R语言在陆面建模系统上的应用研究

罗立辉

luolh@lzb.ac.cn

中国科学院寒区旱区环境与工程研究所

天山站
临泽站
阿拉善站
皋兰站
奈曼站
敦煌站
庙尔沟冰川
七一冰川
祁连站
沙坡头站
平凉站
格尔木站
黑河上游站
北麓河站
甘南站
绒布冰川
绒布冰川
绒布冰川
玉龙站
图例
非寒区
寒区
半干旱区
干旱区
基础定位站
新增定位站
冰川半定位站
冻土监测点
沙尘监测点

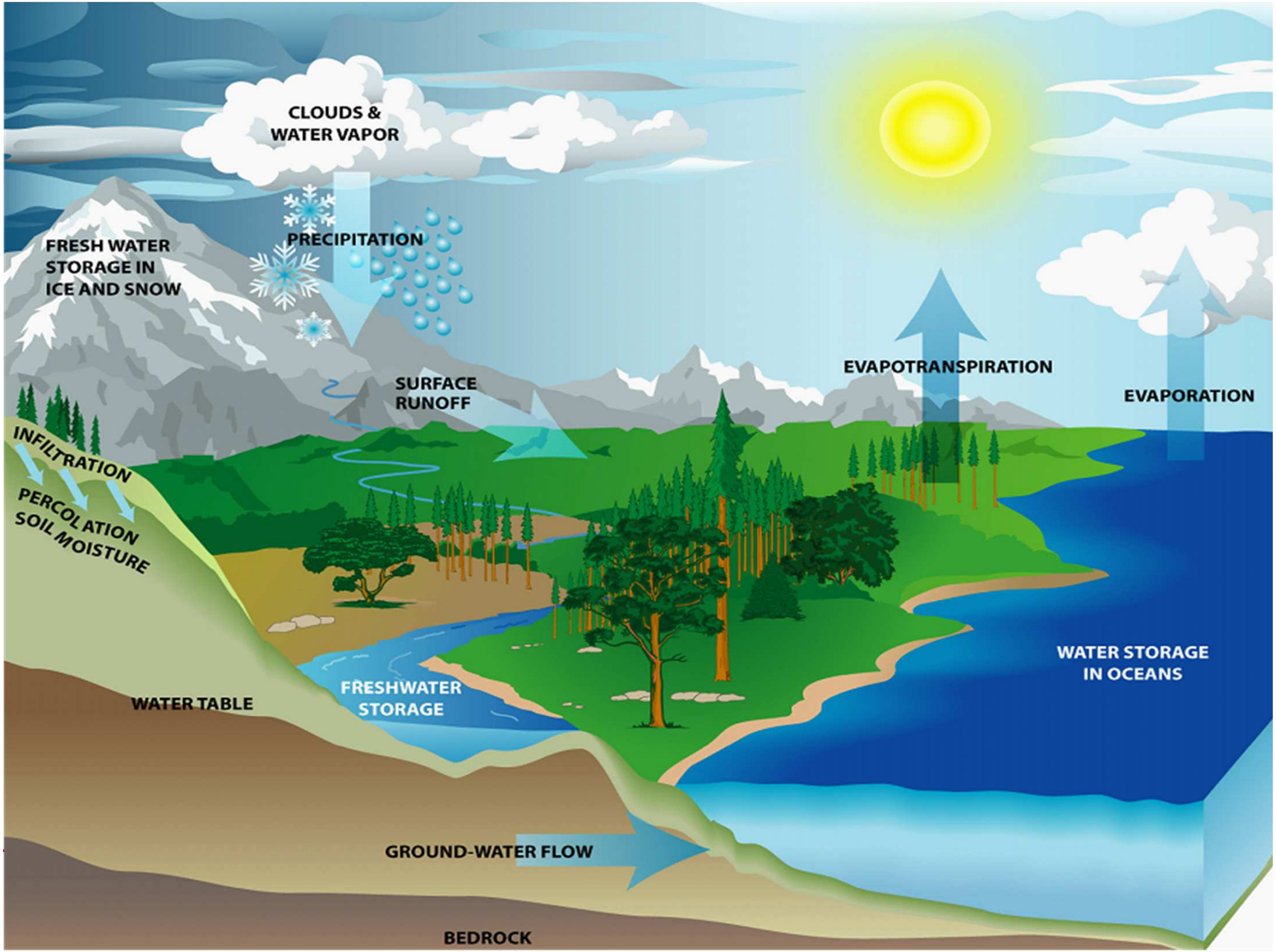
中国西北行，出发在兰州

报告内容

- 1 研究背景
- 2 陆面建模系统的关键技术
- 3 R语言在陆面建模系统上的应用
- 4 其它脚本语言集成
- 5 研究展望

研究背景

“大气边界层1500m以下到地表以下几十米”


CLOUDS &
WATER VAPOR
FRESH WATER
STORAGE IN
ICE AND SNOW
PRECIPITATION
SURFACE
RUNOFF
EVAPOTRANSPIRATION
EVAPORATION
INFILTRATION
PERCOL ATION
SOIL MOISTURE
WATER TABLE
FRESHWATER
STORAGE
WATER STORAGE
IN OCEANS
GROUND-WATER FLOW
BEDROCK

研究背景/陆面过程模型

模型通用组件	典型的陆面过程模型举例
系统边界	低层大气，土壤厚度， 土壤质地
强迫输入	气温、短波和长波辐射，降雨，风速，气压、大气CO2浓度
初始状态	生物地球化学池（碳、氮和甲烷），植被（PFT, LAI, SAI），土壤温度和湿度
参数	化学过程速率常量，物理常量，生物参数
模型结构	流程定义及连接
模型状态	生物地球化学池（碳、氮和甲烷），植被（PFT，LAI，SAI），土壤温度和湿度
模拟输出	生物地球化学通量（潜热、感热、净辐射、地热、CO2等），模型状态的动态变化

碳排放->碳交易

工业->欧洲

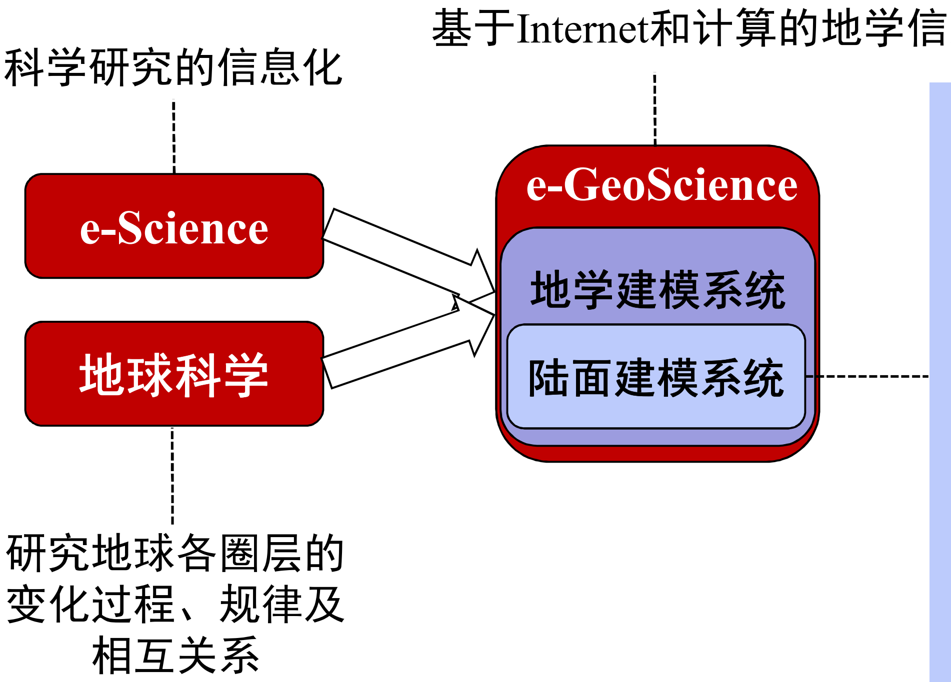
科技革命->美国

研究背景/陆面过程模型

时期	年代	特点	代表模型
第一阶段	70s	无清晰的植被处理； “水箱”水文	Bucket[1]， MIAMI[2] Köppen[3], Holdridge[3]
第二阶段	80s	“大叶”方法； 多层水文； 植被功能型的概念； 植物竞争； 卫星遥感数据；	BATS[1], SiB[1], ISBA[1],SECHIBA[1] CASA[2], TURC[2], TEM[2], BIOME-BGC[2], CARAIB[2], SILVAN[2], CENTURY[2], FBM[2] BIOME[3]
第三阶段	90s	植物生理；**碳循环**； 植被功能型；次网格； 植被动态； “二叶”方法	SiB2[12], LSM[12], MOSES[12] DEMETER[23], BIOME2/3[23], LPJ[23] IBIS[123], **CLM**[123], CoLM[123], ORCHIDEE[123], **JSBACH**[123]

陆面过程模型的发展阶段。1 物理过程模块；2 生物地球化学模块；3 生物地理模块

研究背景/陆面建模系统



- 在陆面过程模型的基础上
- 集成地球观测系统、陆面模拟、高性能计算、数据处理和分析方法，以及可视化等技术手段
- 从多角度来精确评估陆面动量、能量和水分状况等陆面特征
- 为改善陆面过程模型的模拟性能提供指导
- 更为重要的是为研究者提供一套完整的从数据处理到模拟分析的陆面建模环境

开展集成的自动的数据处理、模拟分析和可视化等理论与技术手段支撑的有关陆面建模系统的开发与研究已成为国际地球科学最活跃的研究前沿

国内外主要研究进展—地学建模系统

- 近10年来地球科学建模系统将各种已有的有关地学的（包含陆面过程模型）专用软件框架和工具包加以统一、标准化和扩展，从而达到充分共享资源、快速构建地学建模应用的目的
- 目前地学建模系统的主要研究内容包括：
  - 数据管理、交换、转化（前处理）
  - 模型组件各部分的连接、各种软件工具库、脚本库
  - 数据同化或者模型数据融合
  - 模拟输出的数据处理、结果分析与可视化（后处理）
  - 模拟结果的共享与访问
- 地球科学建模系统减轻了地球科学研究者的软件开发负担，使他们得以腾出更多精力用于科学研究
- 对应用软件可重用性、互操作性及性能可移植性的支持，有效地促进了不同机构间的交流与合作，加快了地球科学研究成果向业务应用的转化

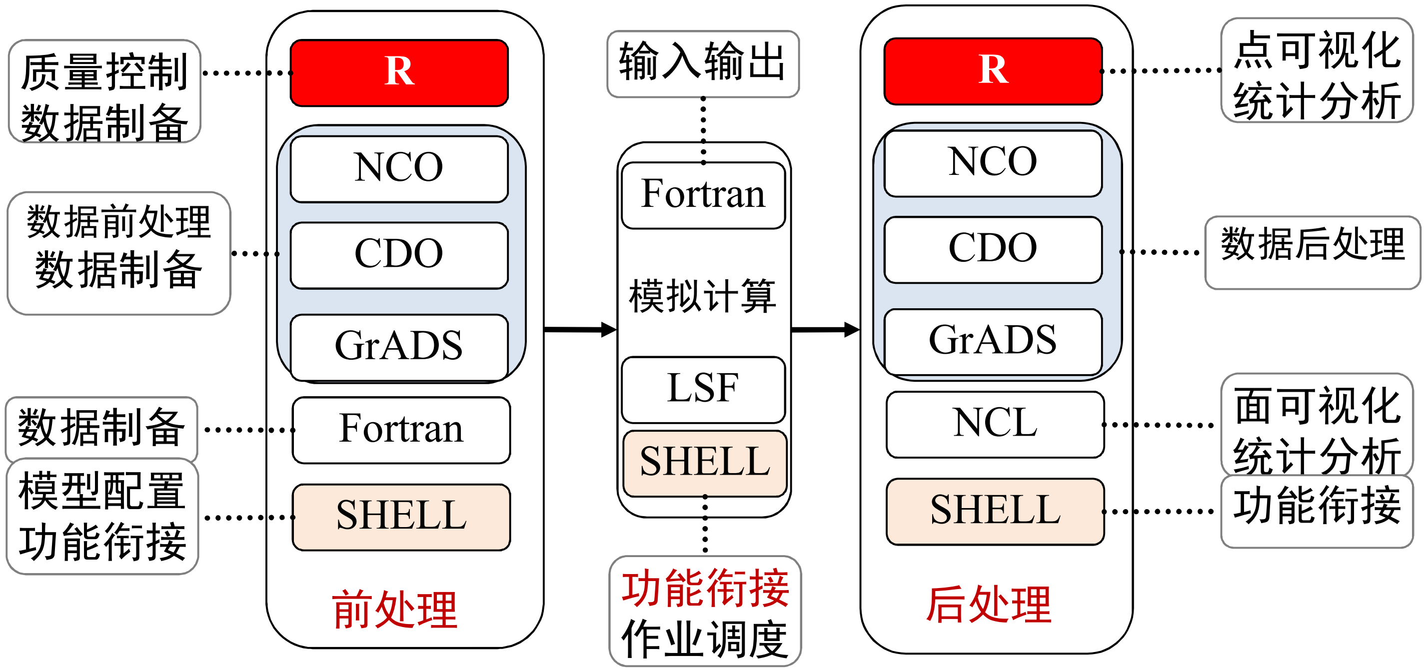
AUTOMATIC?

