
vector [nr] tot_loss;
int w;
vector [nr] alpha_vals;            // Watershed indicator
real t;

vector [nr] A;                     //Agriculture vector
vector [nr] D_lc_pre;              //pre-1980  vector
vector [nr] D_lc_post;             //post-1980  vector
vector [nr] W_lc;                  //Undeveloped  vector
vector [nr] Disch;                 //point source dischargers
vector [nr] C;                     //chickens for adding PIC
vector [nr] H;                     //hogs for adding PIC
vector [nr] Cw;                    //cows for adding PIC
vector [nr] C_r;                   //chickens for aggregating subwatersheds
vector [nr] H_r;                   //swine for aggregating subwatersheds
vector [nr] Cw_r;                  //cows  for aggregating subwatersheds

vector [nr] ly_hat;                        // log of TN load with transformation and offset
vector [nr] y;                             //TN load

// Loop to determine export for each watershed-year
for(i in 1:nr){

        //Looping to aggregate subwatersheds and corresponding stream and reservoir retention.
        A_lc[i]= sum((ag[l_start[i]:l_end[i]]) .* (exp((-Sn/(1+PIC_q*av_prec[i])) *
        l_loss_str[l_start[i]:l_end[i]])) .* (exp((-Sn2/(1+PIC_q*av_prec[i])) ./l_loss_res[l_start[i]:l_end[i]])));
        D_lc_pre[i]= sum((devpre[l_start[i]:l_end[i]]) .* (exp((-Sn/(1+PIC_q*av_prec[i])) *
        l_loss_str[l_start[i]:l_end[i]])) .* (exp((-Sn2/(1+PIC_q*av_prec[i])) ./l_loss_res[l_start[i]:l_end[i]])));
        D_lc_post[i]= sum((devpost[l_start[i]:l_end[i]]) .* (exp((-Sn/(1+PIC_q*av_prec[i])) *
        l_loss_str[l_start[i]:l_end[i]])) .* (exp((-Sn2/(1+PIC_q*av_prec[i])) ./l_loss_res[l_start[i]:l_end[i]])));
        W_lc[i]= sum((wild[l_start[i]:l_end[i]]) .* (exp((-Sn/(1+PIC_q*av_prec[i])) *
        l_loss_str[l_start[i]:l_end[i]])) .* (exp((-Sn2/(1+PIC_q*av_prec[i])) ./l_loss_res[l_start[i]:l_end[i]])));
```

```
        Disch[i]=sum((d_vals[d_start[i]:d_end[i]]) .* exp((-Sn/(1+PIC_q*av_prec[i])) *
        d_loss_str[d_start[i]:d_end[i]]) .* exp((-Sn2/(1+PIC_q*av_prec[i])) ./
        d_loss_res[d_start[i]:d_end[i]]));
        C_r[i]= sum((chick[l_start[i]:l_end[i]]) .* (exp((-Sn/(1+PIC_q*av_prec[i])) *
        l_loss_str[l_start[i]:l_end[i]])) .* (exp((-Sn2/(1+PIC_q*av_prec[i])) ./l_loss_res[l_start[i]:l_end[i]])));
        H_r[i]= sum((hog[l_start[i]:l_end[i]]) .* (exp((-Sn/(1+PIC_q*av_prec[i])) *
        l_loss_str[l_start[i]:l_end[i]])) .* (exp((-Sn2/(1+PIC_q*av_prec[i])) ./l_loss_res[l_start[i]:l_end[i]])));
        Cw_r[i]= sum((cow[l_start[i]:l_end[i]]) .* (exp((-Sn/(1+PIC_q*av_prec[i])) *
        l_loss_str[l_start[i]:l_end[i]])) .* (exp((-Sn2/(1+PIC_q*av_prec[i])) ./l_loss_res[l_start[i]:l_end[i]])));

