

SUPPLEMENTARY MATERIAL

SM 1

#####Total phosphorus (TP) in summer-autumn (SA)- Case1 #####

#Categorical variables are converted to factors
V1 = Variable 1

```
dataSA$V1=as.factor(dataSA$V1)
dataSA$V2=as.factor(dataSA$V2)
```

###GLM ###

```
attach(dataSA)
glm.dataSA=glm(TP~(V1)+(V2)+(V3)+(V4)+(V5)+(V6)+(V7),dataSA,family="gaussian")
summary(glm.dataSA)
AIC(glm.dataSA)
```

#Estimate pseudo R²
require(rcompanion)
nag=nagelkerke(glm.dataSA)
nag\$Pseudo.R.squared.for.model.vs.null
require(DescTools)

#check
plot(glm.dataSA)
VIF(glm.dataSA)

#The best submodel is selected by stepAIC
require(MASS)
stepAIC(glm.dataSA,direction="backward")
glm.optI=glm(TP ~UI+OD+DP,family="gaussian",dataSA)
summary(glm.optI)
nag.optI=nagelkerke(glm.optI)
nag.optI\$Pseudo.R.squared.for.model.vs.null

##ERROR - estimated per test sample - GLM model -
set.seed(101)
K=25
error=list(matrix(NA, K))
varianza=list(matrix(NA, K))
n = nrow(dataSA)
for(k in 1:K) {
 smp=sample(n,round(n/10))
 learn=dataSA[-smp,]
 learn=learn[,1:11]
 test=dataSA[smp,]
 test=test[,1:11]
 glm.optI=glm(TP~(V1)+(V2)+(V3)+(V4)+(V5)+(V6)+(V7),family="gaussian",data=learn)
 pred= predict(glm.optI,test)
 error[k] = ((sqrt(mean((test[,1]-pred)^2)))/mean(test[,1]))*100
 varianza[k]=1-((mean((test[,1]-pred)^2))/var(test[,1]))
}

#Error and Var
error.glm=mean(unlist(error))
error.sd.glm=sd(unlist(error))
var.glm=mean(unlist(varianza))
var.sd.glm=sd(unlist(varianza))
error.glm
error.sd.glm

```
var.glm
var.sd.glm
```

```
#####
```

```
###GAM - example to TP ###
```

```
require(mgcv)
gam.dataSA=gam(TP~(V1)+(V2)+s(V3)+s(V4)+s(V5)+s(V6)+s(V7)+s(V8)+s(V9),dataSA,family="gaussian")
summary(gam.dataSA)
AIC(gam.dataSA)
plot(gam.dataSA)
```

```
#check
```

```
gam.check(gam.optI)
```

```
##ERROR - estimated per test sample - GAM model -
```

```
gam.optI<-gam.dataSA
set.seed(55)
K=200
error=list(matrix(NA, K))
varianza=list(matrix(NA, K))
n = nrow(dataSA)
for(k in 1:K) {
  smp=sample(n,round(n/10))
  learn=dataSA[-smp,]
  learn=learn[,1:11]
  test=dataSA[smp,]
  test=test[,1:11]
  gam.optI=gam(TP~(V1)+(V2)+s(V3)+s(V4)+s(V5)+s(V6)+s(V7)+s(V8)+s(V9),family="gaussian",data=learn)
  pred= predict(gam.optI,test)
  error[k] = sqrt(mean((test[,1])-pred)^2)
  varianza[k]=1-((mean((test[,1])-pred)^2)/var(test[,1]))
}
```

```
#Error and Var
```

```
error.gam=mean(unlist(error))
error.sd.gam=sd(unlist(error))
var.gam=mean(unlist(varianza))
var.sd.gam=sd(unlist(varianza))
error.gam
error.sd.gam
var.gam
var.sd.gam
```