陆面建模系统的关键技术

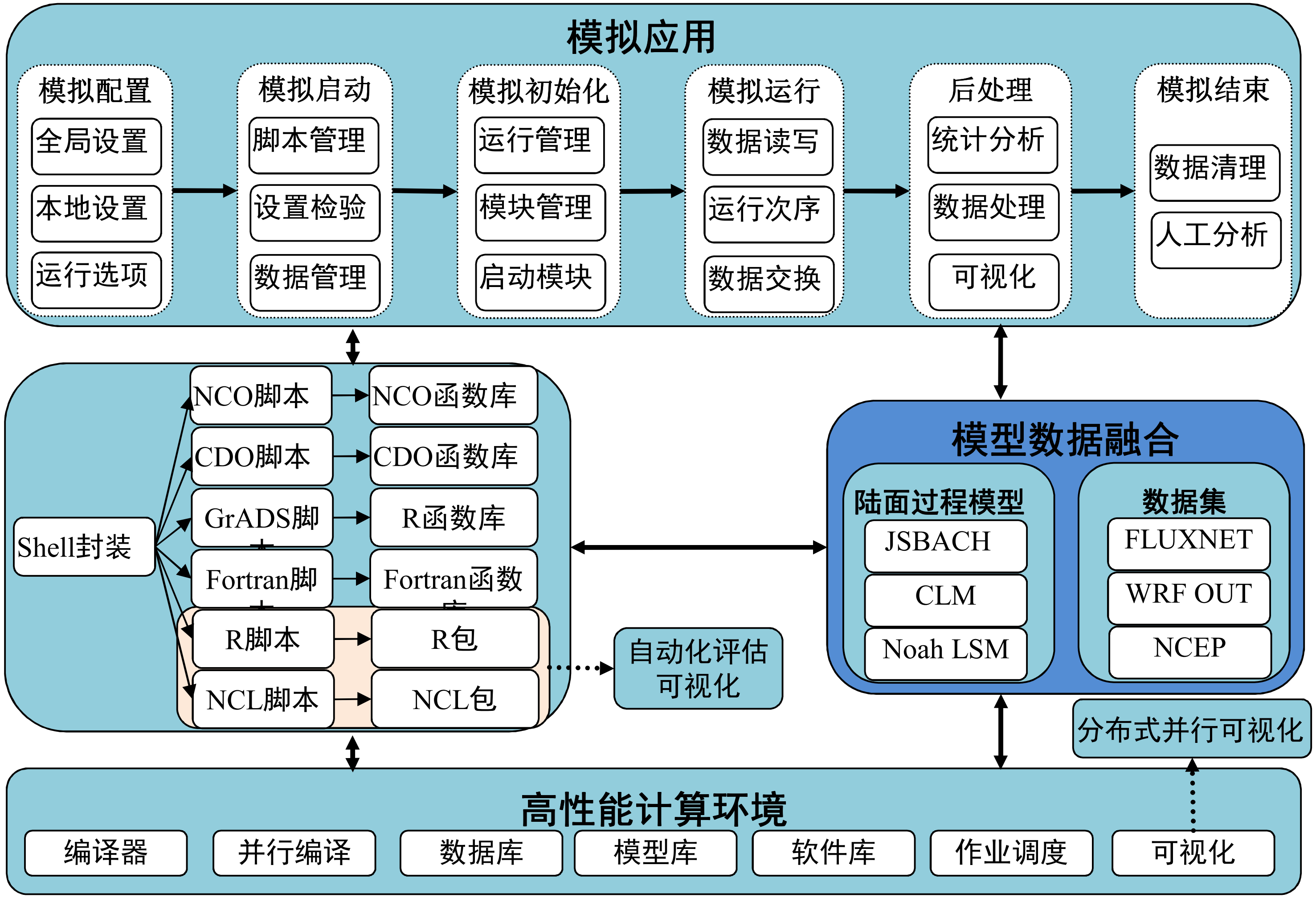
陆面建模系统关键技术：基于脚本语言的建模

◆ 为什么要采用基于脚本语言的建模？而不是GUI！

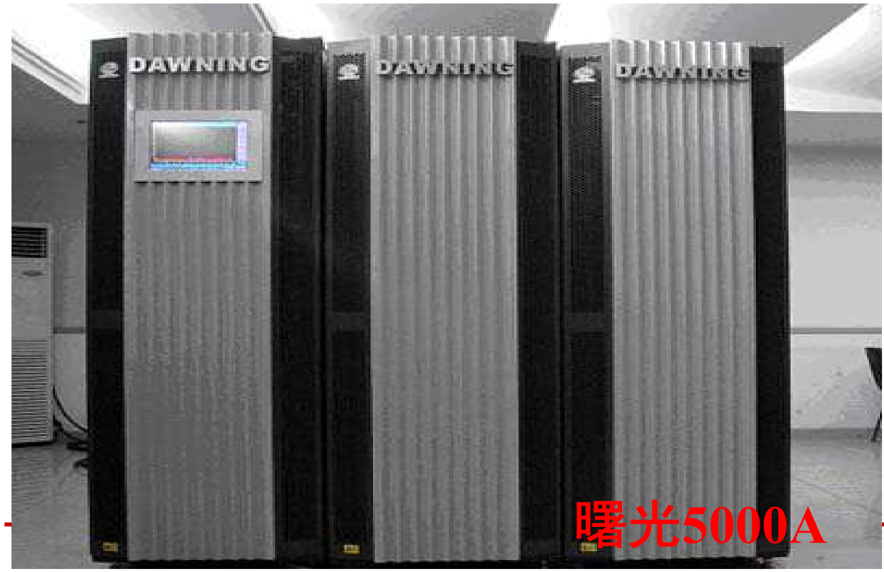
- 由高性能计算的计算模式所决定（SSH连接和job调度）
- 由陆面过程模型的代码和模拟运行所决定（Fortran代码和SHELL控制）
- 开源脚本语言的支持（R、NCL、GrADS、NCO、CDO …）

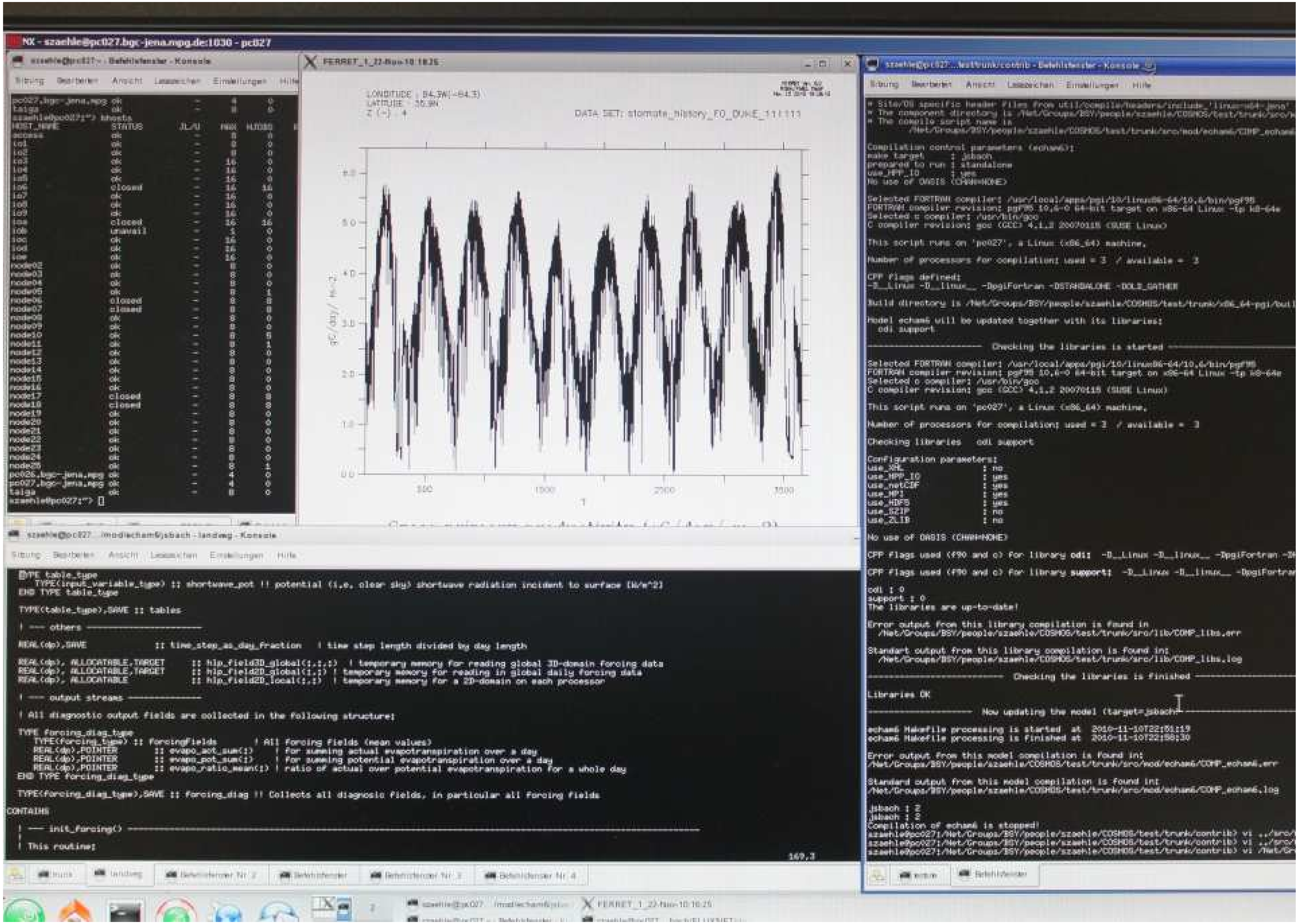


陆面建模系统关键技术—系统结构图



R离不开高性能计算环境

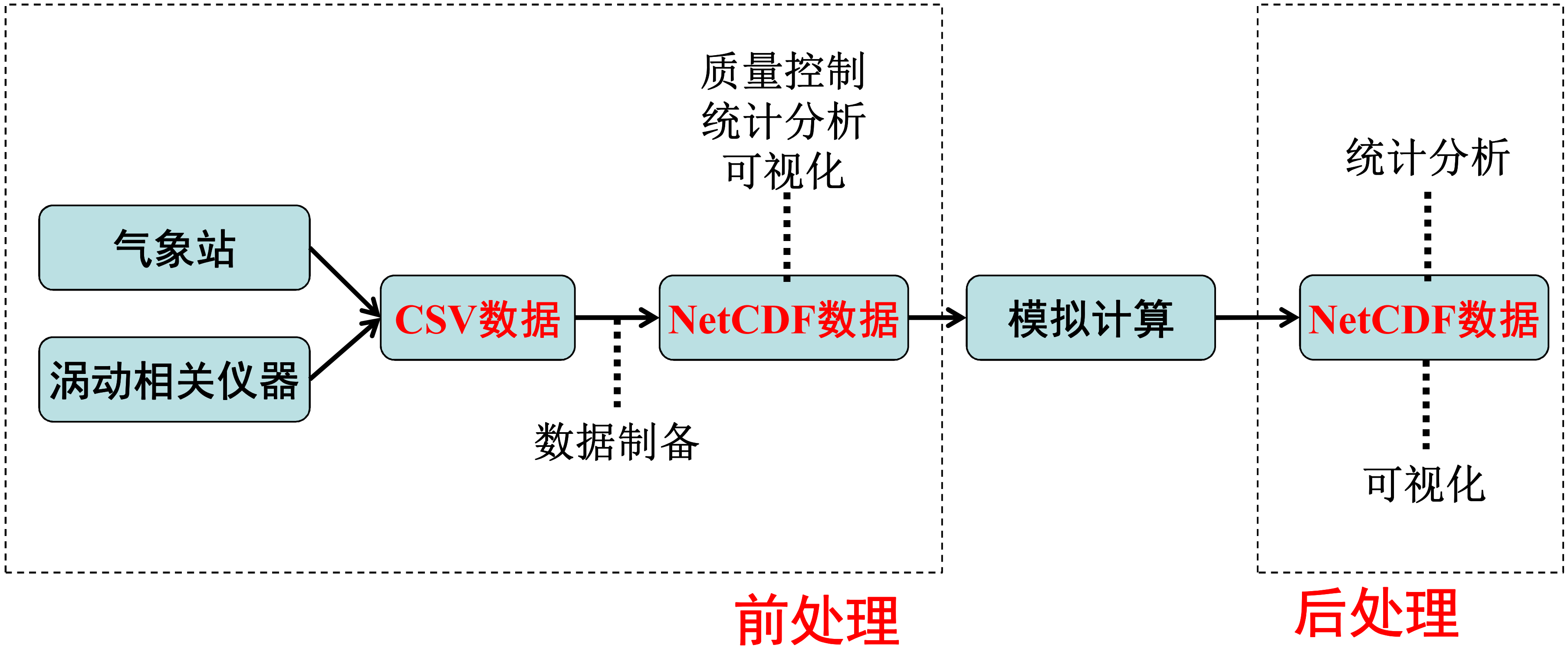




NX - szaehle@pc027.bgc-jena.mpg.de:1030 - pc027
FERRET_1_22-Nov-10 18:25
LONGITUDE : 84.3W(-84.3)
LATITUDE : 35.9N
DATA SET: stomate_history_FO_DUKE_11111
Checking the libraries is started
The libraries are up-to-date!
Checking the libraries is finished
Libraries OK
Compilation of echam6 is stopped!

