



## Setting up SWAT for the Upper Amazon

Michael Strauch, Martin Volk

2014 International SWAT Conference  
Porto de Galinhas, July 30 – August 1, 2014

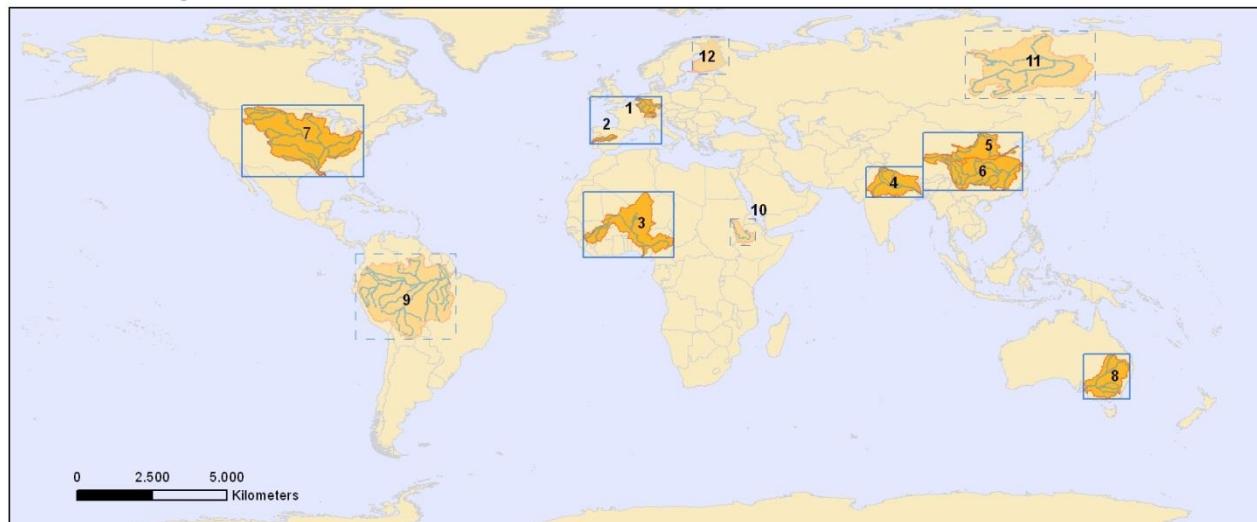


- community-driven modelling effort
- bringing together impact models across sectors and scales
- to create consistent and comprehensive projections of the impacts of different levels of global warming

coordinated by



### Focus regions



**Europe**  
1. Rhine  
2. Tagus  
12. Finland

**Africa**  
3. Niger  
10. Blue Nile

**Asia**  
4. Ganges  
5. Yellow  
6. Yangtze  
11. Lena

**N.America**  
7. Mississippi

**Australia**  
8. Murray Darling

**S.America**  
9. Amazon

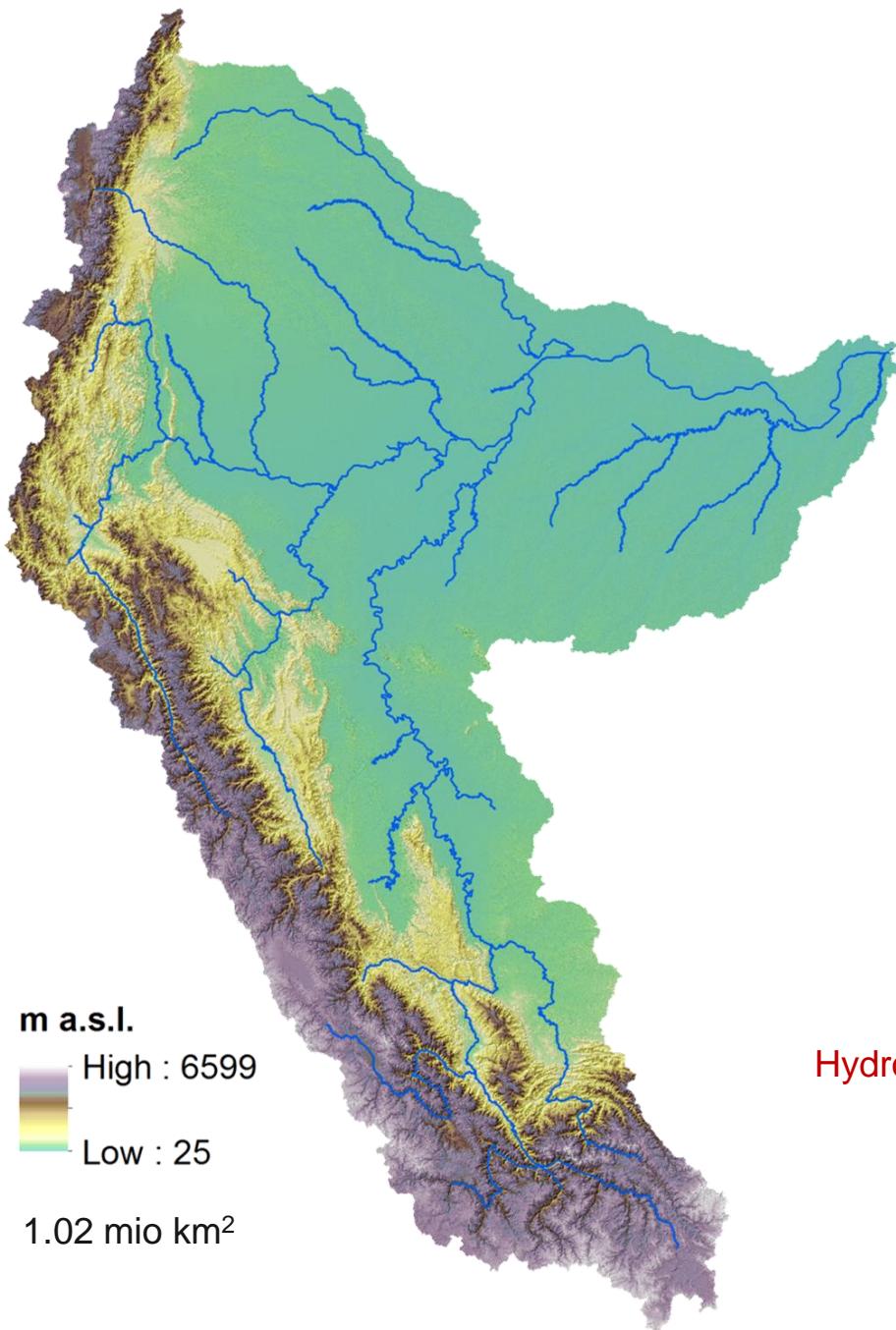
**ISI-MIP Framework:**  
Common set of input data  
Common modelling protocol  
Central output data archive



# SWAT

 Soil & Water Assessment Tool

Input data	Data source
DEM	HydroSHEDS (SRTM, 90 m, hydrol. conditioned, Lehner et al. 2008)
Soil	Harmonized World Soil Database, Version 1.2, SOL_K calc. using ROSETTA (Schaap et al., 2001)
Land use	GLC2000 product South America (Eva et al., 2002)
Climate	WFD ERA 40 (Weedon et al., 2010)
Calibration/validation	Global Runoff Data Centre (GRDC), Hidroweb ( <a href="http://hidroweb.ana.gov.br/">http://hidroweb.ana.gov.br/</a> )



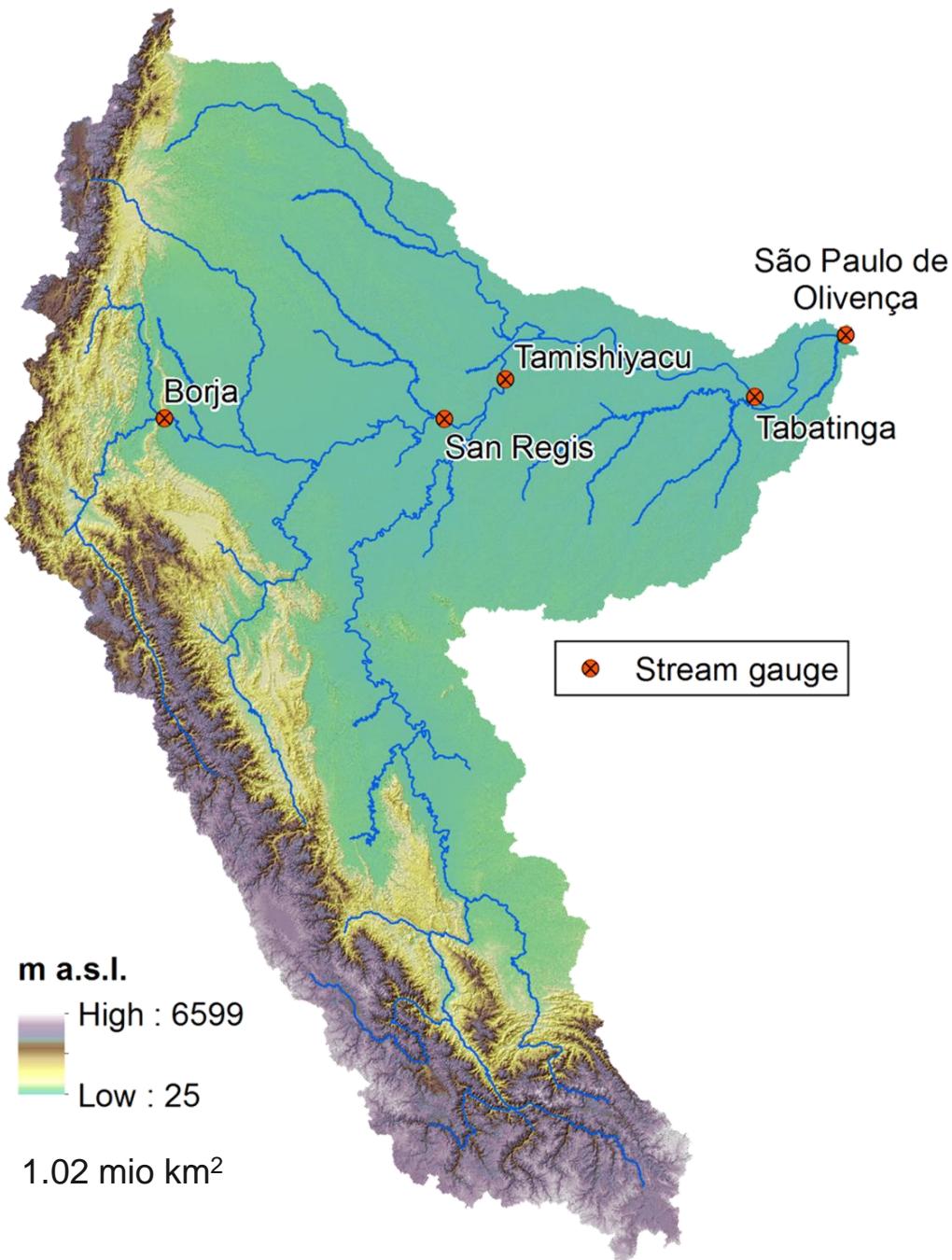
## Digital Elevation Model

SRTM 90 m,  
hydrologically conditioned

Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales (Lehner et al., 2008)