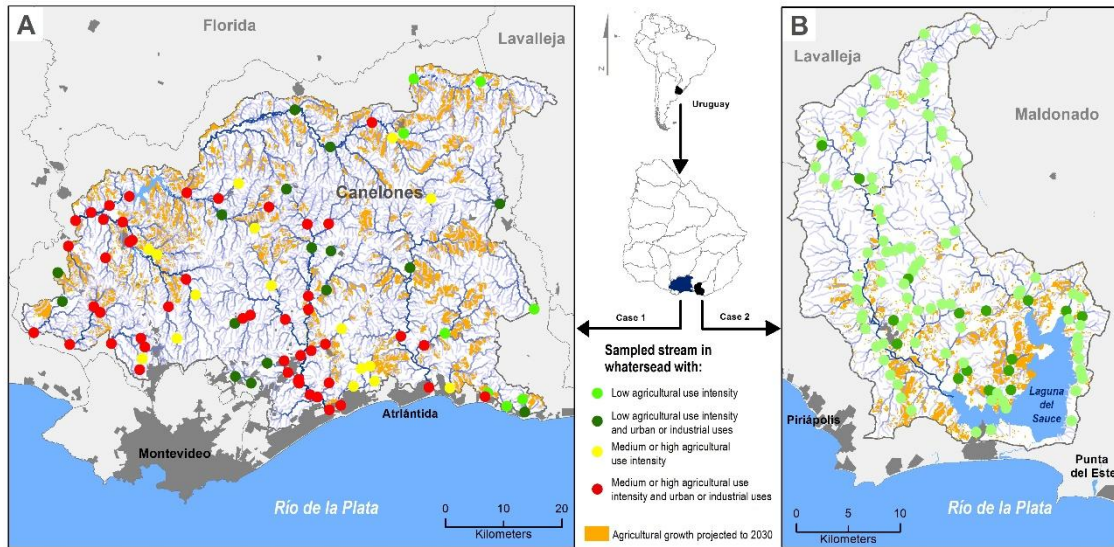


Figure S1. Projected agricultural growth to 2030 and water quality sampling points in lotic systems in Case 1 (A) and Case 2 (B). In both maps, urban areas are shown in dark gray and course order according to Strahler's method is shown in colours from light blue (1) to dark blue (8). The projection for 2030 scenario was generated with information from Achkar et al. (2012) INE (INE, 2014) and OPP (OPP, 2009), and agricultural intensity of the soil was taken from Díaz et al. (2018).

Achkar, M., Blum, A., Bartesaghi, L., Ceroni, M., 2012. Escenarios de cambio de uso del suelo en Uruguay.

Díaz, I., Ceroni, A., López, G., Achkar, M., 2018. Análisis espacio-temporal de la intensificación agraria y su incidencia en la productividad primaria neta. *Rev. Electrónica Medioambiente*. UCM 19, 24–40.

INE, 2014. Estimaciones y proyecciones de la población de Uruguay: metodología y resultados Revisión 2013.

OPP, 2009. Estrategia Uruguay III SIGLO. Aspectos Productivos.

SM 3

A Generalized Linear Model (GLM) was developed in order to identify and then select the best performance model, prior to the development of the Generalized Additive Model (GAM) (McCullagh and Nelder, 1989). To elaborate the GLM and to comply with the assumptions of normality and homoscedasticity of residues, log (X+1) transformation was used. It was found that, for both areas, for both samples (in two stations) and for both nutrients analyzed (TP and TN), GAM presented a higher percentage of the explained variance, better adjustment and less error than GLM.

Sample	WS			SA				WS			SA		
Parameter	TP			TP				TN			TN		
Statistical	psR ²	ΔAIC	NRMSE	psR ²	ΔAIC	ECM	VE	psR ²	ΔAIC	NRMSE	psR ²	ΔAIC	NRMSE
Caso 1													
Section A	0.32	-	20.8	0.37	-	15.7		0.30	-	12.2	0.22	-	24.1
Section B	0.42	-13	20.2	0.53	-24	13.3		0.48	-26	10.0	0.33	-13	20.8
Caso 2													
Section A	0.47	-1.5	13.3	0.30	-	38.7		0.21	-2	17.1	0.32	-3	17.8
Section B	0.48	-	13.6	0.33	3	37.6		0.21	-	17.2	0.26	-	18.4

Table S1. GLM models for total phosphorus (TP) and total nitrogen (NT) in Case 1 (Canelones) and Case 2 (Laguna del Sauce), for Section A (basin) and B (basin + course), in winter-spring (WS) and summer-autumn (SA). The correlation between predicted and response values (R²), the difference according to the information criterion and Akaike (ΔAIC) between the models that include the A and B Section, and the normalized root mean square error (NRMSE) as a % are presented.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models, The R Book. Chapman and Hall.