R语言在陆面建模系统上的应用

陆面过程模型的前、后处理

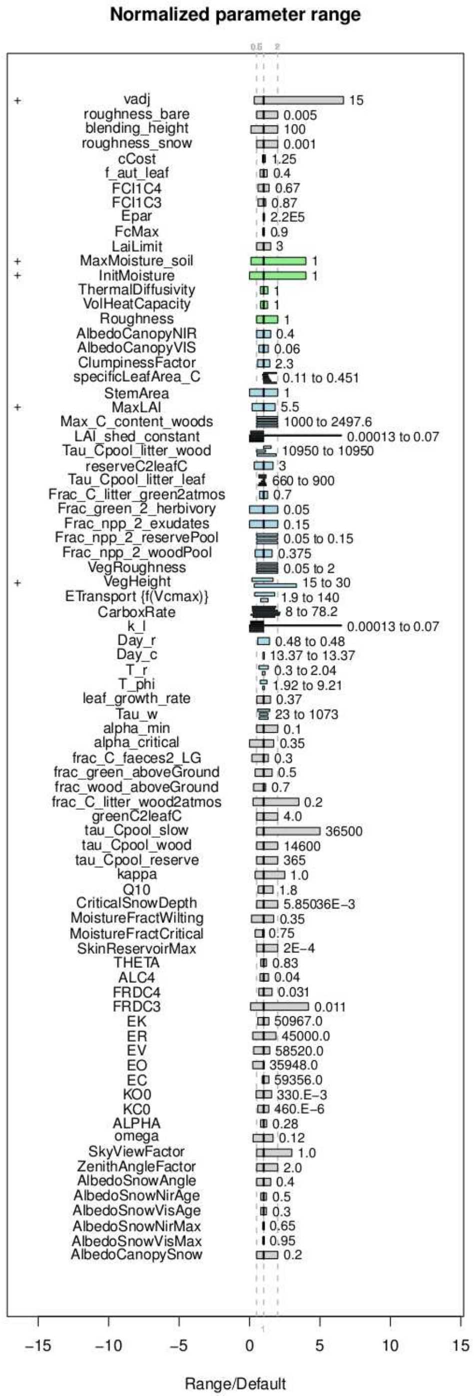


输入

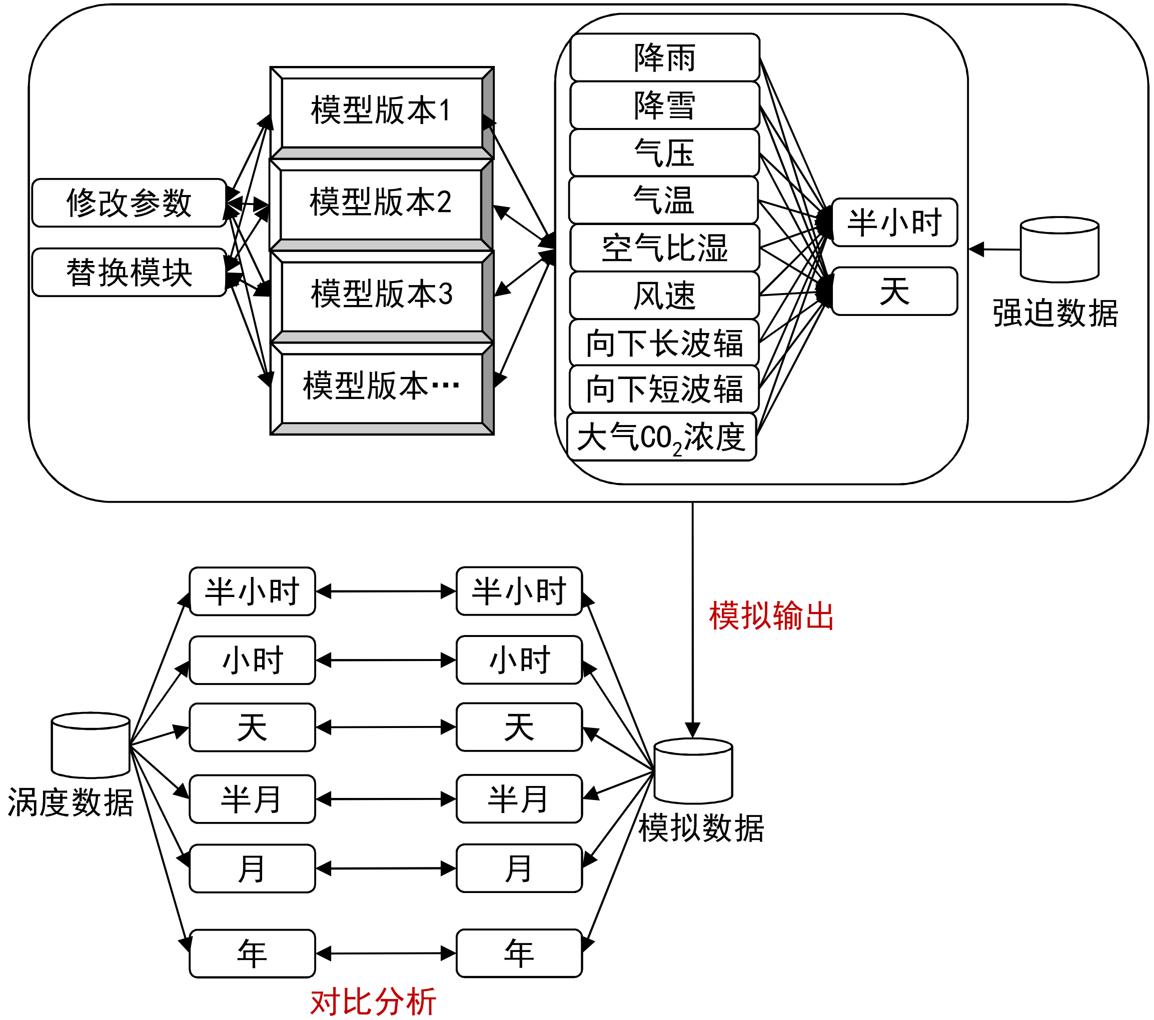
变量	单位	说明
水文条件:		
qair	1	最底层的空气比湿
precip_rain	Kg/m^2s	降雨
precip_snow	Kg/m^2s	降雪
热量条件:		
temp_air	K	最底层的空气温度
力学条件:		
Wind	m/s	最底层的风速
wind10	m/s	地面10m处的风速
cdrag	1	表面拖曳系数
pressure	Kg/m^2	表面大气压
辐射条件:		
lwdown	W/m^2	向下长波辐射
swdown	W/m^2	向下短波辐射
sw_vis_net	W/m^2	净可见光辐射
sw_vis_frac_diffuse	1	净可见光辐射中散射比例
sw_nir_net	W/m^2	净近红外辐射
czenith	1	太阳天顶角余弦
湍流通量条件:		
etAcoef		richtmyer-morton A系数（热能）
etBcoef		richtmyer-morton B系数（热能）
eqAcoef		richtmyer-morton A系数（水汽通量）
eqBcoef		richtmyer-morton B系数（水汽通量）
echam_zchl	1	湍流热传输系数
碳循环条件:		
CO2_concentration	$Kg(CO_2)/Kg(air)$	最底层CO2浓度
表面条件:		
albedo	1	土壤反照率
albedo_soil_nir	1	近红外光谱土壤反射率
albedo_soil_vis	1	可见光光谱土壤反射率
albedo_veg_nir	1	近红外光谱植被反射率
albedo_veg_vis	1	可见光光谱植被反射率

输出

缩写	中文名
PPFD	光合有效辐射
Rn	净辐射
Qh	感热
Qle	潜热
gscan	冠层导率
GPP	总初级生产力
Reco	生态系统呼吸
NEE	净生态系统碳交换量
fAPAR	吸收光合有效辐射比例
vegC	植物的碳池
litterC	枯枝落叶的碳池
soilC	土壤有机质的碳池


Normalized parameter range
vadj
roughness_bare
blending_height
roughness_snow
cCost
f_aut_leaf
FCl1C4
FCl1C3
Epar
FcMax
LaiLimit
MaxMoisture_soil
InitMoisture
ThermalDiffusivity
VolHeatCapacity
Roughness
AlbedoCanopyNIR
AlbedoCanopyVIS
ClumpinessFactor
specificLeafArea_C
StemArea
MaxLAI
Max_C_content_woods
LAI_shed_constant
Tau_Cpool_litter_wood
reserveC2leafC
Tau_Cpool_litter_leaf
Frac_C_litter_green2atmos
Frac_green_2_herbivory
Frac_npp_2_exudates
Frac_npp_2_reservePool
Frac_npp_2_woodPool
VegRoughness
VegHeight
ETransport {f(Vcmax)}
CarboxRate
k_l
Day_r
Day_c
T_r
T_phi
leaf_growth_rate
Tau_w
alpha_min
alpha_critical
frac_C_faeces2_LG
frac_green_aboveGround
frac_wood_aboveGround
frac_C_litter_wood2atmos
greenC2leafC
tau_Cpool_slow
tau_Cpool_wood
tau_Cpool_reserve
kappa
Q10
CriticalSnowDepth
MoistureFractWilting
MoistureFractCritical
SkinReservoirMax
THETA
ALC4
FRDC4
FRDC3
EK
ER
EV
EO
EC
KO0
KC0
ALPHA
omega
SkyViewFactor
ZenithAngleFactor
AlbedoSnowAngle
AlbedoSnowNirAge
AlbedoSnowVisAge
AlbedoSnowNirMax
AlbedoSnowVisMax
AlbedoCanopySnow
15
0.005
100
0.001
1.25
0.4
0.67
0.87
2.2E5
0.9
3
1
1
1
1
1
0.4
0.06
2.3
0.11 to 0.451
1
5.5
1000 to 2497.6
0.00013 to 0.07
10950 to 10950
3
660 to 900
0.7
0.05
0.15
0.05 to 0.15
0.375
0.05 to 2
15 to 30
1.9 to 140
8 to 78.2
0.00013 to 0.07
0.48 to 0.48
13.37 to 13.37
0.3 to 2.04
1.92 to 9.21
0.37
23 to 1073
0.1
0.35
0.3
0.5
0.7
0.2
4.0
36500
14600
365
1.0
1.8
5.85036E-3
0.35
0.75
2E-4
0.83
0.04
0.031
0.011
50967.0
45000.0
58520.0
35948.0
59356.0
330.E-3
460.E-6
0.28
0.12
1.0
2.0
0.4
0.5
0.3
0.65
0.95
0.2
-15
-10
-5
0
5
10
15
Range/Default

全面评估



陆面建模系统中的R功能函数

全面评估：

- 不同时间尺度
- 不同站点
- 不同变量在不同时间尺度
- 不同时间尺度强迫数据
- 不同模型版本
- 小波分析
- 植被高度的敏感性分析

模型改进：

- 水箱模型和五层土壤对比分析

R功能函数

- 前处理，R采用相邻7值的方法质量控制，设置旗标，RNetCDF
- 后处理，ncdf、gdata、gplots、plotrix、Hmisc、lattice和泰勒R包
- 代码普适性，模拟结果参数等的变化调整其代码
- SHELL可以根据实际需要调用R脚本，以pdf、eps、emf、ps等格式输出，利用x11()的RGui的图形界面输出

图形函数：	
JSBACH_FLUXNET_GRAPHIC	基本的图像函数，包括绘图函数plo 差分析，箱图等功能。
JSBACH_FLUXNET_GRAPHIC_FAPAR	专门为fAPAR数据定制的图像函数
JSBACH_FLUXNET_GRAPHIC_FORCING	与强迫数据集进行对比分析的图形函
JSBACH_FLUXNET_GRAPHIC_ERR	误差分析的图形函数
JSBACH_FLUXNET_STATS_GRAPHIC	统计分析的图形函数
JSBACH_FLUXNET_TAYLOR	泰勒图
GWS_ANALYSIS	小波分析
数据函数：	
JSBACH_DATA	读写模拟结果
FLUXNET_DATA	读写FLUXNET数据
时间操作：	
JSBACH_FLUXNET_COMPUTE_GROUPING	不同时间分辨率的转换
模型评估：	
JSBACH_FLUXNET_EVALUATION	15种统计分析评估
常规函数：	
CONVERT_SITE_TO_PARAMETER	站点与参数之间的数据类型转换
DIFF_VERSION_COMPARE	比较不同的模型版本的模拟结果
CALC_RELATIVE_HUMIDITY_IN_SOIL	计算不同土壤模型的相对湿度（水和 层土壤模型）
LONLAT	经纬度
主程序：	
MAIN_PROGRAM	控制评估流程及输出模式
DIFF_PARAMETER	不同参数间的对比分析
DIFF_SITE	不同站点间的对比分析

数据制备 CSV→NetCDF

```
#!/usr/bin/env Rscript # command line run in shell
rm(list = ls())        # clear objects
graphics.off()         # close graphics windows
library(ncdf)
library(RNetCDF)
HA94 = read.table('HA94_hh.met', head=T)
netcdf.from.fluxnet <- create.nc("DE-Hai.2000-
2006.obs.halfhourly.fluxnet.nc")
# dimensions
dim.def.nc(netcdf.from.fluxnet, "lon", 1)
dim.def.nc(netcdf.from.fluxnet, "lat", 1)
dim.def.nc(netcdf.from.fluxnet, "time", unlim=TRUE)
```

数据制备 CSV→NetCDF

```
# variables
var.def.nc(netcdf.from.fluxnet, "NEE", "NC_DOUBLE",
c(0,1))
# variables Attribution
att.put.nc(netcdf.from.fluxnet, "NEE", "long_name",
"NC_CHAR", "gapfilled Net Ecosystem Exchange")
att.put.nc(netcdf.from.fluxnet, "NEE", "units",
"NC_CHAR", "umol m-2 s-1")
att.put.nc(netcdf.from.fluxnet, "NEE",
"missing_value", "NC_DOUBLE", -9999.)
att.put.nc(netcdf.from.fluxnet, "NEE", "long_name",
"NC_CHAR", "gapfilled Net Ecosystem Exchange")
```

数据制备 CSV→NetCDF

```
# Define variable values
LON_DATA <- HA94$LON
LAT_DATA <- HA94$LAT
TIME_DATA <- HA94$TIME
NEE_DATA <- HA94$NEE
# Put the data with default start/count
var.put.nc(netcdf.from.fluxnet, "LON", LON_DATA)
var.put.nc(netcdf.from.fluxnet, "LAT", LAT_DATA)
var.put.nc(netcdf.from.fluxnet, "TIME", TIME_DATA)
var.put.nc(netcdf.from.fluxnet, "NEE", NEE_DATA)
close.nc(netcdf.from.fluxnet)
```

质量控制

- 采用7邻方法对相邻的7个数据进行检测画图，以判断奇异值。

```
for ( n in 1:nver )
lines(TIME, JSBACH_V[n,], col='pink' ,pch=19)
for ( n in 1:nver )
lines(TIME, filter(JSBACH_V[n,],rep(1/7,7)),
col=colc[n] ,pch=19)
lines(TIME, filter(FLUXNET_V,rep(1/7,7)), col="black"
```

- **其它的数据质量检验的方法**

- 观测数据因为仪器或者别的原因会有空值、测量误差等，需要在变量的基础上再增加一个关于变量flag的变量，如var1_flag(变量var1)

  0=Original, 1=Diurnal_mean_fill, 2=Daymet,
  3=Daymet_and_dailyNCDC, 4=dailyNCDC, 6=hourlyNCDC,
  7=nearby_tower, 8=multiple_var

自动化评估技术

每一种数学统计的方法应用于模型模拟性能的评价都有其**局限性**，所以在陆面建模系统的模型评估中采用多种评价方法相结合（可视化也存在着同样的局限性）

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S_i - O_i)^2}{n}}$$

$$MEF = \frac{\left[\sum_{i=1}^{n}(O_i - \overline{O})^2 - \sum_{i=1}^{n}(S_i - O_i)^2\right]}{\left[\sum_{i=1}^{n}(O_i - \overline{O})^2\right]}$$

$$\mathrm{VR} = \frac{\mathrm{V_p}^2}{\mathrm{V_o}^2}$$

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{\overline{O}}$$

$$RSR = \frac{RMSE}{STDEV_O} = \frac{\left[\sqrt{\sum_{i=1}^{n}(O_i - S_i)^2}\right]}{\left[\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}\right]}$$

$$RS = \frac{\left[\sum_{i=1}^{n}(O_i - \overline{O})^2\right]}{\left[\sum_{i=1}^{n}(S_i - \overline{O})^2\right]}$$

$$NAE = \left[\frac{(\overline{S} - \overline{O})}{\overline{O}}\right]$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2}\right]$$

$$PBIAS = \left[\frac{\sum_{i=1}^{n}(O_i - S_i)*100}{\sum_{i=1}^{n}(O_i)}\right]$$

总共13种数学统计方法

S_i表示i时刻的模拟值

O_i表示i时刻的观测值

$\overline{S}$是模拟值的平均值

$\overline{O}$是观测值的平均值

n 为样本数

自动化评估技术

```
JSBACH_FLUXNET_EVALUATION <- function(method, JSBACH_V,
FLUXNET_V){

  d1=FLUXNET_V[!is.na(FLUXNET_V)&!is.na(JSBACH_V)]
  d2=JSBACH_V[!is.na(FLUXNET_V)&!is.na(JSBACH_V)]

  if(length(d1)>0) {

   # Root Mean Square Error
   if (method=="rmse") r_value <- sqrt(mean((d1-d2)^2,
na.rm=TRUE))

   # Normalized RMSE
   if (method=="nrmse") r_value <- (sqrt(mean((d1-d2)^2,
na.rm=TRUE)))/mean(d1)

   # Coefficient of determination (R Squared)
   if (method=="RR") r_value <- summary(lm(d2 ~ d1,
na.action=na.omit))[c("r.squared")]
```

自动化评估技术

```
    # Pearson's correlation coefficient
    if (method=="cor_test") r_value <- cor.test(d1, d2,
na.action=na.omit )$p.value
      #cor.test(FLUXNET_V, JSBACH_V, method="spearman")
      #cor.test(FLUXNET_V, JSBACH_V, method="kendall")

    # Nash-Sutchliffe efficiency (NSE)
    if (method=="nse") r_value <- 1- sum((d1 - d2)^2)/sum((d1 -
mean(d2))^2)

    # RMSE-observations standard deviation ratio (RSR)
    if (method=="rsr") r_value <- sqrt(sum((d1 -
d2)^2))/sqrt(sum((d1 - mean(d2))^2))

    # Percent bias (PBIAS)
    if (method=="pbias") r_value <- sum(d2 - d1)*100/sum(d1)

    # Modelling efficiency
    if (method=="mef") r_value <- (sum((d1-mean(d1))^2)-sum((d2-
d1)^2))/sum((d1-mean(d1))^2)
```