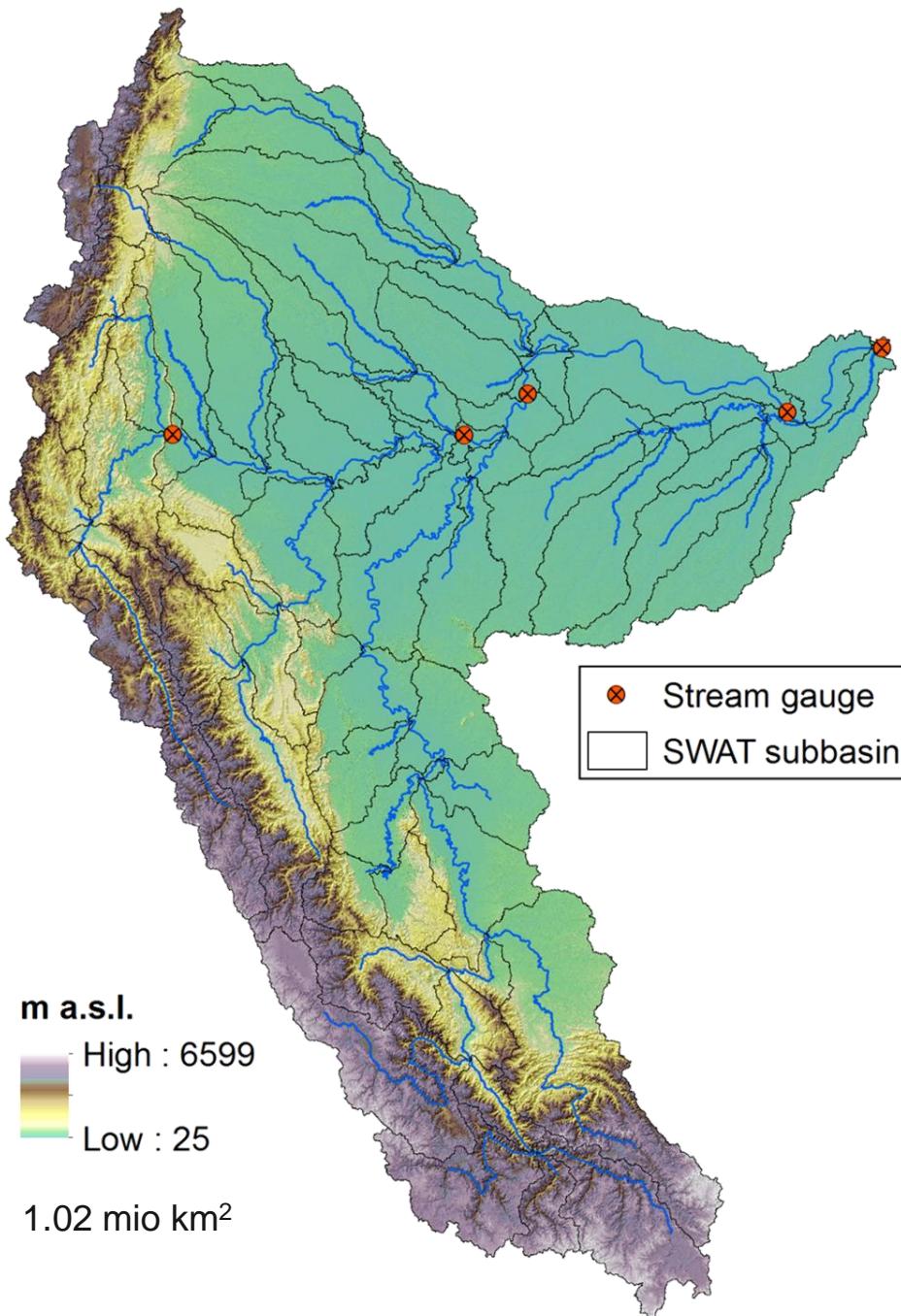
HydroSHEDS



## Streamflow data

Global Runoff Data Centre



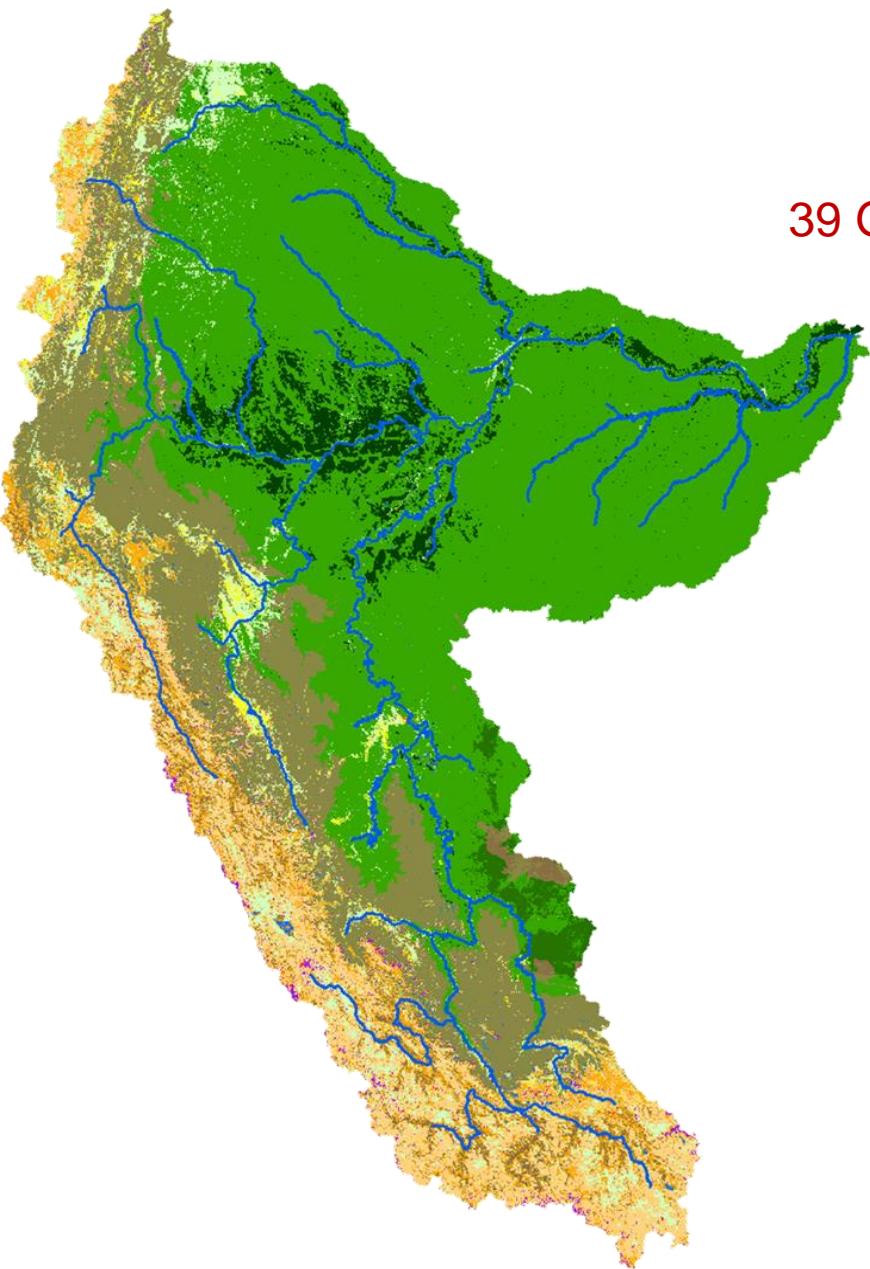


## Subbasin delineation

Based on suggested minimum size: 719,600 ha

⇒ 82 subbasins ( $\varnothing 12,500 \text{ km}^2$ )

⇒ too large for physically meaningful Muskingum routing???



## Land use

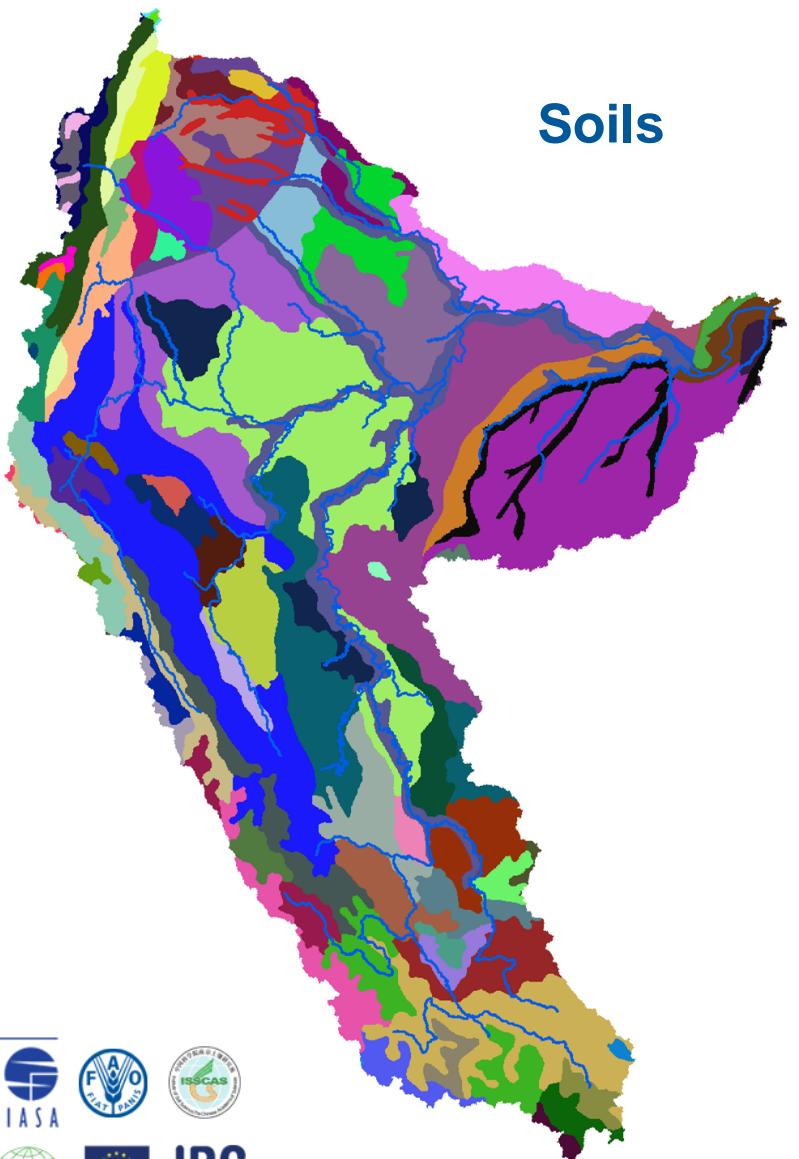
39 GLC classes aggregated to 16 SWAT classes