        //Adding precipitation impact coefficient to land export
        A[i]=  Be_a * pow(av_prec2[i],pic_p[1]) .* A_lc[i];
        Dpr[i] = Be_d_pre * pow(av_prec2[i],pic_p[2]) .* D_lc_pre[i];
        Dpt[i] = Be_d_post * pow(av_prec2[i],pic_p[3]) .* D_lc_post[i];
        D[i]   = Dpr[i]+Dpt[i];
        W[i] = Be_w * pow(av_prec2[i],pic_p[4]) .* W_lc[i];
        C[i] = Be_ch * pow(av_prec2[i],pic_p[5]) .* C_r[i];
        H[i] = Be_h * pow(av_prec2[i],pic_p[6]) .* H_r[i];
        Cw[i] = Be_cw * pow(av_prec2[i],pic_p[7]) .* Cw_r[i];
        Dch[i] = Be_dch * Disch[i];
}

//Loop to determine random effect for each watershed
for (i in 1:nr){
        w= wsd[i];
        sigma[i] = sqrt(pow(SD[i],2)+pow(sigma_res,2));   //with regular sigma values
        alpha_vals[i] = alpha[w];
}

//Sum loadings from all sources
tot =  A + D + W + C + H + Cw + Dch;

//Add random effects to source loadings and subtract losses from upstream loads
y_hat = tot + 10000 * alpha_vals - up_t_load1 .* (1-(exp((-Sn ./ (1+PIC_q*av_prec)) .* str_loss_load1) .* exp((-Sn2
./ (1+PIC_q*av_prec)) ./ res_loss_load1))) - up_t_load2 .* (1-(exp((-Sn ./ (1+PIC_q*av_prec)) .* str_loss_load2) .*
exp((-Sn2 ./ (1+PIC_q*av_prec)) ./ res_loss_load2))) - up_t_load3 .* (1-(exp((-Sn ./(1+PIC_q*av_prec)) .*
str_loss_load3) .* exp((-Sn2 ./ (1+PIC_q*av_prec)) ./ res_loss_load3)));

//Priors for the model
Be_a ~ normal(900,700);                          //Prior for agriculture, TN/km2/yr
Be_d_pre ~ normal(800,300);                      //Prior for pre-1980 development, TN/km2/yr
Be_d_post ~ normal(800,300);                     //Prior for post-1980 development, TN/km2/yr
Be_w ~ normal(200,200);                          //Prior for undeveloped, TN/km2/yr
Be_ch ~ normal(0.001,0.0003);                    //Prior for chickens (TN/count)
Be_h ~ normal(0.04,0.02);                        //Prior for swine  (TN/count)
Be_cw ~ normal(0,5);                             //Uninformed Prior for cow
Be_dch ~ normal(1,.10);                          //Prior for point source delivery, unitless
sigma_res ~ normal(0,20);                        //st error of the model
sigma_w ~ normal(0,300);                         //st. deviation of random effect hyperdistribution
alpha ~ normal(0,sigma_w);                       //watershed random effects
sigma_B1 ~ normal(0,1);                          //st. deviation for PIC hyperdistribution
Bp_mean ~ normal(1,1);                           //mean for PIC hyperdistribution
pic_p[1] ~ normal(Bp_mean,sigma_B1);             //PIC ag distribution
pic_p[2] ~ normal(Bp_mean,sigma_B1);             //PIC pre distribution
pic_p[3] ~ normal(Bp_mean,sigma_B1);             //PIC post distribution
pic_p[4] ~ normal(Bp_mean,sigma_B1);             //PIC undeveloped distribution
```

```
pic_p[5] ~ normal(Bp_mean,sigma_B1);          //PIC chicken distribution
pic_p[6] ~ normal(Bp_mean,sigma_B1);          //PIC hog distribution
pic_p[7] ~ normal(Bp_mean,sigma_B1);          //PIC cow distribution
Sn ~ normal(.14,.05);                         //Prior for stream retention rate, 1/d
Sn2 ~ normal(11,2);                           //Prior for waterbody retention rate, m/y
PIC_q ~ normal(0,1);                          //PIC for retention

// with log transformation
ly_hat=log((y_hat ./10000) + 10);     // vector to get log of TN load (y_hat)
ly ~ normal(ly_hat,sigma_res);        //parameter that calibrates ly_hat (log of TN load) with ly ()
y=exp(ly)-10;                         // parameter that calibrates to load
load ~ normal(y,SD);                  // load = WRTDS estimate, SD = WRTDS sd
}
```