自动化评估技术

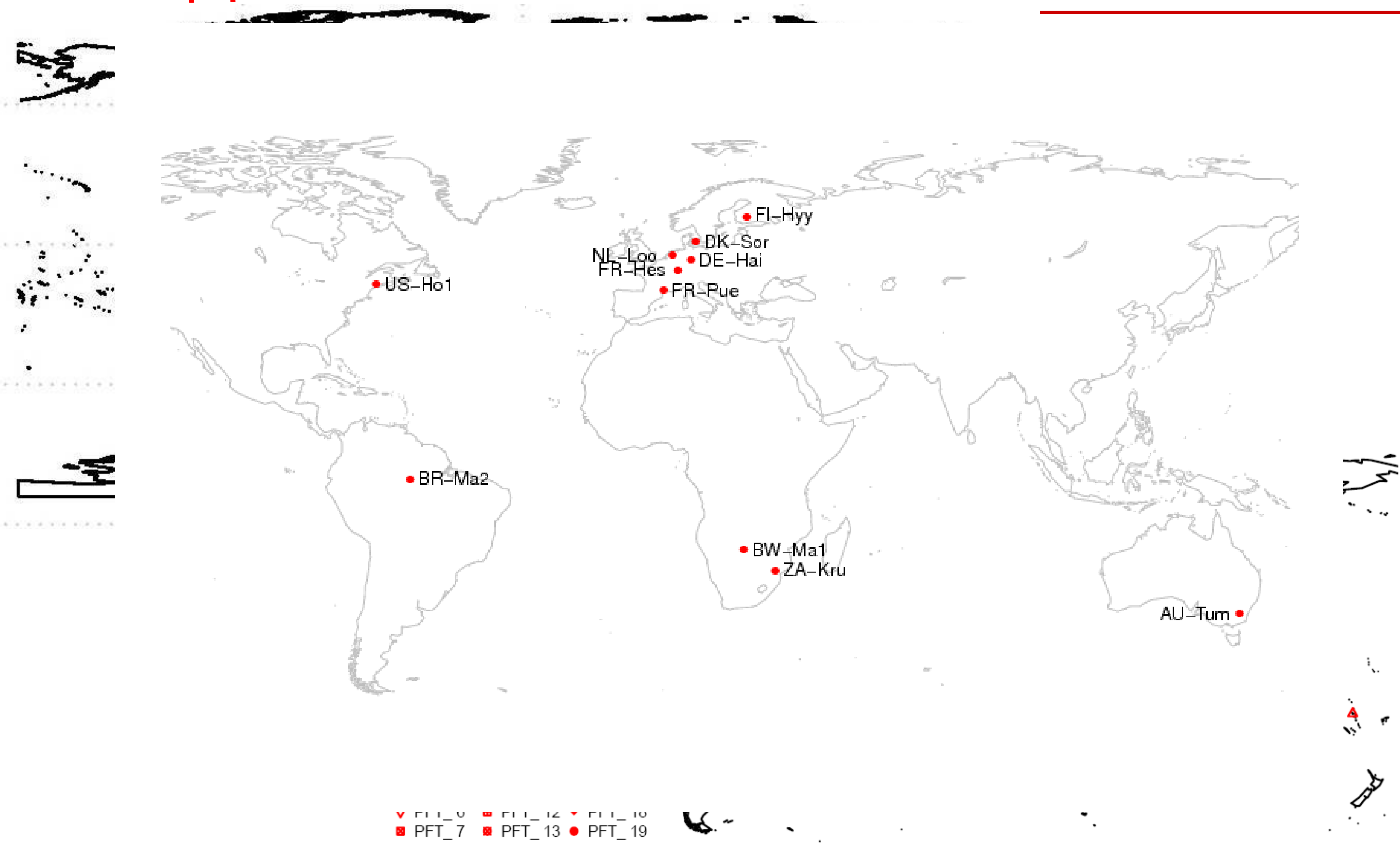
```
    # Normalized average error
    if (method=="nae") r_value <- (mean(d2)-mean(d1))/mean(d1)

    # Variance Ratio
    if (method=="vr") r_value <- var(d2)/var(d1)

  } else {
    r_value <- NA
  }

  return(sprintf("%.2f", r_value))
}
```

GIS图



FI–Hyy
DK–Sor
NL–Loo
FR–Hes
DE–Hai
FR–Pue
US–Ho1
BR–Ma2
BW–Ma1
ZA–Kru
AU–Tum
PFT_7
PFT_13
PFT_19

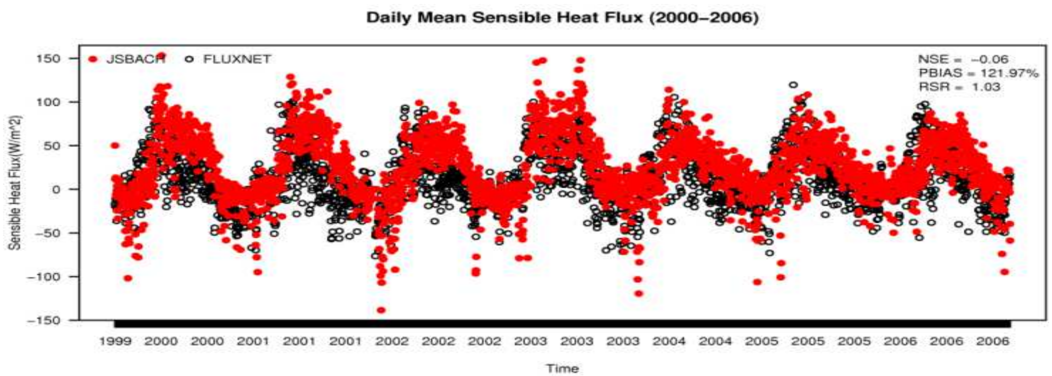
GIS图

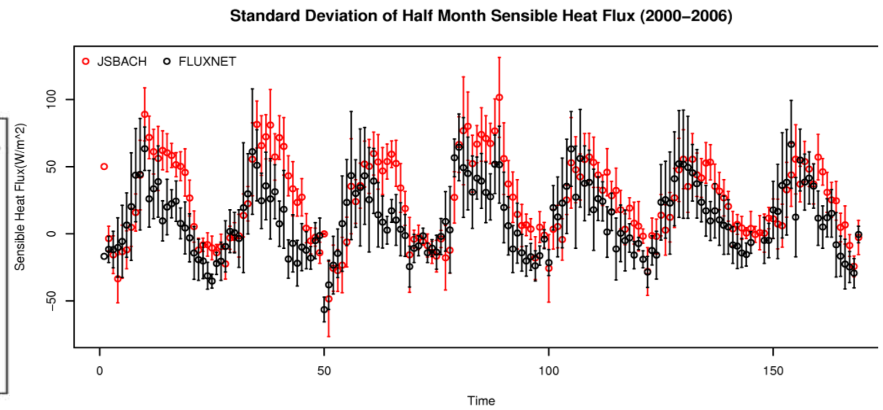
从所有站点的NetCDF文件中获取其经纬度信息

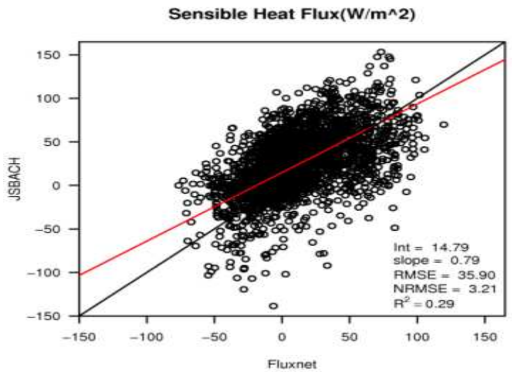
```
for (S in 1:length(SITE_NAME)) {
   site_nc <- open.ncdf(paste(site_surface_dir, "/", SITE_NAME[S]
".", NTILES[S], "_tiles.surface.nc", sep=""))
   print(SITE_NAME[S])
   lonlat[1,S] <- get.var.ncdf(site_nc, "lon")
   lonlat[2,S] <- get.var.ncdf(site_nc, "lat")[1]
}
# Plot lon&lat in world map
opar<-par( cex=0.6, pin=c(6.0, 3.0))
load("TM_WORLD_BORDERS_SIMPL-0.2.RData")
plot(wrld_simpl, axes = TRUE, xlim = c(-180,180), ylim = c(-
120, 90), asp = 1)
grid()
for (L in 1:length(LON)){
points(LON[L],LAT[L], col=rainbow(length(LON))[L],
pch=substr(SITE_NAME[L],1,1))
}
legend('bottom', SITE_NAME, pch=substr(SITE_NAME,1,1),
col=1:length(SITE_NAME), bty="n", ncol=8, cex=0.5)
```

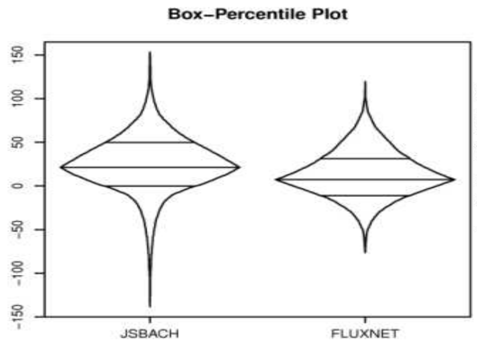
不同时间尺度的模拟性能评估

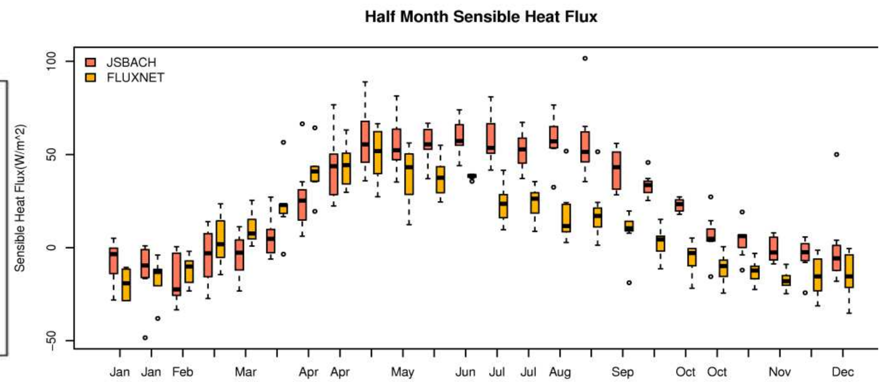
- 不同时间尺度对模型的模拟性能存在差异
- 6种时间分辨率（半小时、天、半月、月、年、每月每小时）的模拟结果与FLUXNET涡动数据进行对比分析

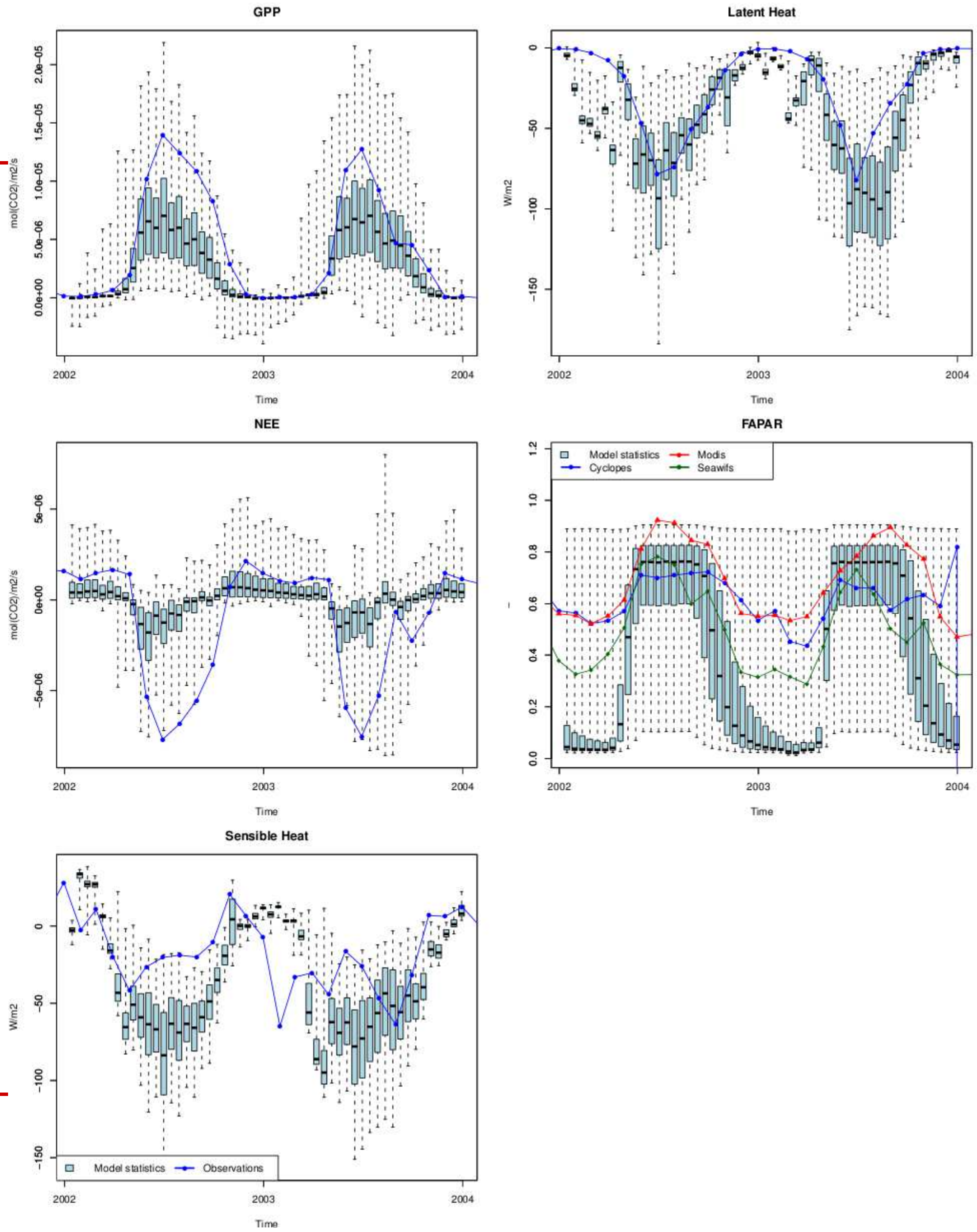










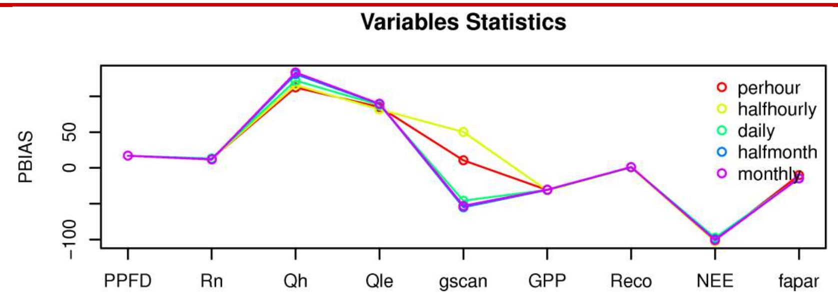

GPP
mol(CO2)/m2/s
Time
Latent Heat
W/m2
Time
NEE
mol(CO2)/m2/s
Time
FAPAR
Model statistics
Cyclopes
Modis
Seawifs
Time
Sensible Heat
W/m2
Model statistics
Observations
Time
2002
2003
2004

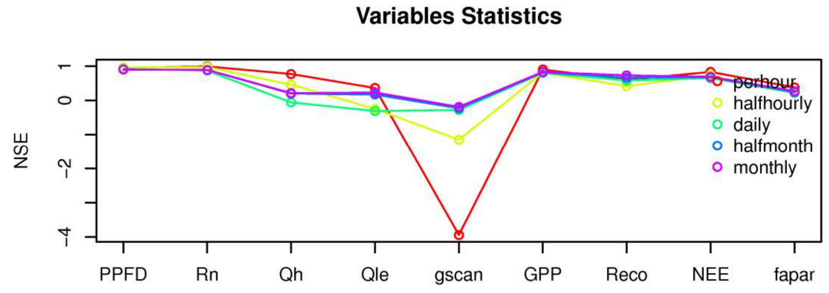
不同时间尺度的模拟性能评估

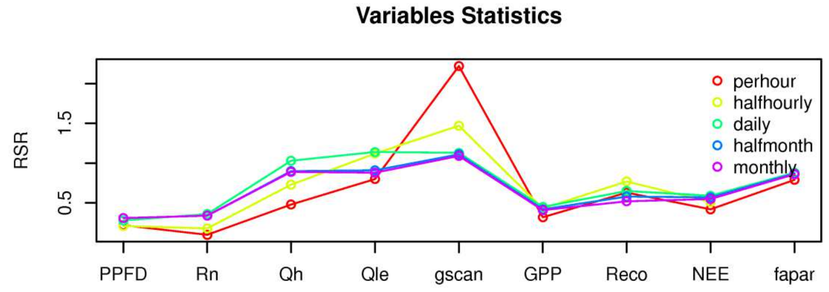
年平均对比分析。NA为空值。

参数	模拟值	观测值	单位
PPFD	267.9624313	229.1798109	$umol/m^2s$
Rn	64.01994312	56.74520001	W/m^2
Qh	22.24316819	11.7292733	W/m^2
Qle	41.03596916	21.15859903	W/m^2
Gscan	0.003159619	0.00813919	m/s
GPP	2.598816044	3.948835397	$umol/m^2s$
Reco	2.585311374	2.672223771	$umol/m^2s$
NEE	-0.01350467	-1.276611627	$umol/m^2s$
fAPAR	0.434984523	0.437642477	-
vegC	7563.983507	NA	gC/m^2
litterC	4353.142782	NA	gC/m^2
soilC	22985.14643	NA	gC/m^2