- AGRL - Agriculture - intensive
- PAST - Mosaic agriculture / degraded vegetation
- BARR - Barren / bare soil; Desert; Permanent snow/ice
- FRSD - Deciduous forest
- FRSE - Evergreen tropical forest
- FRST - Bamboo dominated and semi-deciduous forest
- FSDM - Montane deciduous forests (>500m)
- FSEM - Montane evergreen forests (>500m)
- FSTM - Montane bamboo / semi-deciduous (>500m)
- RNGB - Shrub savannah and shrublands
- RNGE - Steppe and montane grasslands
- URLD - Urban
- WATR - Water bodies
- WETF - Flooded / permanent swamp forest
- WETL - Periodically flooded savannah
- WTFM - Montane flooded forests (>500m)

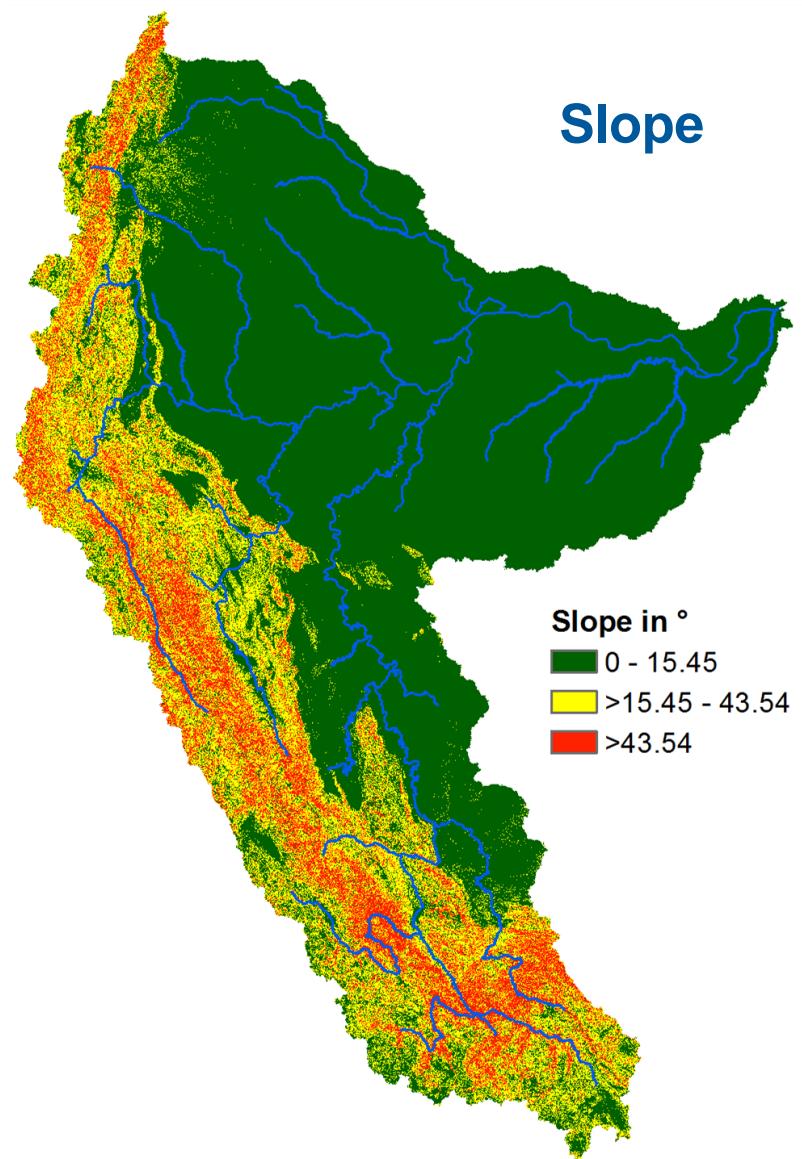
Global Land Cover 2000 product South America (Eva et al., 2002)



JOINT RESEARCH CENTRE  
Land Resource Management Unit



Soil class	PE25
	BR214
	PE28
	BR215
	PE3
	BR217
	PE30
	BR221
	PE31
	BR223
	PE34
	BR228
	PE35
	BR233
	PE36
	BR234
	PE37
	BR280
	PE38
	CO27
	PE39
	CO34
	PE4
	CO45
	PE41
	CO56
	PE43
	CO8
	PE44
	EC1
	PE45
	EC10
	PE46
	EC11
	PE47
	EC12
	PE48
	EC13
	PE53
	EC14
	PE54
	EC15
	PE55
	EC17
	PE56
	EC18
	PE57
	EC2
	PE58
	EC29
	PE59
	EC3
	PE61
	EC38
	PE62
	EC39
	PE64
	EC4
	PE65
	EC40
	PE66
	EC41
	PE69
	EC43
	PE7
	EC5
	PE70
	EC6
	PE71
	EC7
	PE72
	EC8
	PE73
	EC9
	PE74
	PE1
	PE75
	PE10
	PE76
	PE11
	PE77
	PE12
	PE78
	PE13
	PE79
	PE14
	PE8
	PE15
	PE80
	PE17
	PE9
	PE2



**Slope in °**

0 - 15.45
>15.45 - 43.54
>43.54



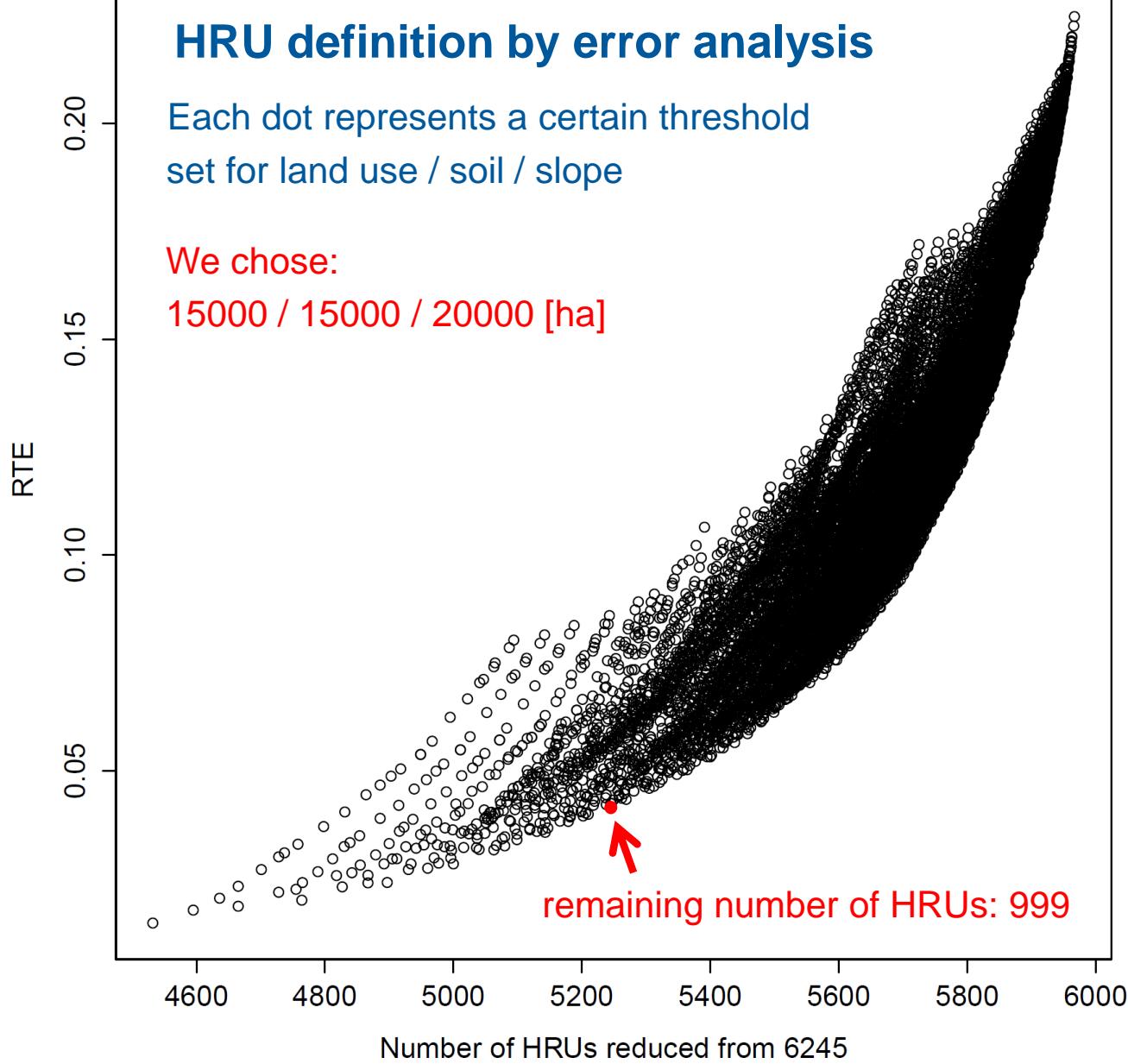
Harmonized World Soil Database, Version 1.2

## HRU definition by error analysis

Each dot represents a certain threshold  
set for land use / soil / slope

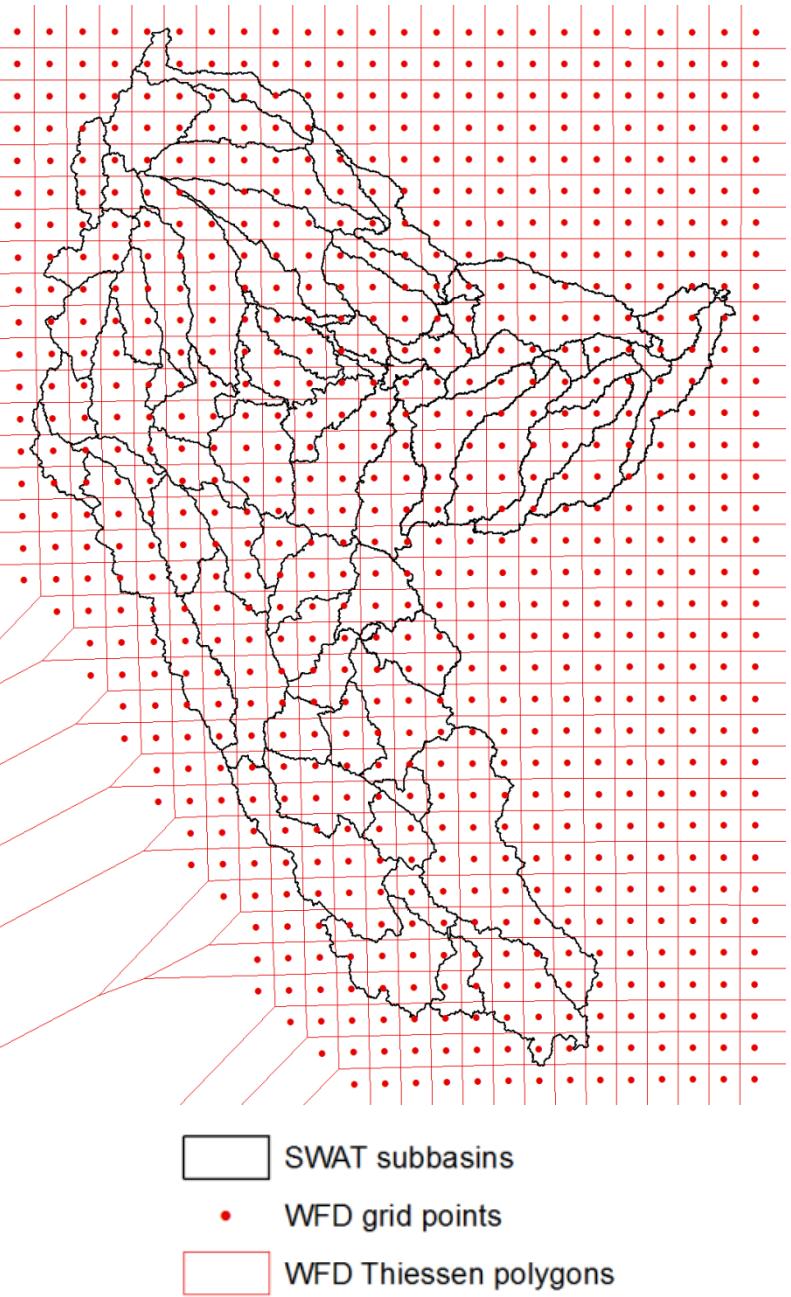
We chose:

15000 / 15000 / 20000 [ha]



RTE:

Relative Total Error  
(of land use, soil, and  
slope distribution com-  
pared to full HRU map)



## Weather data

WFD – WATCH Forcing Data (Basis ERA 40),  
daily for time period 1/1/1958 – 31/12/2001

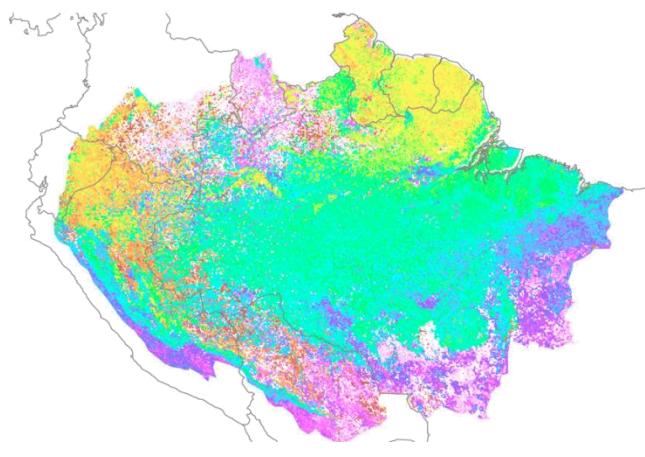
- Temperature (max, min in K)
- Precipitation (in kg/m<sup>2</sup>/s)
- Wind speed (in m/s)
- Solar radiation (in W/m<sup>2</sup>)
- Specific humidity (in kg/kg)
- Surface pressure (in Pa)

Weedon et al. (2010)



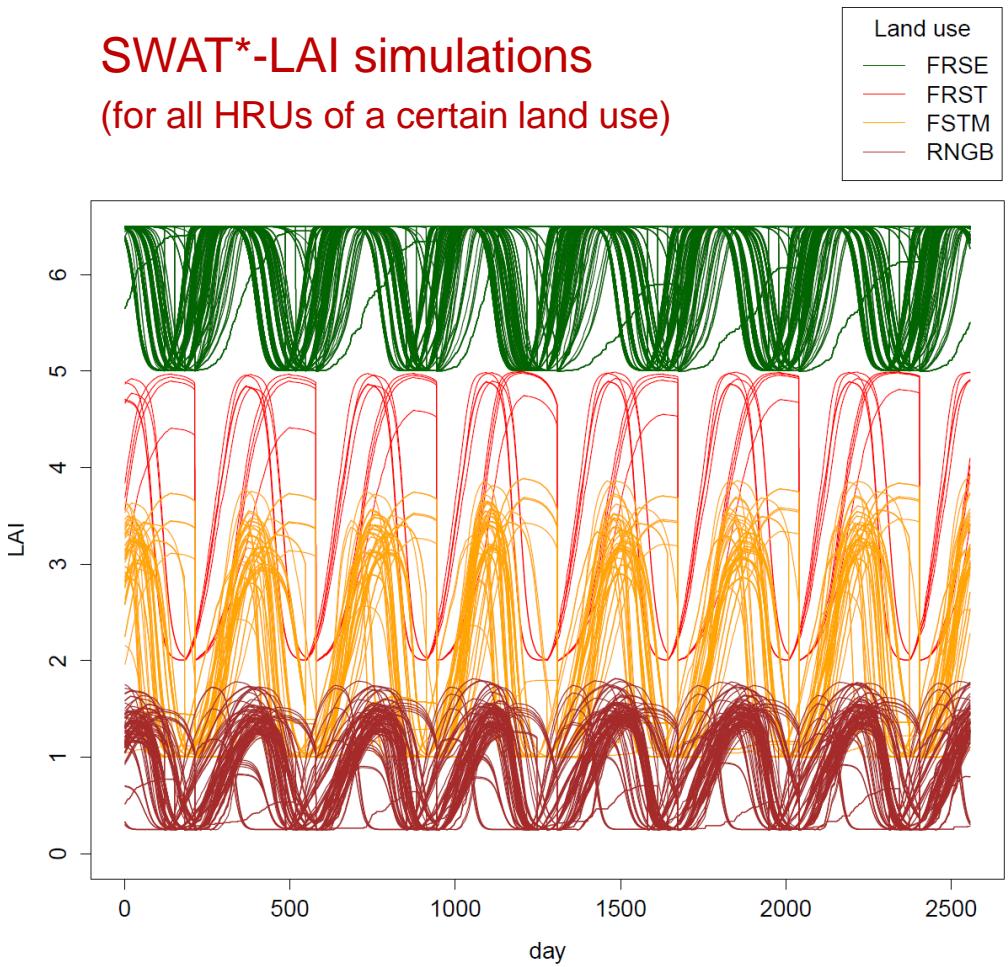
# Model calibration: Plant growth

Time of NDVI peaks  
(Brando et al., 2010)



early Jan	early Jul
late Jan	late Jul
early Feb	early Aug
late Feb	late Aug
early Mar	early Sep
late Mar	late Sep
early Apr	early Oct
late Apr	late Oct
early May	early Nov
late May	late Nov
early Jun	early Dec
late Jun	late Dec

SWAT\*-LAI simulations  
(for all HRUs of a certain land use)

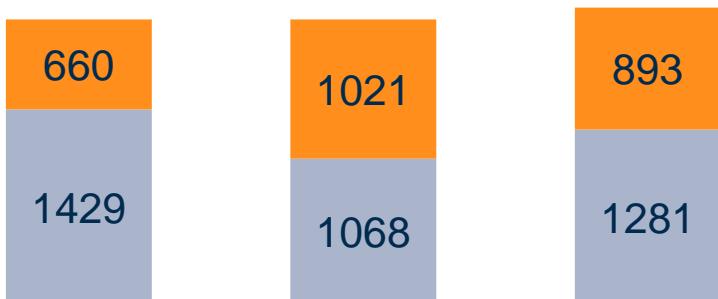


\*Seasonal pattern depending on latitude & soil moisture using  
the modified SWAT version for tropics! (Strauch & Volk, 2013)

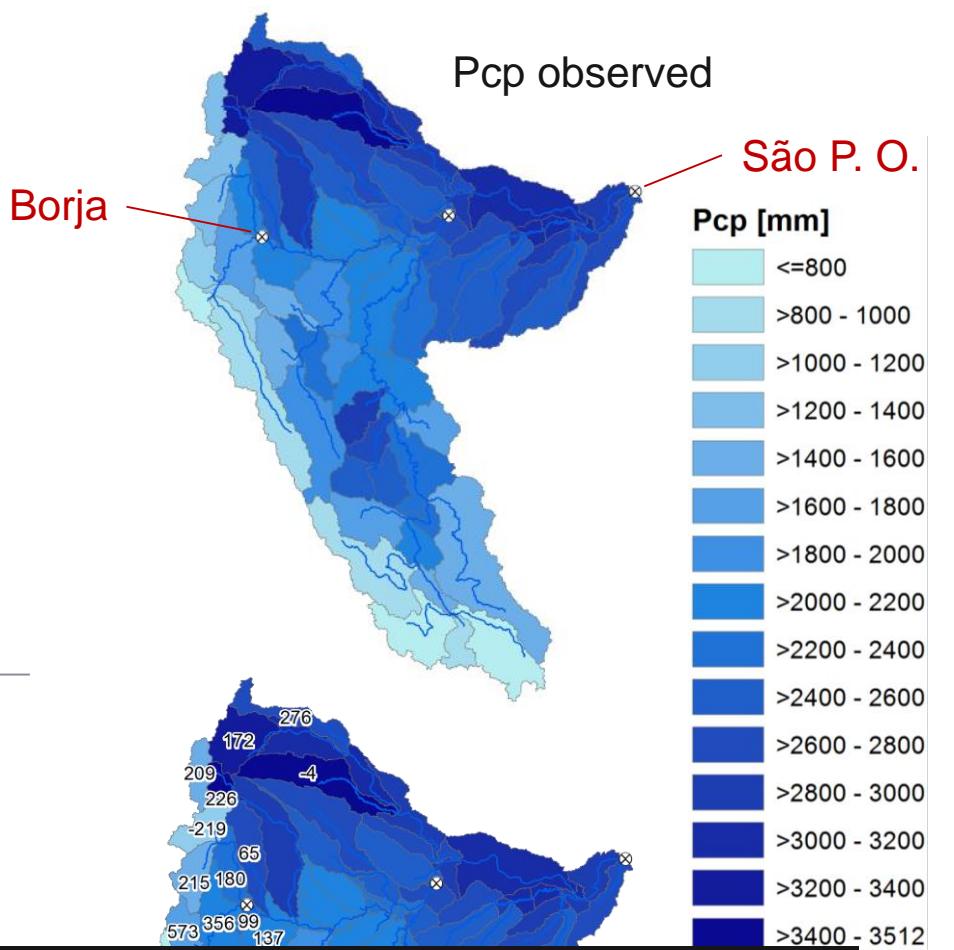
# Model calibration: Water balance (1986-1994)

## São Paulo de Olivença

■ R (mm)   ■ ET (P-R, mm)