- 不同的统计分析图表对数据的统计的意义和侧重点不同
- 月和每月每小时尺度的感热模拟的性能相对较好，天时间尺度的感热的模拟性能较差

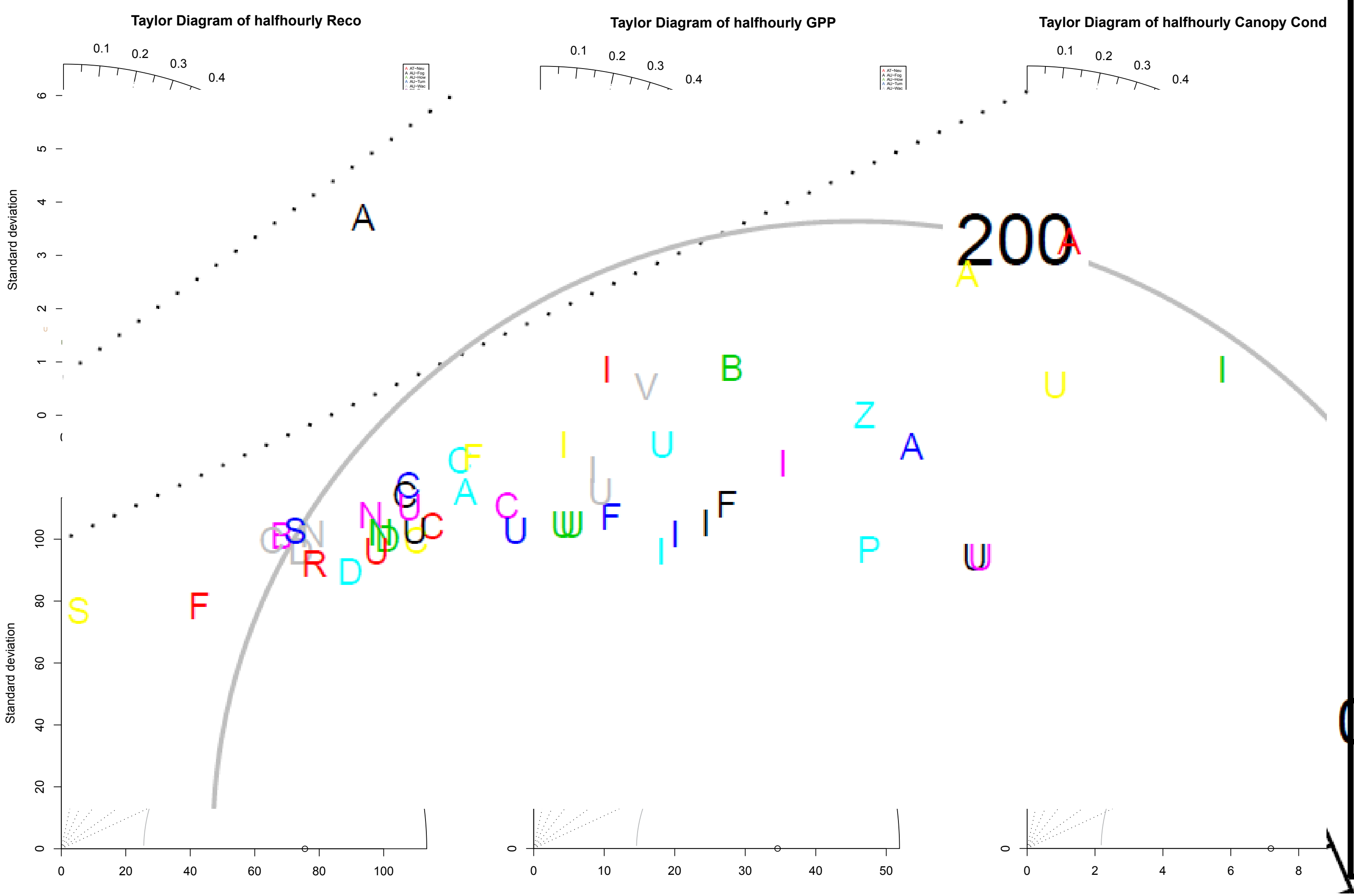


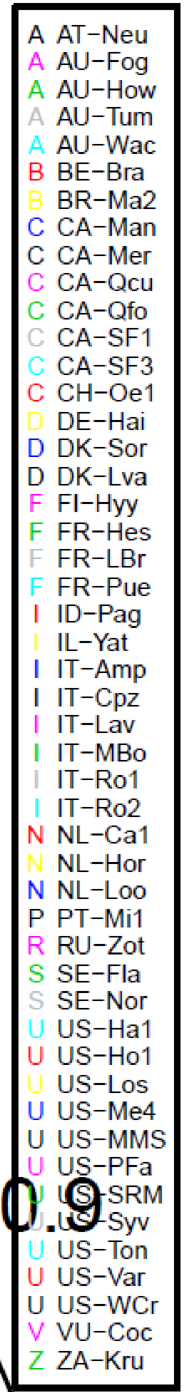




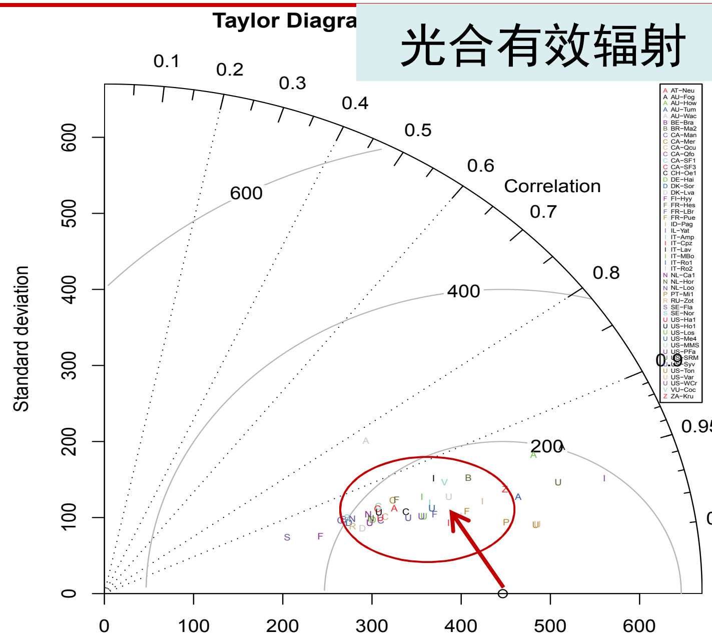
不同站点的模拟性能评估

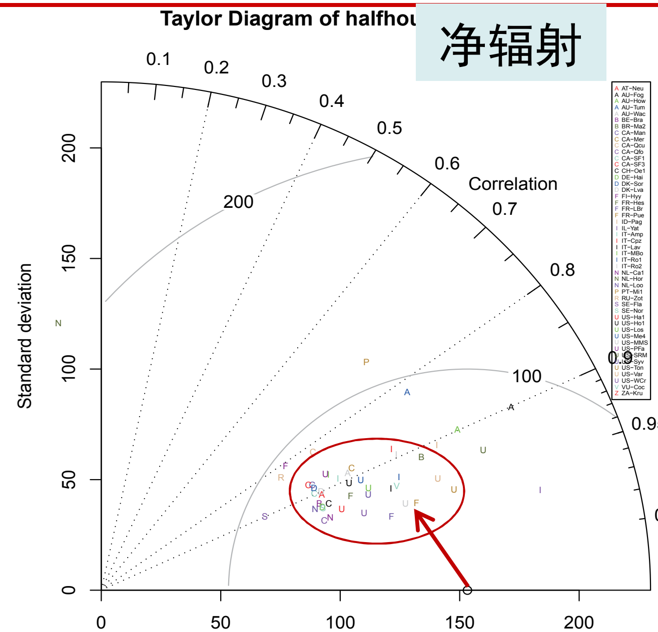
- 不同站点的模拟性能的对比分析能揭示JSBACH在不同植被功能型和气境中的模拟性能的优劣，从整体上评价JSBACH模型的适用性以及优缺





不同站点的模拟性能评估



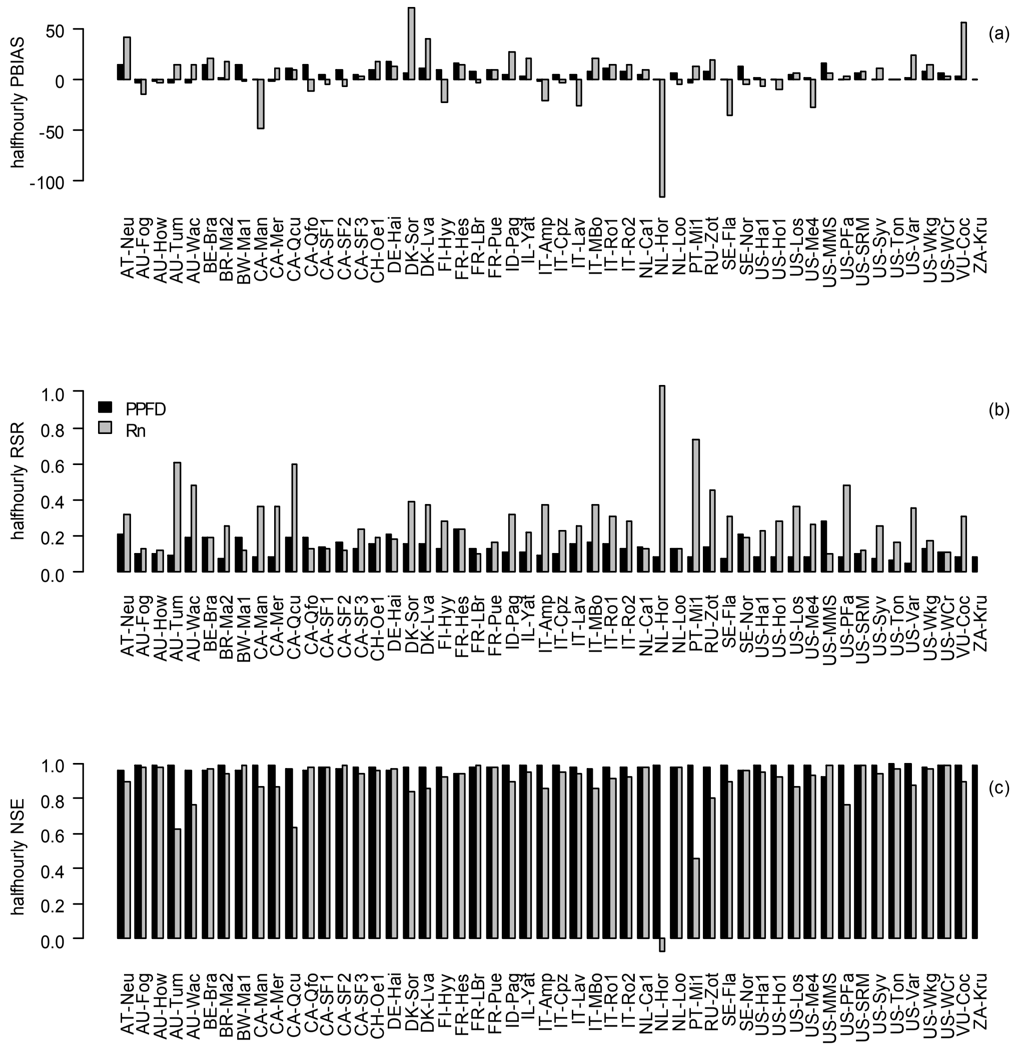


- 从这8个变量的半小时模拟尺度的泰勒图可以看出，净辐射(Rn)和光合有效辐射(PPFD)在52个FLUXNET站点的模拟性能最好
- 这两个变量在各个站点的统计值离“参考点”的距离最近，且离散的程度较小
- AT-Neu、AU-Tum、CA-Man、Dk-Sor、DK-Lva、NL-Hor、SE-Fla、US-Pfa和ZA-Kru等FLUXNET站点模拟的这8个变量的性能相对较差
- 说明JSBACH在不同的植被功能型和气候环境下的适应性较好