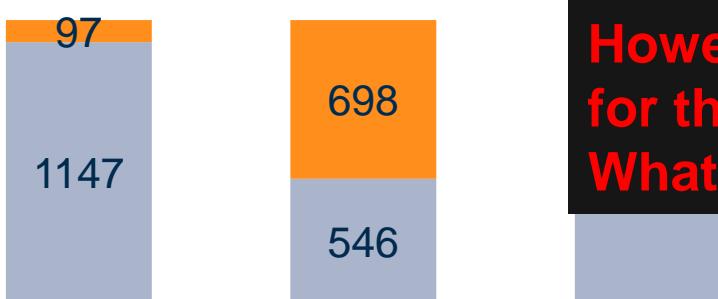


observed      sim initial      sim elevb



## Borja

■ R (mm)   ■ ET (P-R, mm)

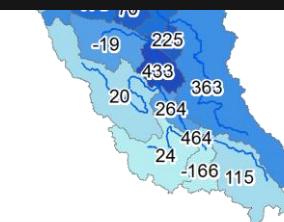


observed      sim initial      sim elevb

**However, elevation bands not justified  
for the type of data used in this study.  
What to do?**

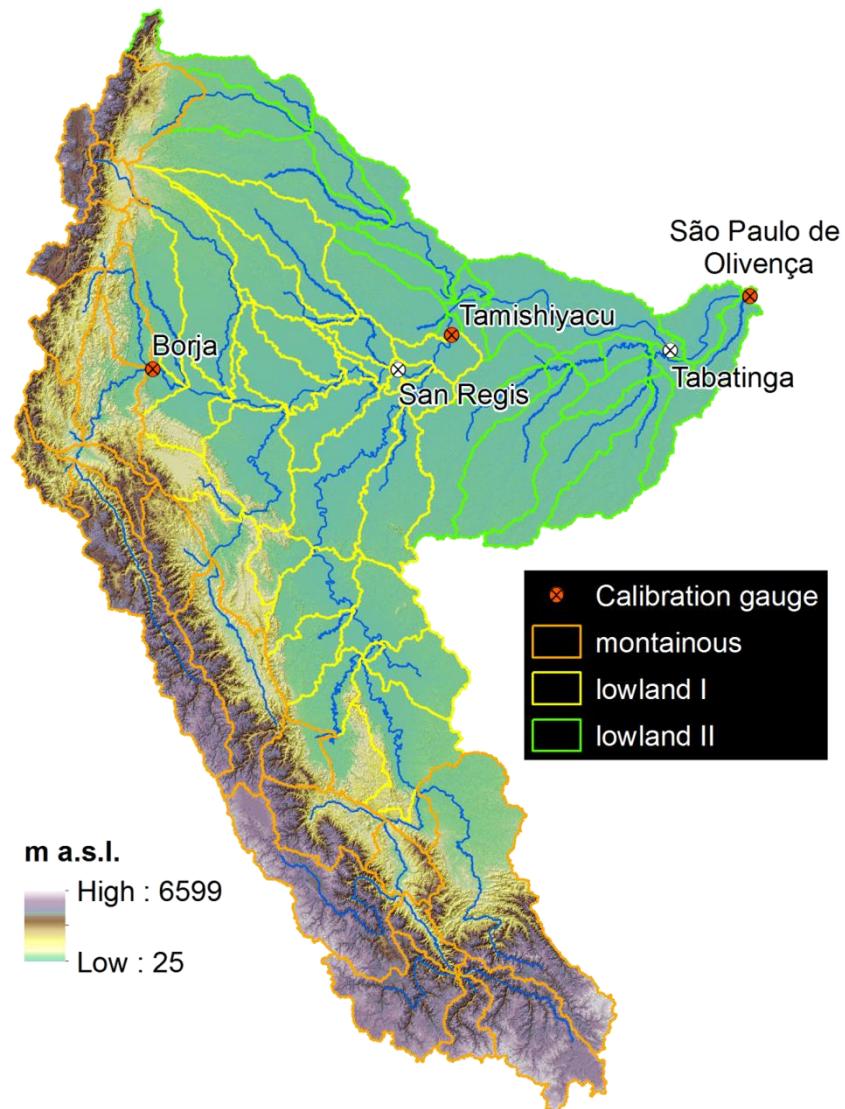
ation  
s in

e. b. from DEM  
PLAPS: 500 mm/km  
TLAPS: -8 °C/km



# Model calibration

## Multi-site autocalibration



- Latin-Hypercube sampling (n=3000)
- 32 parameters (5 basin wide + 8 + 11 + 7 regional parameters)
- Best solution defined by **minimum rank sum of 11 performance metrics** over all three calibration gauges
- Uncertainty interval (UI) by considering as many (sorted) solutions as necessary that:  
 $\text{width}_{\text{UI}} = \sigma_{\text{obs}}$

## Model evaluation

Rank sum of 11 performance metrics over three gauges

PBIAS Percent Bias

NSE Nash-Sutcliff Efficiency

mNSE Modified Nash-Sutcliff Efficiency

rNSE Relative Nash-Sutcliff Efficiency

d Index of Agreement

md Modified Index of Agreement

rd Relative Index of Agreement

cp Persistence Index

bR2  $R^2$  multiplied by coefficient of regression line between sim and obs

KGE Kling-Gupta Efficiency

VE Volumetric Efficiency

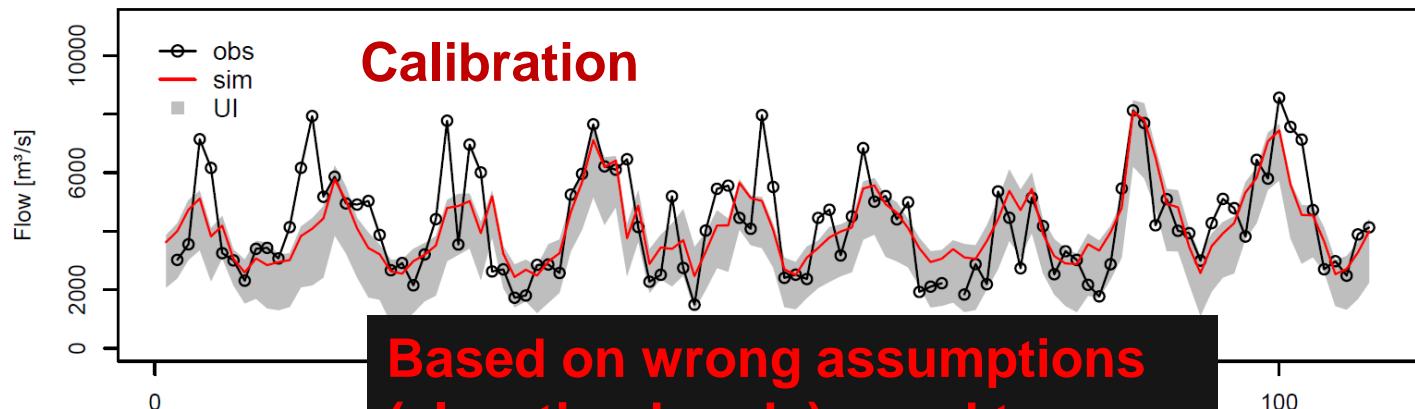
Borja

NSE = 0.52

PBIAS = 3.5%

d-factor = 1.11

POC = 61%



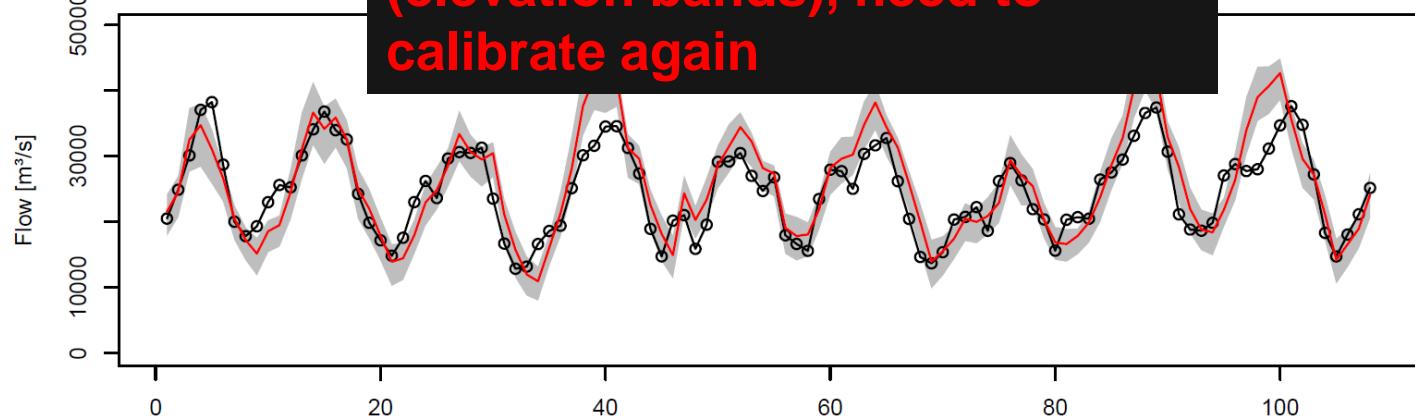
Tamishiyacu

NSE = 0.63

PBIAS = -4.8%

d-factor = 1.01

POC = 59%



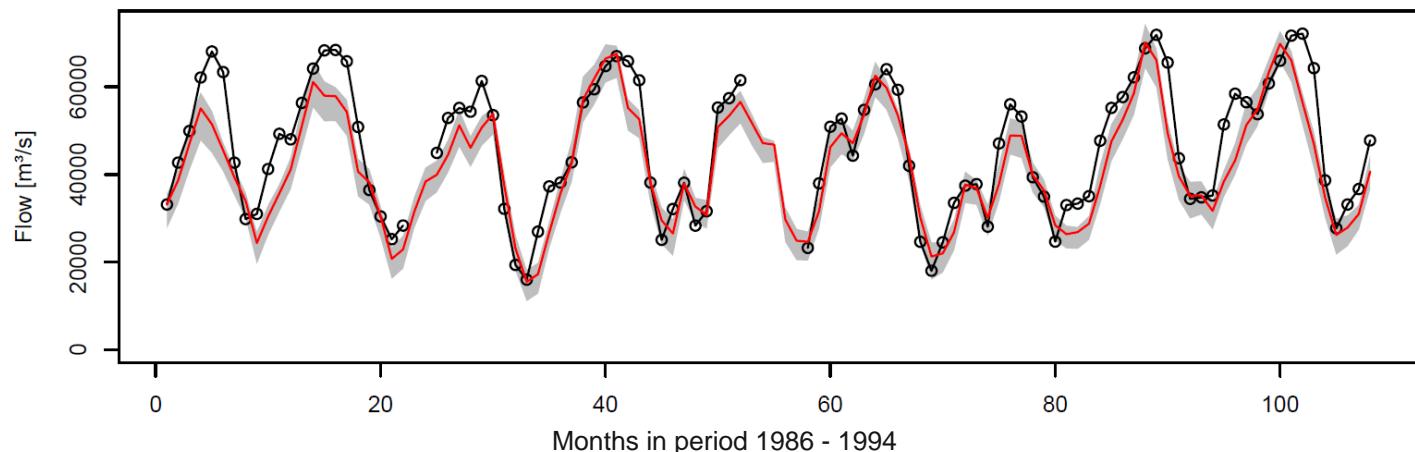
São Paulo de O.

NSE = 0.78

PBIAS = 8.8%

d-factor = 0.51

POC = 43%



Borja

NSE = 0.65

PBIAS = -3.5%



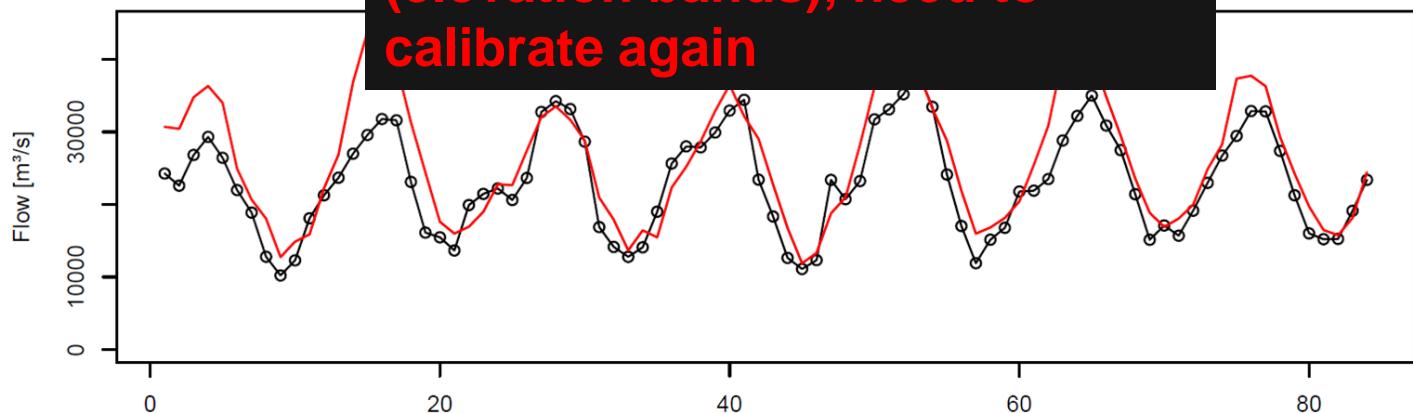
## Validation

Based on wrong assumptions  
(elevation bands), need to  
calibrate again

Tamishiyacu

NSE = 0.51

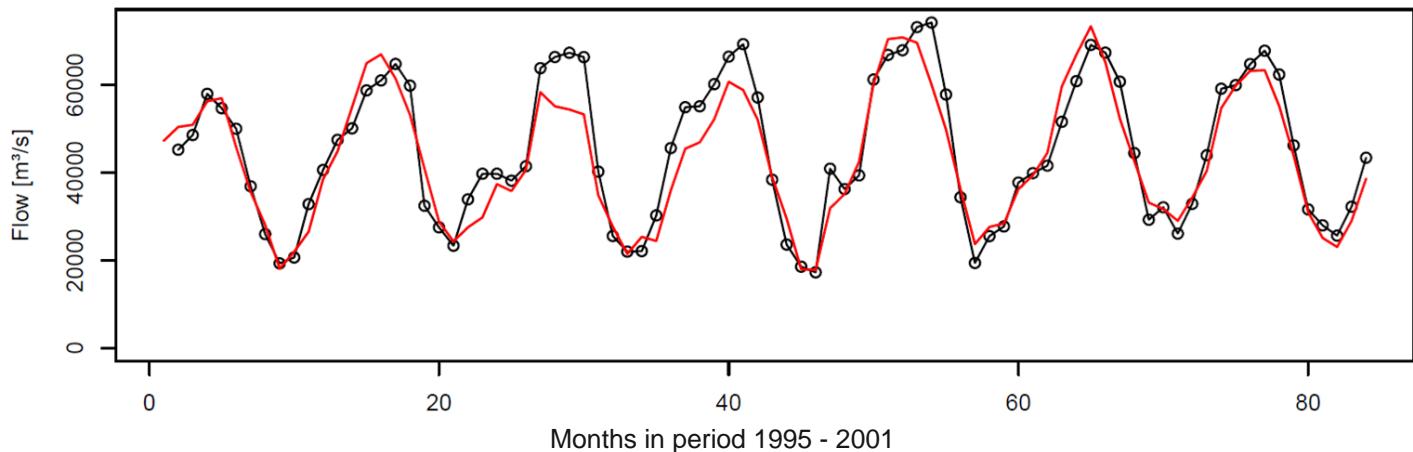
PBIAS = -13.5%

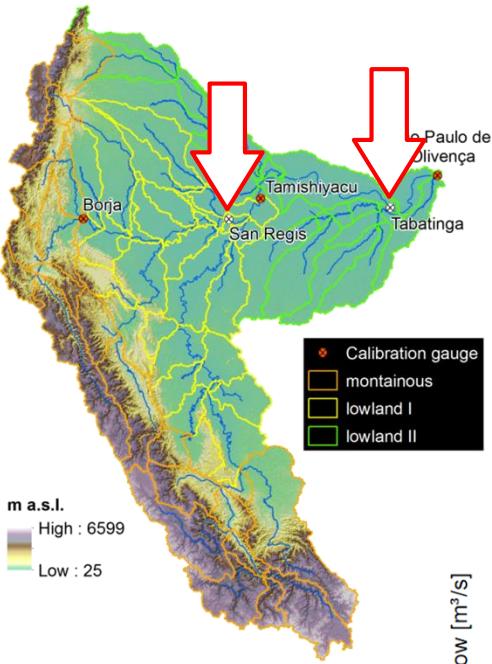


São Paulo de O.

NSE = 0.89

PBIAS = 4%

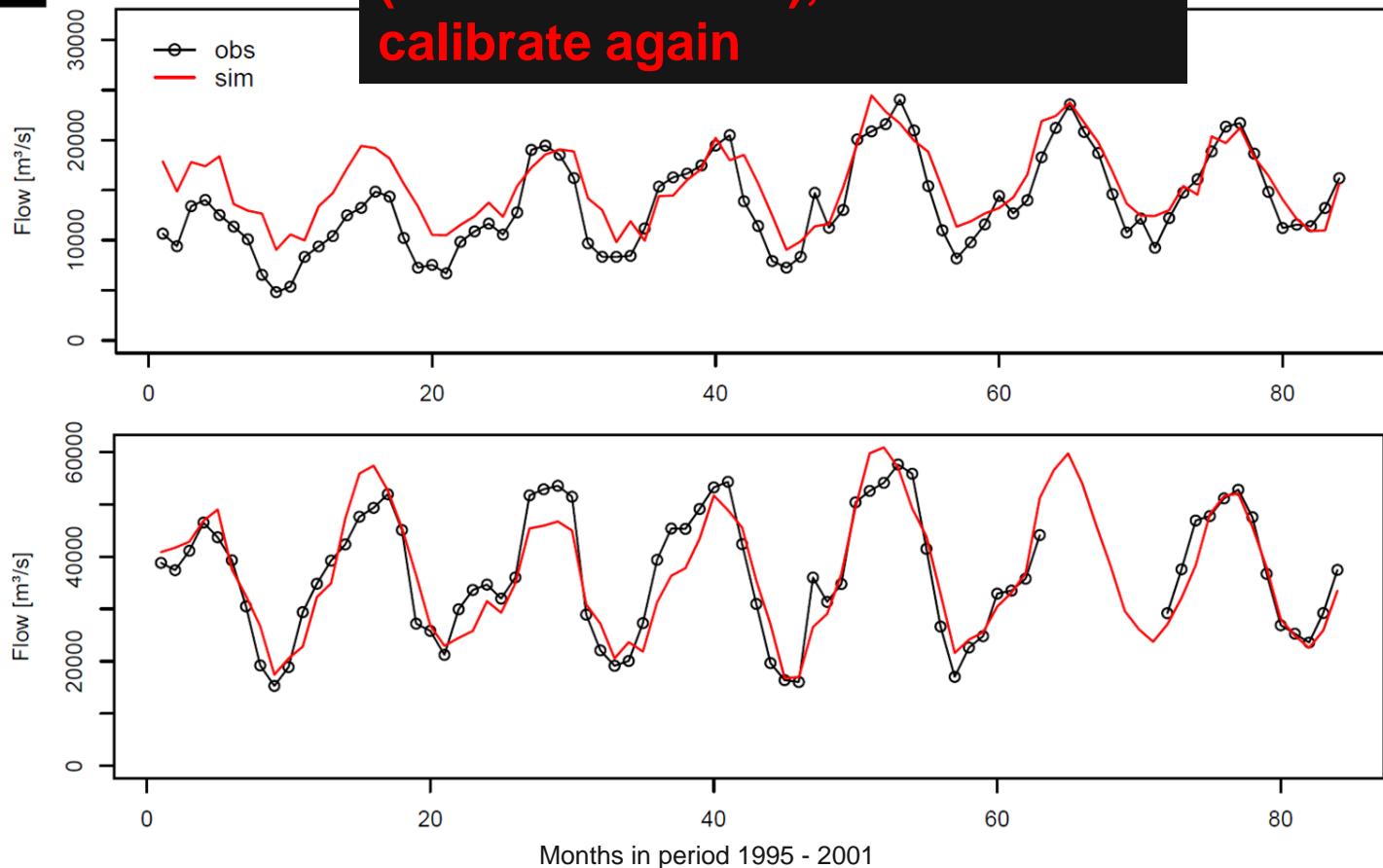




## Cross Validation (at gauges not used for calibration)

San Regis

**Based on wrong assumptions  
(elevation bands), need to  
calibrate again**



Tabatinga  
NSE = 0.84  
PBIAS = 0.8%

## Summary / conclusion

- **Ready-to-use global datasets** for fast model setup **available**
- **However, many tricky issues:**
  - Maximum area of subbasins in large-scale applications?
  - Assigning tropical vegetation types to SWAT land use classes
  - Plant growth seasonal pattern (modified version useful!)
  - Abnormal water balance according observations
  - Elevation bands significantly improved model performance but are not justified when using grid data averaged over subbasins
- **Multi-site autocalibration** and **multi-metrics model evaluation** to provide robust calibration results and uncertainty intervals

Muito obrigado!