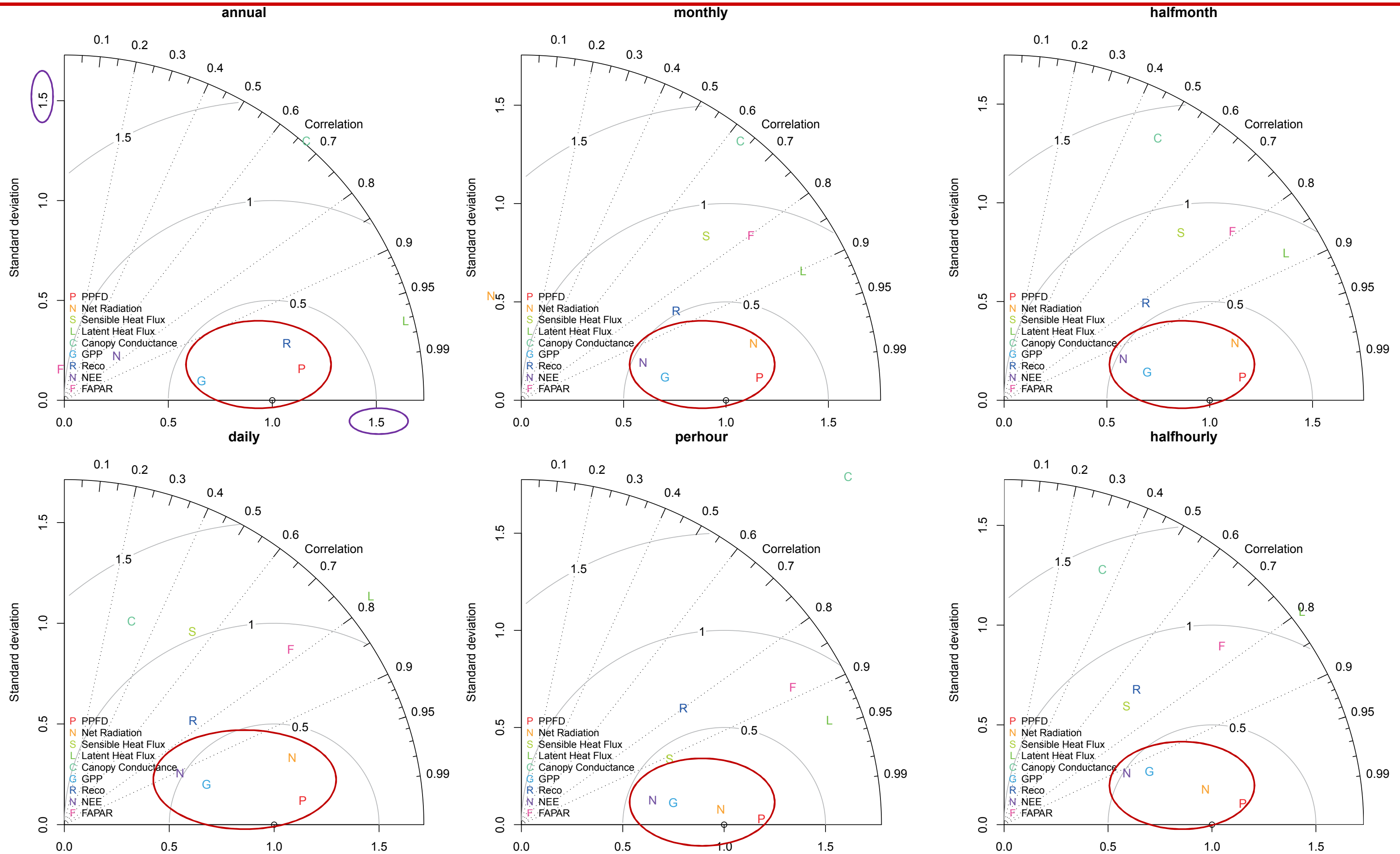
不同站点的模拟性能评估

光合有效辐射（PPFD）和净辐射（净辐射）在大多数站点：

- PBIAS的绝对值<20%
- RSR值<0.25
- NSE值接近1

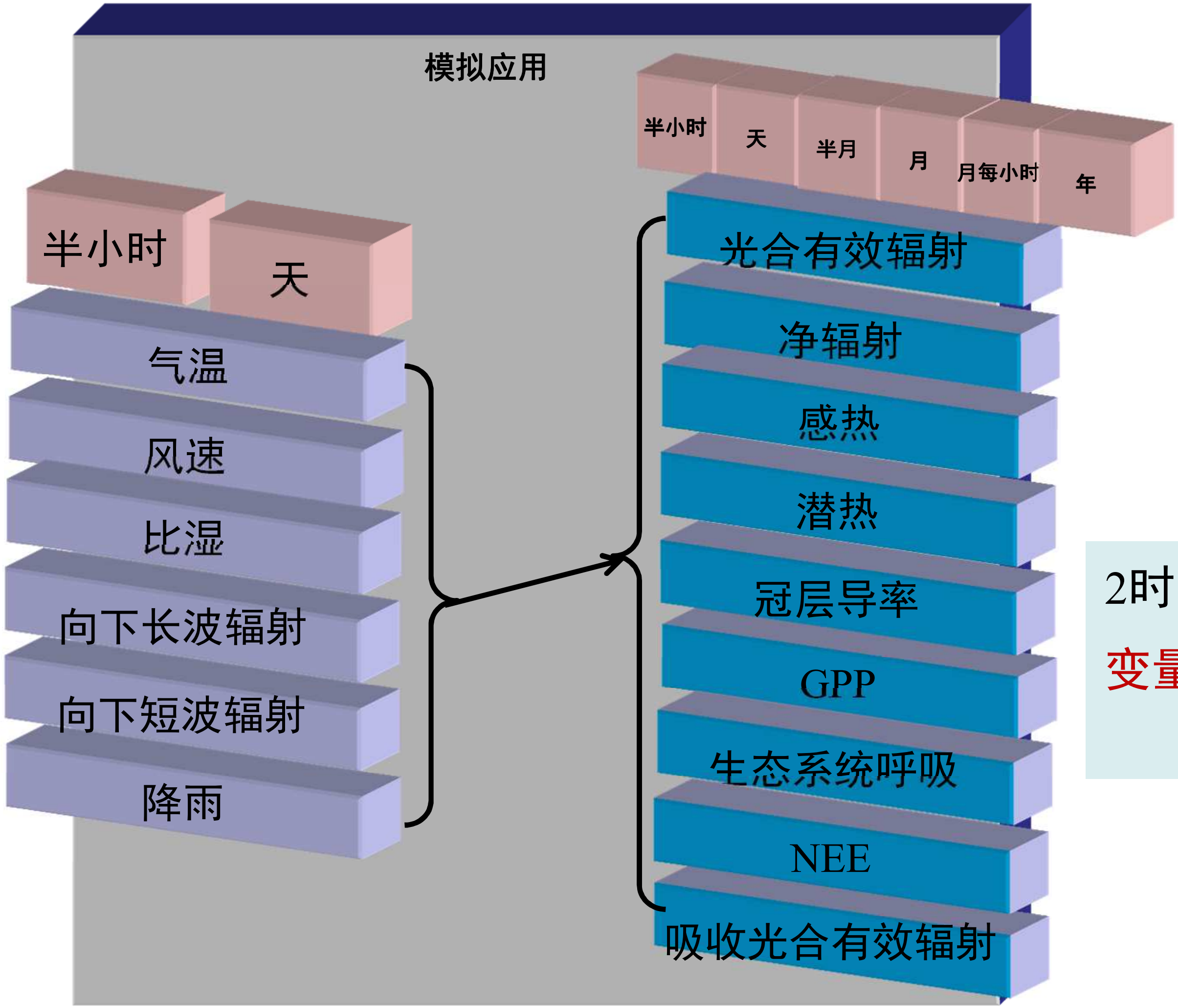


不同变量在不同时间尺度的模拟性能评估



- 变量的单位不一样，需要在统一的标准下进行处理分析（标准偏差处理）
- DE-Hai站点为例，半小时、天、半月、月、月每小时（年数据太少）NEE、GPP、Rn和PPFD这四种变量在DE-Hai站点的不同时间尺度中模拟性能较好

不同时间尺度强迫数据对模拟性能评估

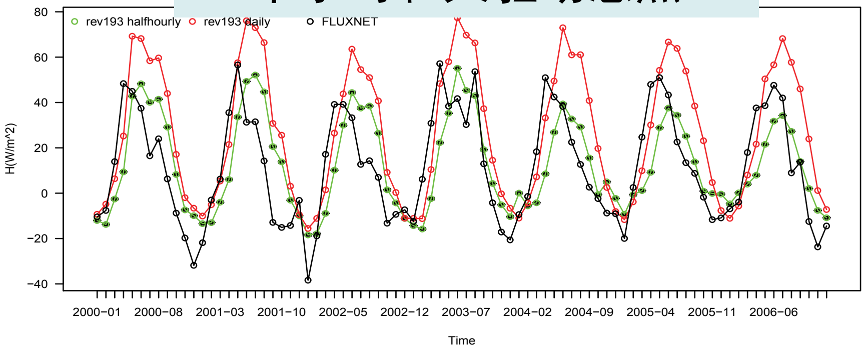


2时间*6强迫->6时间*9变量

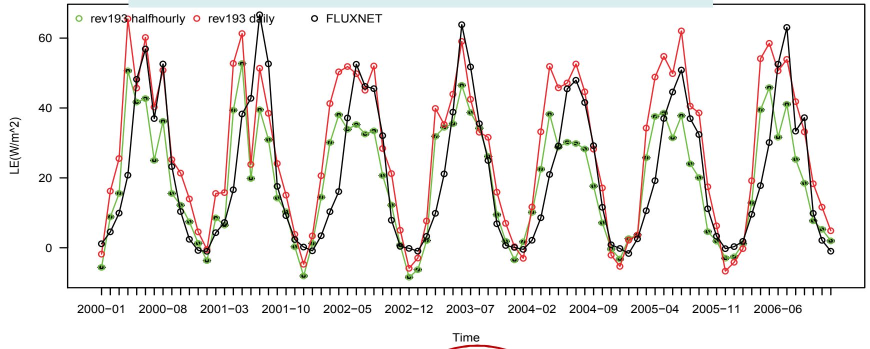
变量之间的时间尺度不一样

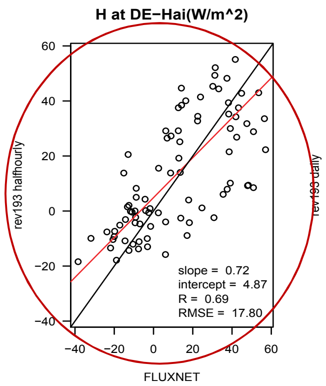
不同时间尺度强迫数据对模拟性能评估

半小时和天驱动感热

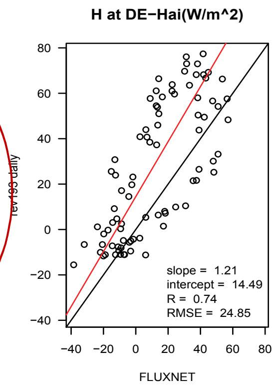


半小时和天驱动潜热

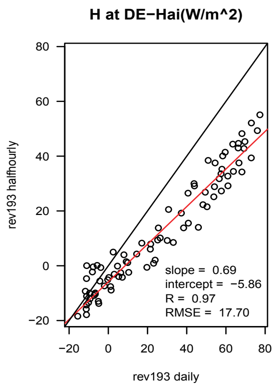




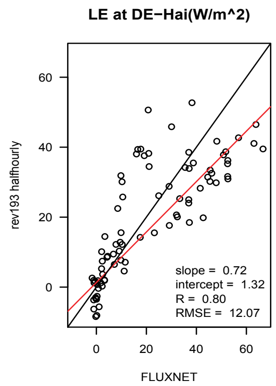
半小时

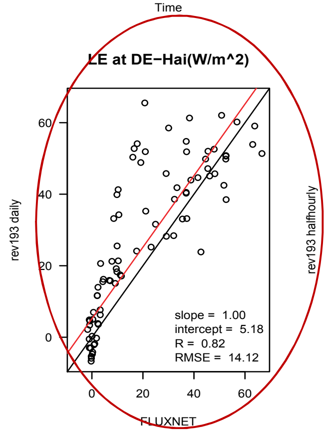


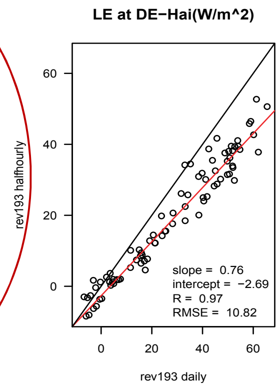
天



相互





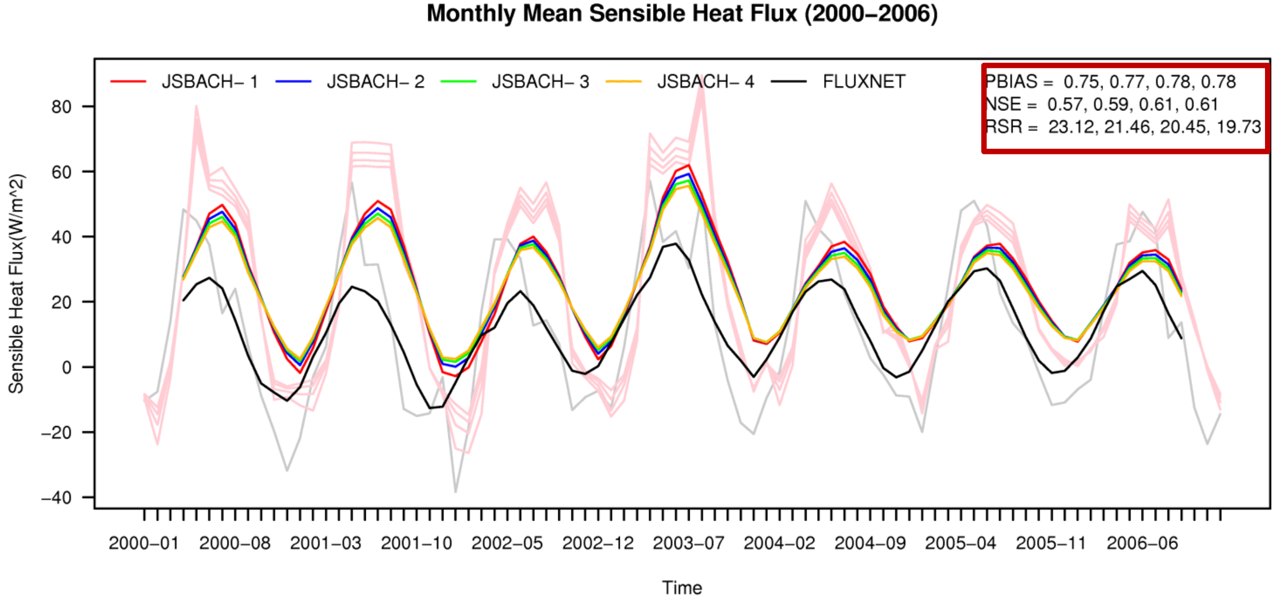


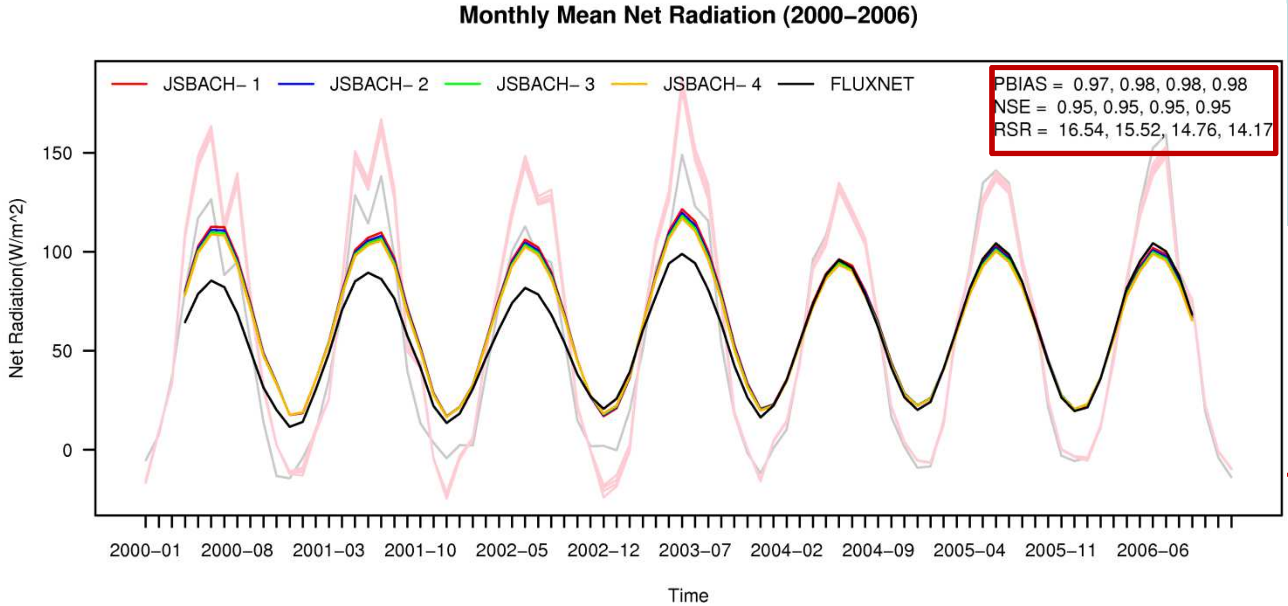
不同时间尺度强迫数据对模拟性能评估

参数	输入时间尺度	Intercept	Slope	r	R^2	RMSE	SD_JSB	SD_FLX	NAE	VR	PBIAS	NSE	RSR
PPFD	HalfHourly	-7	0.91	1	0.99	36.14	163.01	179.31	-0.12	0.83	12.12	0.96	0.2
	Temp_HalfHourly	0.49	1.19	0.99	0.99	56.97	206.7	179.31	0.2	1.33	-19.63	0.9	0.31
	Wind_HalfHourly	-8.56	0.89	1	0.99	42.55	159.33	179.31	-0.15	0.79	14.93	0.95	0.23
	Humi_HalfHourly	-8.56	0.89	1	0.99	42.55	159.33	179.31	-0.15	0.79	14.93	0.95	0.23
	Long_HalfHourly	0.49	1.19	0.99	0.99	56.97	206.7	179.31	0.2	1.33	-19.63	0.9	0.31
	Shor_HalfHourly	-0.85	1.17	0.99	0.99	50.63	204.85	179.33	0.17	1.3	16.92	0.92	0.28
	Prec_HalfHourly	0.49	1.19	0.99	0.99	56.97	206.7	179.31	0.2	1.33	-19.63	0.9	0.31
	Daily	0.49	1.19	0.99	0.99	56.97	206.7	179.31	0.2	1.33	-19.63	0.9	0.31
Rn	HalfHourly	-9.62	0.77	NA	0.88	32.87	45.79	59.89	-0.42	0.58	41.51	0.74	0.51
	Temp_HalfHourly	-7.52	1.12	NA	0.88	21.46	61.78	59.89	-0.01	1.06	1.08	0.87	0.36
	Wind_HalfHourly	-16.09	0.86	NA	0.86	33.85	48.17	59.89	-0.43	0.65	43.24	0.73	0.52
	Humi_HalfHourly	-233.4	0.62	NA	0.65	256.4	42.08	59.89	-4.48	0.49	448.35	0.03	0.98
	Long_HalfHourly	0.09	0.99	NA	0.89	19.72	57.87	59.89	-0.03	0.93	2.6	0.89	0.33
	Shor_HalfHourly	-8.25	1.32	0.95	0.91	28.32	75.44	59.9	0.2	1.58	19.52	0.78	0.46
	Prec_HalfHourly	-7.58	1.11	NA	0.88	21.58	61.46	59.89	-0.02	1.05	2.41	0.87	0.36
	Daily	-7.65	1.12	NA	0.88	21.73	61.67	59.89	-0.02	1.06	1.79	0.87	0.36