# Additional slides...

Three RStudio windows are shown side-by-side, each displaying R code for a script named '02\_rewrite.R'. The code is identical across all three windows.

```

419 #read in list of directories (MGT)
420 filenames<-dir(paste(getwd(),"\\backup",sep=""),pattern=".mgt")
421 sapply(1:length(filenames),adj.mgt)
422
423 #read in list of directories (BSN)
424 filenames<-dir(paste(getwd(),"\\backup",sep=""),pattern=".bsn")
425 sapply(1:length(filenames),adj.bsn)
426
427 #read in list of directories (.sub)
428 filenames<-dir(paste(getwd(),"\\backup",sep=""),pattern=".sub")
429 sapply(1:length(filenames),adj.sub)
430
431 #read in list of directories (.rte)
432 filenames<-dir(paste(getwd(),"\\backup",sep=""),pattern=".rte")
433 sapply(1:length(filenames),adj.rte)
434
435 #execute SWAT
436 setwd(paste(basepath,"/txtinout_1",sep ""))
437 system("swat2012.exe")
438
439 #store output.rch files
440 filenamesfirst<-dir(pattern="output.std")
441 newnumber<-sprintf("%05d", t)
442 filenamenew<-paste("output",newnumber,".std",sep="")
443 file.rename(filenamesfirst, filenamenew)
444
445 #store reach outputs
446 dat <- read.table(paste(basepath,"/txtinout_1/output.rch",sep ""),
447 dat_1 <- dat[,6][dat[,2]==REACH_1]
448 write(file = paste("out_",REACH_1,".txt",sep=""), c(paste("#",t,sep="",
449 ,dat_1, append = TRUE)
450 write(file = paste("out_",REACH_2,".txt",sep=""), c(paste("#",t,sep="",
451 ,dat_2 <- dat[,6][dat[,2]==REACH_2]
452 write(file = paste("out_",REACH_3,".txt",sep=""), c(paste("#",t,sep="",
453 ,dat_3 <- dat[,6][dat[,2]==REACH_3]
454 write(file = paste("out_",REACH_3,".txt",sep=""), c(paste("#",t,sep="",
455 ,dat_3, append = TRUE)
456
457 SWAT2012
458 Rev. 588
459 Modified for tropics by M. Strauch
460 Soil & Water Assessment Tool
461 PC Version
462 Program reading from file.cio . . . executing
463
464 Executing year 1
465 Executing year 2
466 Executing year 3
467 Executing year 4
468 Executing year 5
469 Executing year 6
470 Executing year 7
471 Executing year 8
472 Executing year 9

```

The consoles below each window show the execution of the R code. The first two windows show the execution of the R code, while the third window shows the execution of the SWAT2012 executable.

```

Console ~ / ~
+ dat_1 <- dat[,6][dat[,2]==REACH_1]
+ write(file = paste("out_",REACH_1,".txt",sep=""), c(paste("#",t,sep="",
+ ,dat_1, append = TRUE)
+ dat_2 <- dat[,6][dat[,2]==REACH_2]
+ write(file = paste("out_",REACH_2,".txt",sep=""), c(paste("#",t,sep="",
+ ,dat_2, append = TRUE)
+ dat_3 <- dat[,6][dat[,2]==REACH_3]
+ write(file = paste("out_",REACH_3,".txt",sep=""), c(paste("#",t,sep="",
+ ,dat_3, append = TRUE)
+
+ )
SWAT2012
Rev. 588
Modified for tropics by M. Strauch
Soil & Water Assessment Tool
PC Version
Program reading from file.cio . . . executing
Executing year 1
Executing year 2
Executing year 3
Executing year 4
Executing year 5
Executing year 6
Executing year 7
Executing year 8
Executing year 9

```

```

Console ~ / ~
+ dat <- read.table(paste(basepath,"/txtinout_3/output.rch",sep ""),
+ skip=9, as.is=TRUE)
+ dat_1 <- dat[,6][dat[,2]==REACH_1]
+ write(file = paste("out_",REACH_1,".txt",sep=""), c(paste("#",t,sep="",
+ ,dat_1, append = TRUE)
+ dat_2 <- dat[,6][dat[,2]==REACH_2]
+ write(file = paste("out_",REACH_2,".txt",sep=""), c(paste("#",t,sep="",
+ ,dat_2, append = TRUE)
+ dat_3 <- dat[,6][dat[,2]==REACH_3]
+ write(file = paste("out_",REACH_3,".txt",sep=""), c(paste("#",t,sep="",
+ ,dat_3, append = TRUE)
+
+ )
SWAT2012
Rev. 588
Modified for tropics by M. Strauch
Soil & Water Assessment Tool
PC Version
Program reading from file.cio . . . executing
Executing year 1
Executing year 2
Executing year 3
Executing year 4
Executing year 5
Executing year 6
Executing year 7
Executing year 8

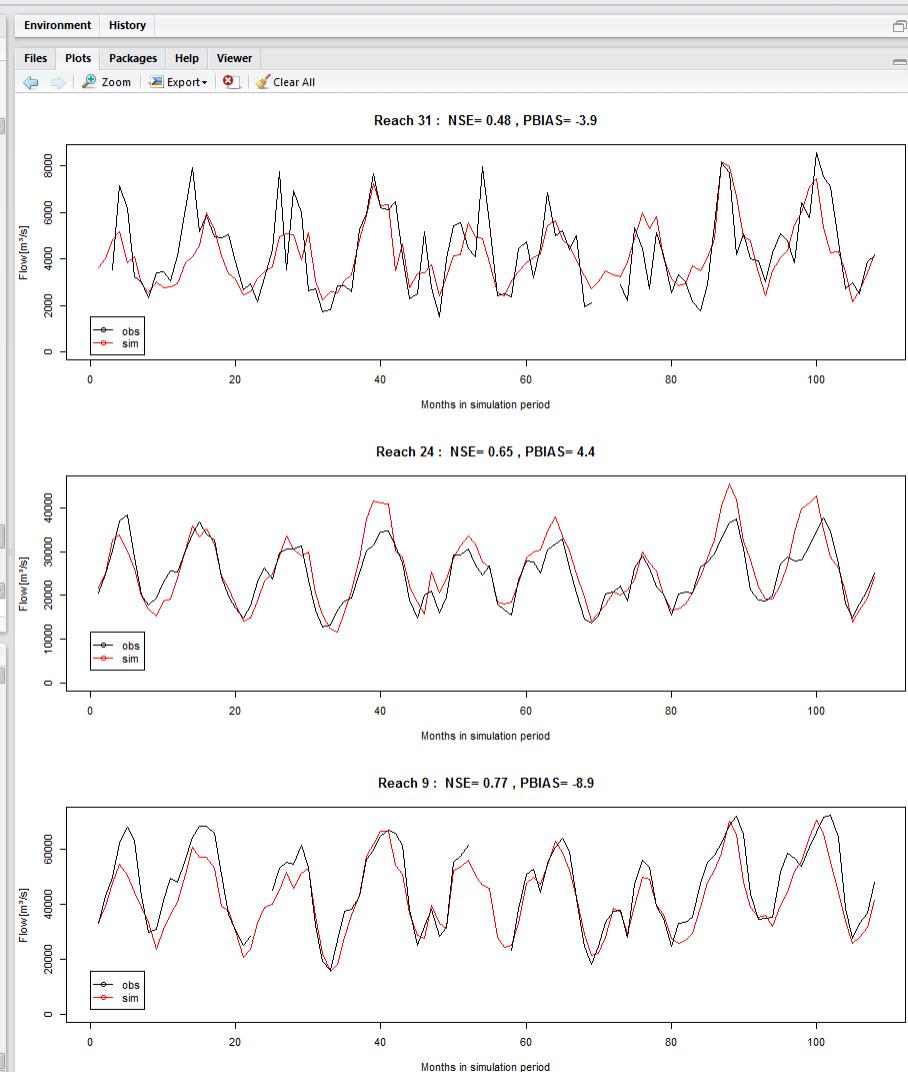
```

## Autocalibration using R

```
REACH
Next | Prev | _3 | Replace | All | Wrap | In selection | Match case | Whole word | Regex | Wrap

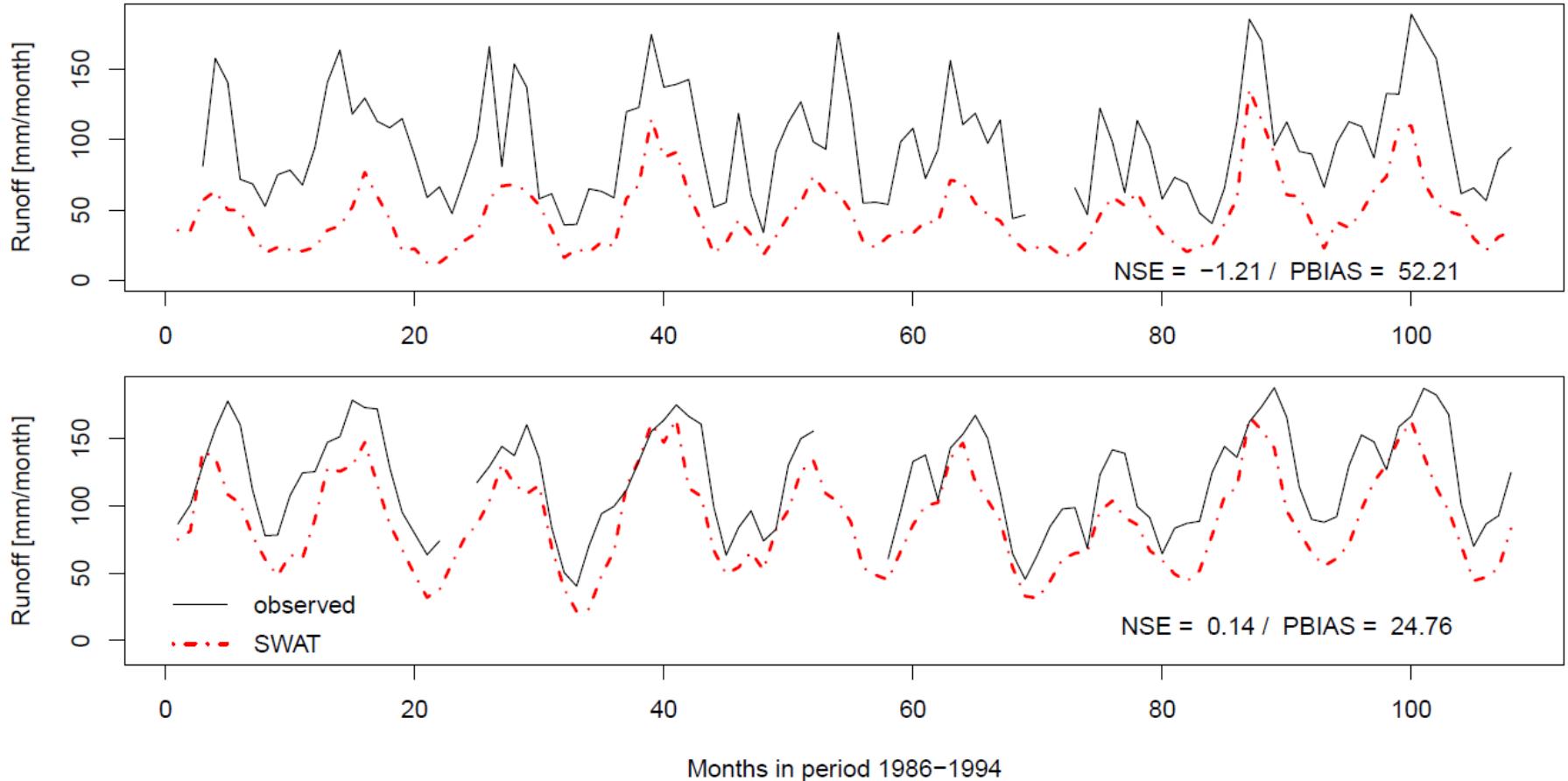
410 sim_j.mon_2 <- as.numeric(out[,1*tlength+j[1]+1):(j[1]*tlength+j[1])])
411 out.best <- as.matrix(cbind(paste(format.Date(obsdat.time,"%Y"),format.Date(obsdat.time,"%m"),sep="."))
412 ,obsdat.data,sim_j_2))
413 sim_j.mon_2 <- as.matrix(tapply(as.numeric(out.best[,3]),out.best[,1],mean,na.rm = FALSE))
414
415 #Reach 3
416 sim_j_3 <- as.numeric(out[,1*tlength+j[1]+1):(j[1]*tlength+j[1])])
417 out.best <- as.matrix(cbind(paste(format.Date(obsdat.time,"%Y"),format.Date(obsdat.time,"%m"),sep="."))
418 ,obsdat.data,sim_j_3))
419 sim_j.mon_3 <- as.matrix(tapply(as.numeric(out.best[,3]),out.best[,1],mean,na.rm = FALSE))
420
421 ##Plot monthly hydrographs
422 par(mfrow = c(3, 1))
423 x <- seq(1:length(obs.mon_1))
424 plot(x, sim_j.mon_1, type="l", ylab="Flow [m³/s]", xlab="Months in simulation period", col="red",
425 ylim=c(0,(max(sim_j.mon_1$obs.mon_1,na.rm = TRUE))), main=paste("Reach",REACH_1," NSE=",metrics.mon[j,9],
426 lines(x, obs.mon_1, type="l", xlab="", ylab="", col="black")
427 legend(0, min(obs.mon_1:sim_j.mon_1,na.rm = TRUE), c("obs", "sim"), col = c("black","red"),
428 lty = c(1, 1), pch = c(1, 1),merge = TRUE)
429
430 plot(x, sim_j.mon_2, type="l", ylab="Flow [m³/s]",xlab="Months in simulation period", col="red",
431 ylim=c(0,(max(sim_j.mon_2$obs.mon_2,na.rm = TRUE))), main=paste("Reach",REACH_2," NSE=",metrics.mon[j,42],
432 lines(x, obs.mon_2, type="l", xlab="", ylab="", col="black")
433 legend(0, min(obs.mon_2:sim_j.mon_2,na.rm = TRUE), c("obs", "sim"), col = c("black","red"),
434 lty = c(1, 1), pch = c(1, 1),merge = TRUE)
435
436 plot(x, sim_j.mon_3, type="l", ylab="Flow [m³/s]",xlab="Months in simulation period", col="red",
437 ylim=c(0,(max(sim_j.mon_3$obs.mon_3,na.rm = TRUE))), main=paste("Reach",REACH_3," NSE=",metrics.mon[j,75],
438 lines(x, obs.mon_3, type="l", xlab="", ylab="", col="black")
439 legend(0, min(obs.mon_3:sim_j.mon_3,na.rm = TRUE), c("obs", "sim"), col = c("black","red"),
440 lty = c(1, 1), pch = c(1, 1),merge = TRUE)
```

```
Console: C:\R\win/MIP/R/autocal/ftntout3.R
> y=c(0,(max(sim_j.mon_1,obs.mon_1,na.rm = TRUE)), main=paste("Reach",REACH_1,": NSE=",metrics.mon[j,1],sep=""))
> lines(x, obs.mon_1,type="l", xlab="", ylab="", col="black")
> legend(0, min(obs.mon_1,sim_j.mon_1,na.rm = TRUE), c("obs","sim"), col = c("black","red"),
+       lty = c(1, 1), pch = c(1, 1),merge = TRUE)
> par(mfrow = c(3, 1))
> X <- seq(1:length(obs.mon_1))
> plot(x, sim_j.mon_1,type="l", ylab="Flow [m³/s]", xlab="Months in simulation period", col="red",
+       ylim=c(0,(max(sim_j.mon_1,obs.mon_1,na.rm = TRUE))), main=paste("Reach",REACH_1,": NSE=",metrics.mon[j,9]," , PB
IAS=",metrics.mon[j,6],sep=""))
> lines(x, obs.mon_1,type="l", xlab="", ylab="", col="black")
> legend(0, min(obs.mon_1,sim_j.mon_1,na.rm = TRUE), c("obs","sim"), col = c("black","red"),
+       lty = c(1, 1), pch = c(1, 1),merge = TRUE)
>
> plot(x, sim_j.mon_2,type="l", ylab="Flow [m³/s]",xlab="Months in simulation period", col="red",
+       ylim=c(0,(max(sim_j.mon_2,obs.mon_2,na.rm = TRUE))), main=paste("Reach",REACH_2,": NSE=",metrics.mon[j,42]," , P
BIAS=",metrics.mon[j,39],sep=""))
> lines(x, obs.mon_2,type="l", xlab="", ylab="", col="black")
> legend(0, min(obs.mon_2,sim_j.mon_2,na.rm = TRUE), c("obs","sim"), col = c("black","red"),
+       lty = c(1, 1), pch = c(1, 1),merge = TRUE)
>
> plot(x, sim_j.mon_3,type="l", ylab="Flow [m³/s]",xlab="Months in simulation period", col="red",
+       ylim=c(0,(max(sim_j.mon_3,obs.mon_3,na.rm = TRUE))), main=paste("Reach",REACH_3,": NSE=",metrics.mon[j,75]," , P
BIAS=",metrics.mon[j,72],sep=""))
> lines(x, obs.mon_3,type="l", xlab="", ylab="", col="black")
> legend(0, min(obs.mon_3,sim_j.mon_3,na.rm = TRUE), c("obs","sim"), col = c("black","red"),
+       lty = c(1, 1), pch = c(1, 1),merge = TRUE)
> |
```

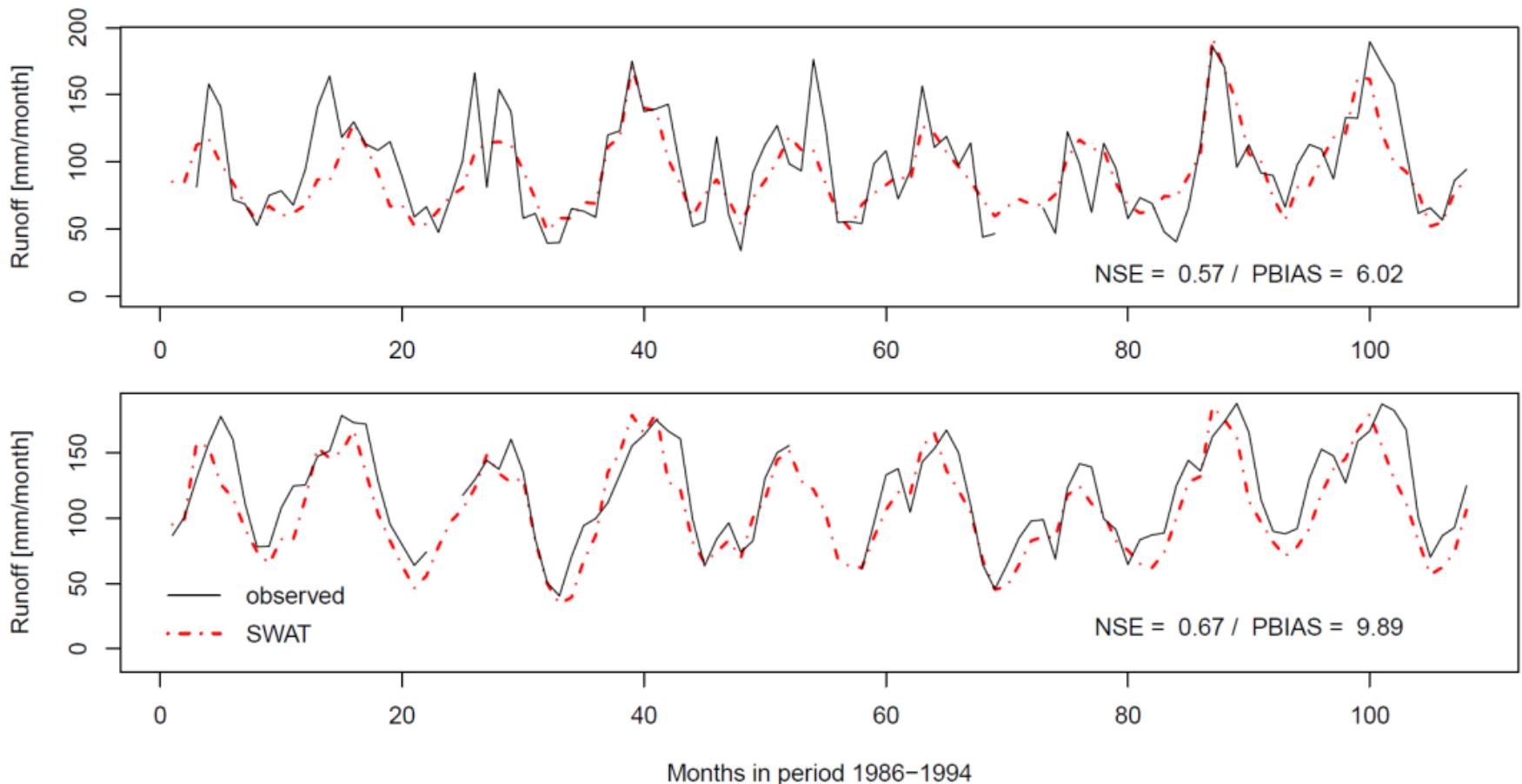


# Autocalibration using R

## Initial simulations – problem with water balance



## Simulations considering elevation bands



## Calibrated parameter values

Regional	Mountains	Lowlands I	Lowlands II	Basin wide	
GW_DELAY	25.36	60.24	14.33	SFTMP	3.88
ALPHA_BF	0.03	0.08	0.12	SURLAG	0.12
GWQMN	875.75	0.00	0.00	MSK_CO1	0.14
GW_REVAP	0.02	0.06	0.01	MSK_CO2	1.46
REVAPMN	409.76	135.35	0.72	MSK_X	0.01
RCHRG_DP	0.00	0.08	0.00		
OV_N	0.10	0.11	0.10		
EPCO	1.00	0.01	1.00		
CH_N2	0.01	0.03	0.01		
CN2	0.00	-0.21	0.00		
PLAPS	648.49	353.86	334.96		
TLAPS	-6.73	-7.16	-8.20		