- 表格的自动生成采用了两种方式：①SHELL的正则表达式；②R脚本
- 整体上其半小时强迫数据集的模拟性能要稍好于其它时间尺度的模拟性能
- 但针对对潜热、生态系统呼吸等天尺度的模拟性能略好于半小时的模拟性能
- 低分辨率的强迫数据集和不同分辨率混合的强迫数据集（更符合野外观测实际）为陆面过程模型提供了一个数据的替代方案。

不同模型版本的评估

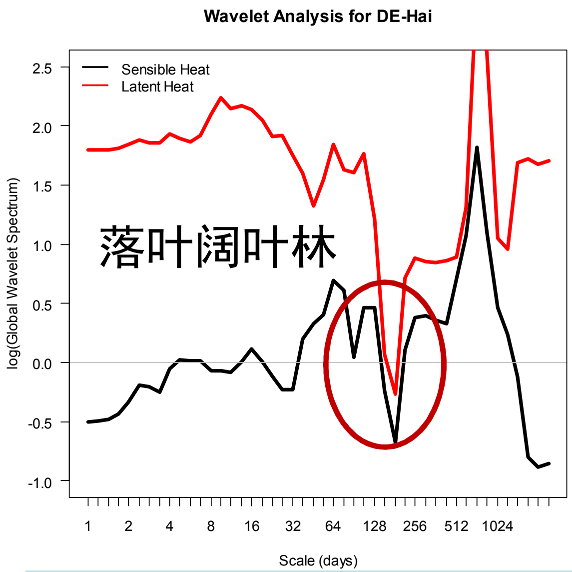


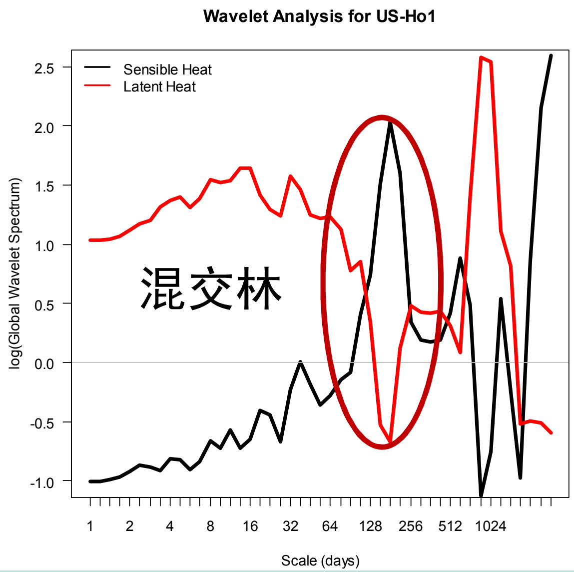


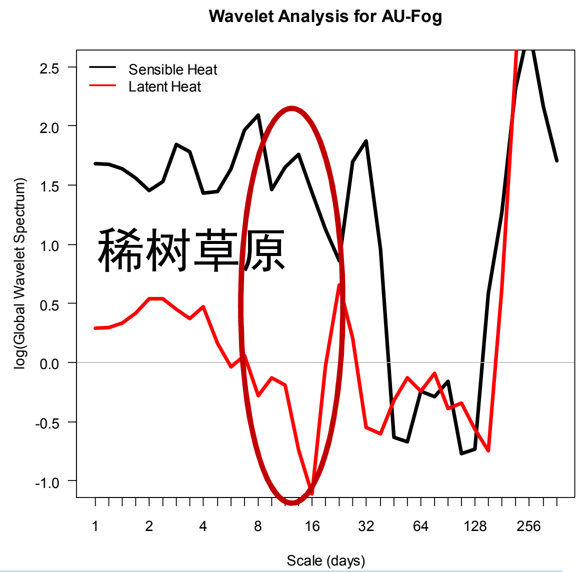
模型的开发过程需要对不同版本模型的模拟性能进行评价，以确定模型的模拟性能随着版本的升级是否得到改善

- 感热随着版本的升级，其峰值和波谷值更收敛
- 净辐射峰值在变小
- 而且潜热和GPP的模拟性能也有所改善
- 其它变量差别不大

小波分析对不同时间尺度的刻画



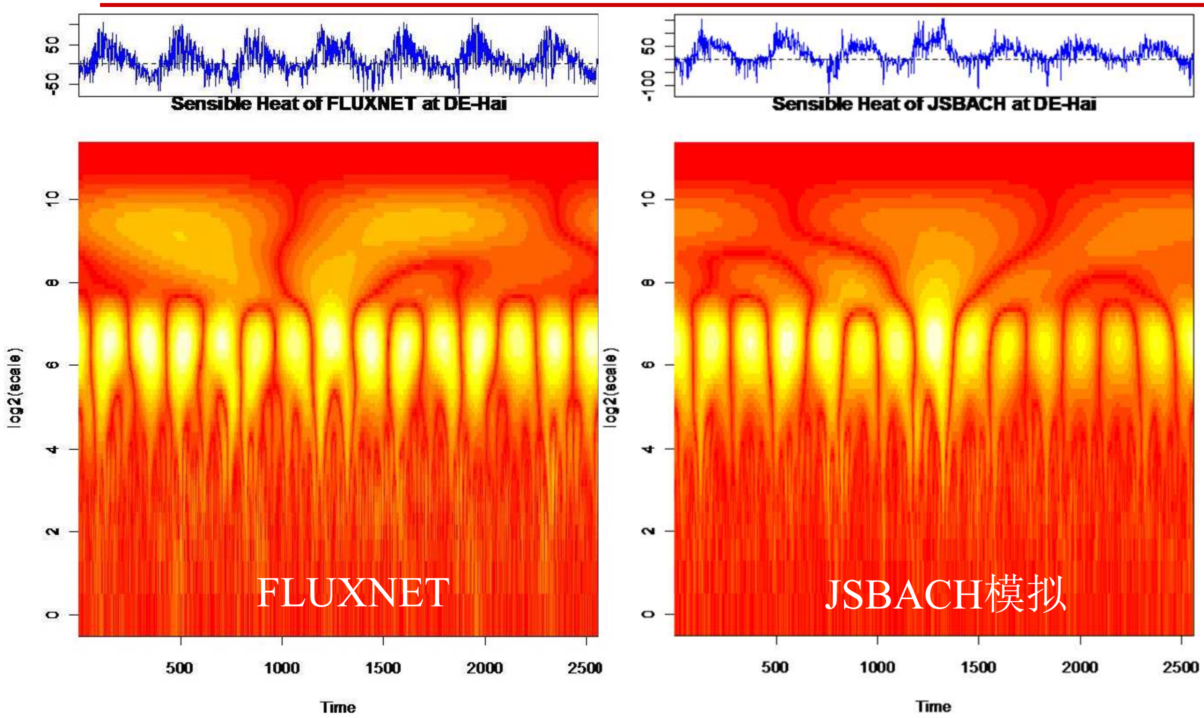




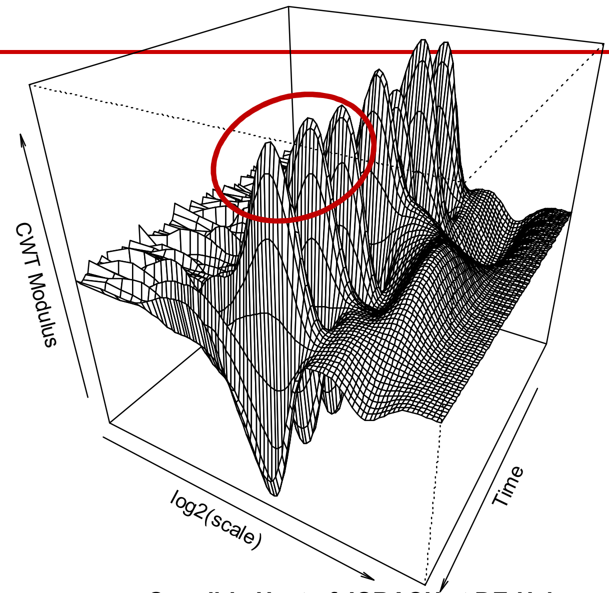
采用了小波全谱GWS转换，时间尺度设置为0.25天，在95%的置信区间，结果用log函数处理以降低GWS的维度

- DE-Hai站点的潜热模拟整体被低估，感热从天到周尺度被低估，其月到半年尺度被高估
- US-Ho1站点的感热和潜热在相同的时间尺度呈现相反的特征，从天到半年尺度感热的模拟被低估，而从半年到年尺度，感热的模拟却显示被高估
- AU-Fog站点，从天到月尺度，感热模拟被高估，从月到半年尺度，其感热和潜热的模拟皆被低估了，潜热从天到周尺度被高估

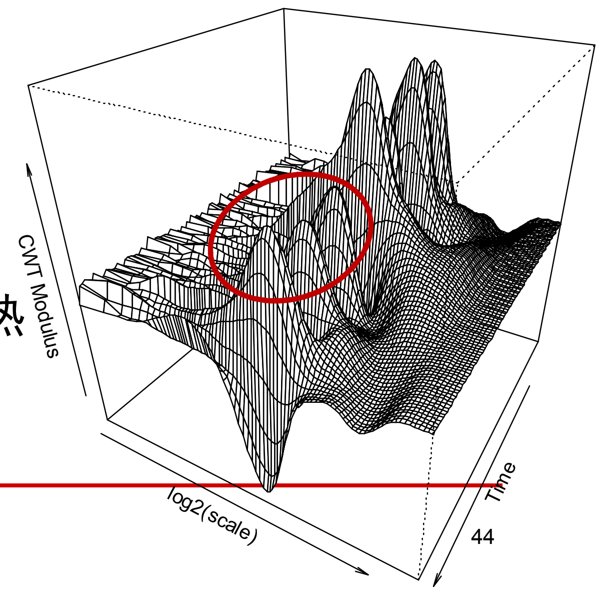
小波分析对不同时间尺度的刻画



Sensible Heat of FLUXNET at DE-Hai



Sensible Heat of JSBACH at DE-Hai



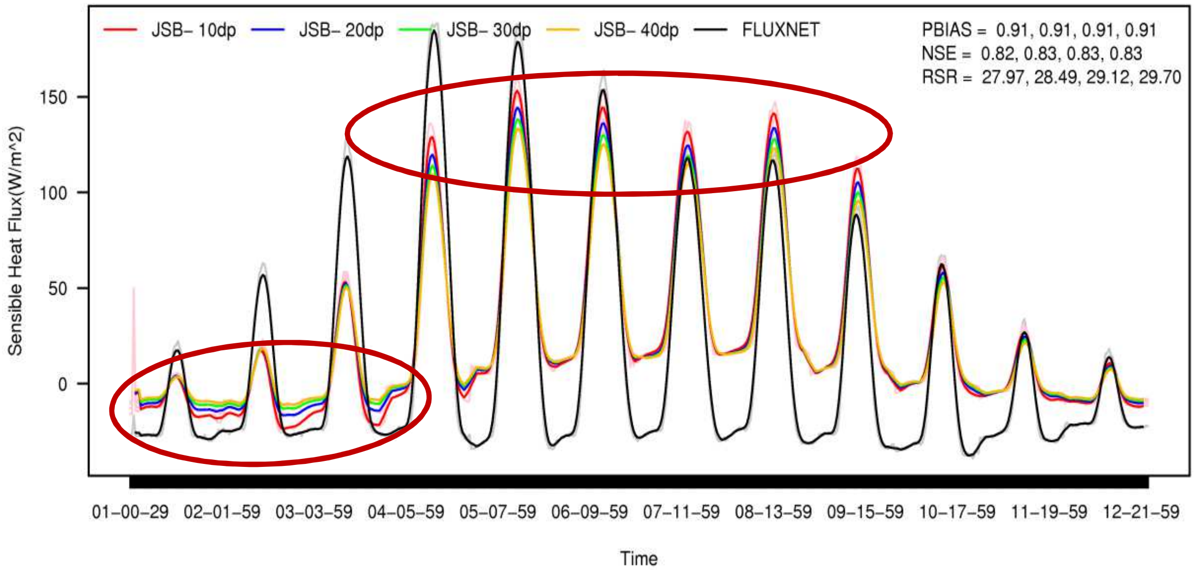
感热连续小波变换

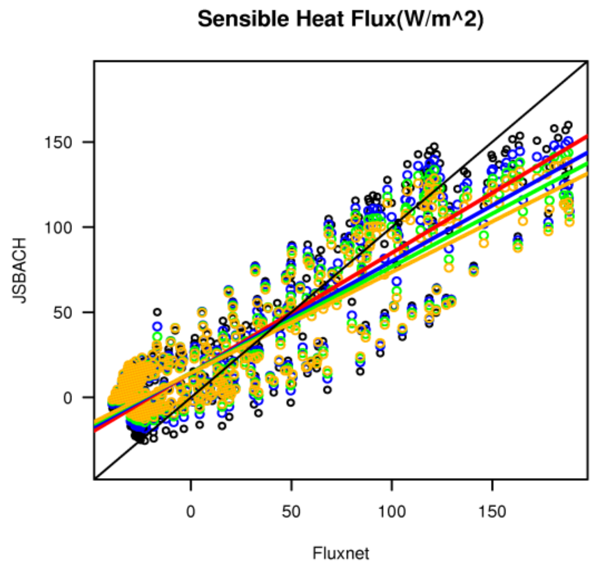
- 模拟小波系数刚开始变化不大，但是随后变化却明显减少
- 而FLUXNET的小波系数的变化却一直都不明显，说明感热模拟的长期变化的稳定程度较差

植被高度对感热和潜热通量的响应

- 地球表面覆被直接影响着地气系统物质、能量的输送和局地环境气候，因此计算和分析植被高度对湍流通量的影响是研究陆面过程的关键
- 不同植被高度下的感热和潜热的敏感性分析，改变植被高度将影响陆面与大气间的湍流交换，设置DE-Hai站点的地表植被高度分别为10、20、30和40cm

每月每小时感热

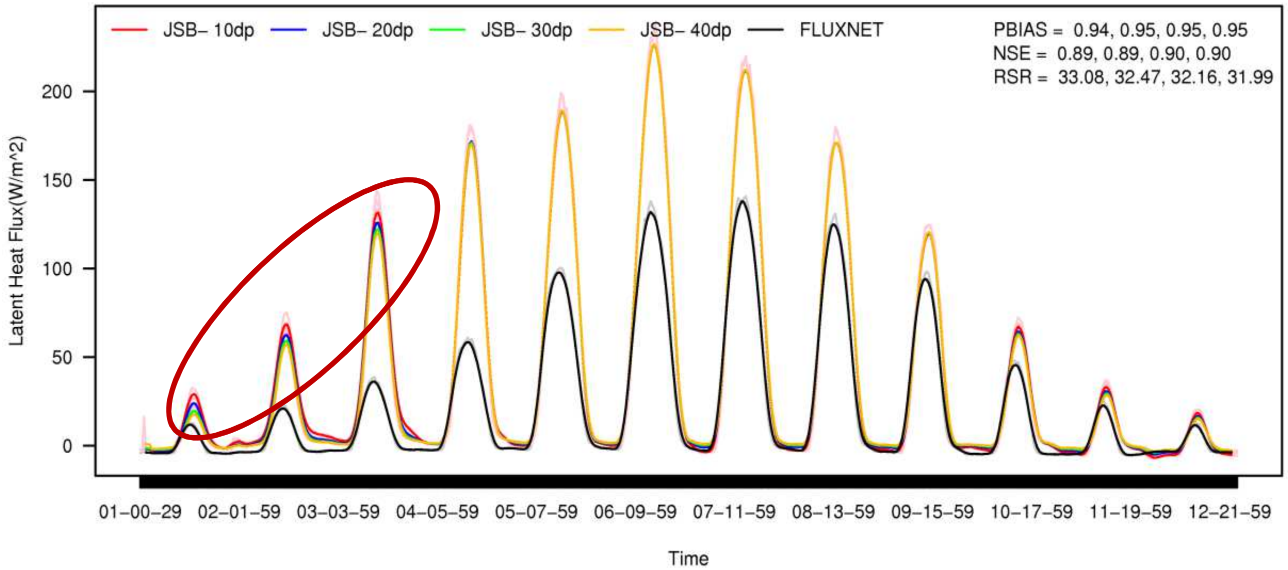


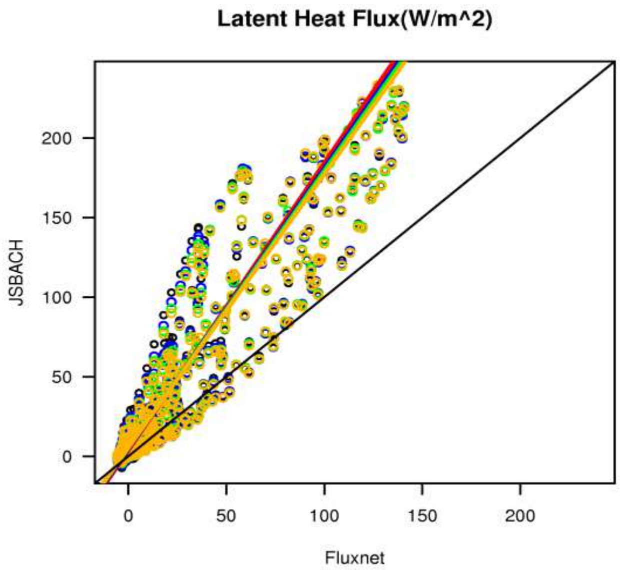


- 感热在1-3月其波谷值差距较大，在4-9月模拟感热通量峰值差距明显
- 其模拟的值与观测值的差距随着植被高度的升高其差距增大
- 差异的最大值出现在4-9月的10点到14点

植被高度对感热和潜热通量的响应

每月每小时潜热



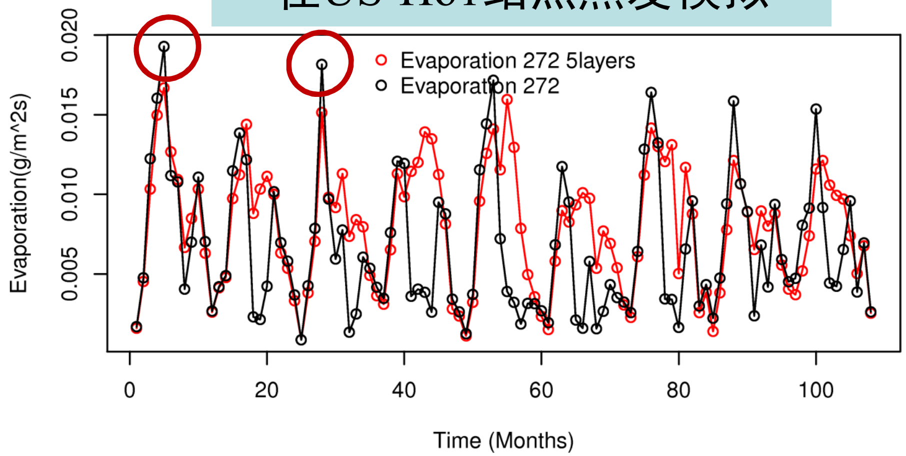


- 潜热在1-3月其峰值值差距略大
- 潜热的变化随着高度的升高并不显著

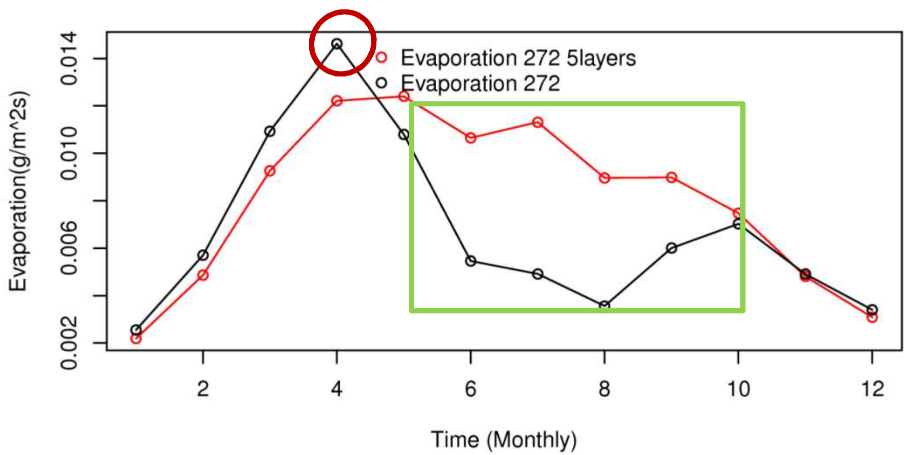
水箱模型与五层土壤模型的对比分析

- JSBACH早期版本采用水箱模型，从272版本开始，采用五层土壤模型
- 水箱模型在JSBACH模型里并没有描述在每一个网格单元的不同植被有不同的根长，而且蒸腾和蒸发的水分有效性并没有被适当的反馈
- 五层土壤模型的土壤分层分别为3cm、19cm、78cm、268cm和698cm

在US-Ho1站点蒸发模拟



US-Ho1 (1996-2004)



- 五层土壤模型对蒸发的模拟从4-10月比水箱模型下降的更为平缓，水箱模型从4-8月显著下降，而从8-10月又呈现显著的上升
- 水箱模型的最大蒸发量要大于五层土壤模型的蒸发量，但五层土壤模型的整体蒸发量要大于水箱模型的蒸发量

R语言程序结构

R脚本输入参数

```
# Set parameters
Args <- commandArgs(trailingOnly=TRUE);
if(length(Args) != 2) {
  message("generate_FLUXNET_diagnostics.R requires
  startyear endyear as input. Terminating");quit()
}
```

R图形函数

```
# Graphic Function (1) FLUXNET STUFF---------------
JSBACH_FLUXNET_GRAPHIC <- function(JSBACH_V, FLUXNET_V,
  TIME, mean, gtype, labs, variable_info,
   RMSE, NRMSE, RR, NSE, PBIAS, RSR){
opar<-par( cex=0.6, pin=c(6.0, 2.0))
split.screen(c(2,1))
split.screen(c(1,2),screen=2)
screen(1)
  plot(TIME, JSBACH_V*NA, type='n',  col="red",
    ylim=c(min(JSBACH_V, FLUXNET_V,na.rm=TRUE),
  max(JSBACH_V, FLUXNET_V,na.rm=TRUE)),
    xlab="Time", ylab=paste(variable_info[1],"(",
  variable_info[2], ")", sep=""),
```

```
main=paste(typestr, " ", variable_info[1], " (",
  variable_info[3],"-",variable_info[4], ")", sep=""),
    las=1, axes=F)
  axis(1, at=1:length(TIME), lab=labs,las=1)
  axis(2,las=1)
  box()
  lines(TIME, FLUXNET_V, type=gtype, col="black")
  lines(TIME, JSBACH_V, type=gtype, col="red" ,pch=19)
  legend("topleft", legend = c("JSBACH","FLUXNET"),
  col=c("red","black"), pch=c(19,21), bty="n", ncol=2)
  legend("topright", legend = c(paste("NSE = ", NSE),
  paste("PBIAS = ", PBIAS, "%", sep=""), paste("RSR = ",
  RSR)), bty="n")
```

```
screen(3)
  if(!all(is.na(FLUXNET_V))) {
    plot(FLUXNET_V,JSBACH_V,
xlab='Fluxnet',ylab='JSBACH',las=1,
          ylim=c(min(JSBACH_V, FLUXNET_V,na.rm=TRUE),
max(JSBACH_V, FLUXNET_V,na.rm=TRUE)),
          xlim=c(min(JSBACH_V, FLUXNET_V,na.rm=TRUE),
max(JSBACH_V, FLUXNET_V,na.rm=TRUE)),
          main= paste(variable_info[1], "(" ,
variable_info[2], ")", sep= "" ))
    line_FJ <- line(FLUXNET_V,JSBACH_V)
    a=NA;b=NA
    if(!any(is.na(coef(line_FJ)))) {
      a=sprintf("%.2f",coef(line_FJ,use.na=false)[1])
      b=sprintf("%.2f",coef(line_FJ,use.na=false)[2])
      abline(line_FJ, col="red")
    }
    abline(0,1)
    legend("bottomright", legend = c(paste("Int = ", a),
paste("slope = ", b), paste("RMSE = ", RMSE),
      paste("NRMSE = ", NRMSE), as.expression(bquote(R^2
== .(RR)))), bty="n")
  }
```

```
screen(4)
  #JF_T <-
c(rep("JSBACH",length(TIME)),rep("FLUXNET",length(TIME)))
  #JF_V <- c(JSBACH_V, FLUXNET_V)
  #bwplot(JF_T ~ JF_V, panel=panel.bpplot, nout=.05,
scat1d.opts=list(frac=.01))
  #bpplot(JSBACH_V, FLUXNET_V, name=c("JSBACH", "FLUXNET"))
  boxplot(JSBACH_V, FLUXNET_V, names=c("JSBACH",
"FLUXNET"),las=1,
    ylab=paste(variable_info[1],"(", variable_info[2], ")",
sep=""),
    main = paste("Boxplot of ", variable_info[1]))

close.screen(1:4)
}
```

Graphic Function (2) SATELLITE STUFF
Graphic Function (3) compare Forcing data with simulation and evaluation
Graphic Function (4) Error Bars
Graphic Function (5) Variables Statistics

不同时间分辨率的转换

```
# Aggregation by time routine ------------------------------------
---------------------------------------
JSBACH_FLUXNET_COMPUTE_GROUPING <- function(Value, TIME, period,
method){

  if (period=="perhour"){
    time <- format(as.POSIXct(as.Date(as.numeric(TIME-1/24),
origin=starttime)),"%m-%H-%M")
  }

  if (period=="perhour_15"){
    time <- format(as.POSIXct(as.Date(as.numeric(TIME-1/24),
origin=starttime)),"%m-%d-%H-%M")
  }

  if (period=="daily"){
    time <- format(as.Date(as.numeric(TIME),origin=starttime),
"%Y-%m-%d")
  }
```

```
if (period=="halfmonth"){
  time <- format(as.Date(as.numeric(TIME),origin=starttime),
"%Y-%m-%d")
  }

  if (period=="monthly"){
    time <- format(as.Date(as.numeric(TIME),origin=starttime),
"%Y-%m")
  }

   。 。 。

time <- paste(substr(time,1,8),
ifelse(as.integer(substr(time,9,10)) %in% c(1:15), 15,
30),sep="")
    JF_V <- data.frame(time, Value=JF_AGG$x)
    JF_AGG <- aggregate(JF_V$Value, list(JF_V$time), method,
na.rm=TRUE)
  }
  return (JF_AGG)
}
```

定义需要对比分析的变量

```
# define required compared variables, can be added later.
nvar=9
aggregation_level = c( "perhour", "daily", "monthly", "annual" )
stats_time <- c("halfhourly", "daily", "monthly")
annual_values=array(NA,c(3,nvar+3))
eval_stats=array(NA, dim=c(14,nvar,4))
variable_names <- array(c("par_acc", "PPFD", "PPFD", "PPFD",
                          "net_radiation", "Rn", "Net Radiation",
"Rn",
                          "sensible_heat_flx", "H", "Sensible
Heat Flux", "Qh" ,
                          "latent_heat_flx", "LE", "Latent Heat
Flux", "Qle",
                          "canopy_cond_limited", "gsurf",
"Canopy Conductance", "gscan" ,
                          "net_assimilation", "GPP", "GPP",
"GPP",
                          "reco", "Reco", "Reco", "Reco",
                          "zCO2_flux_net", "NEE", "NEE", "NEE",
                          "fapar", "fapar", "FAPAR", "fapar" )
                        , dim=c(4,nvar))
```

```
variable_units <- array(c("mol/m^2s", "umol/m^2s",
"umol/m^2s",
                               "W/m^2", "W/m^2", "W/m^2",
                               "W/m^2", "W/m^2", "W/m^2",
                               "W/m^2", "W/m^2", "W/m^2",
                               "m/s","mmol m-2 s-1","m/s",
                               "mol/m^2s", "umol/m^2s",
"umol/m^2s",
                               "mol/m^2s", "umol/m^2s",
"umol/m^2s",
                               "mol/m^2s", "umol/m^2s",
"umol/m^2s",
                               "-","-","-")
                             , dim=c(3,nvar))
variable_units_conversion <- array(c(10^6,1,
                                          1,1,
                                          -1,1,
                                          -1,1,
                                           1,18*10^-6,
                                          10^6,1,
                                          10^6,1,
                                          10^6,1,
                                     1,1)
                             , dim=c(2,nvar))
```

调用

```
if(length(TIME) == length(TIME_F)) {
  for (V in 1:nvar){
     JSBACH_V <- get.var.ncdf( modelfile, variable_names[1,V] )
     JSBACH_V <- JSBACH_V*(variable_units_conversion[1,V])
     JSBACH_V[JSBACH_V<=-9999.]=NA
     if(variable_names[1,V]=='fapar') {
       FLUXNET_V <- get.var.ncdf( datafile, 'fapar_seawifs' )
       FLUXNET_V[FLUXNET_V==-9999.]=NA
       FLUXNET_V_FLAG <- get.var.ncdf( datafile,
'fapar_seawifs_flag' )
       FLUXNET_V_FILTERED=FLUXNET_V
       FLUXNET_V_FILTERED[FLUXNET_V_FLAG>0.5]=NA
     } else {
       FLUXNET_V <- get.var.ncdf( datafile, variable_names[2,V] )
       FLUXNET_V[FLUXNET_V==-9999.]=NA
       FLUXNET_V <- FLUXNET_V*(variable_units_conversion[2,V])
       FLUXNET_V_FILTERED=FLUXNET_V
       if(any(!is.na(FLUXNET_V_FILTERED))) {
         dummy=JSBACH_V
         dummy[is.na(FLUXNET_V_FILTERED)]==NA
```

```
eval_stats[1,V,1]=sprintf("%.2f",coef(line(FLUXNET_V_FILTERED[!i
s.na(FLUXNET_V_FILTERED)], JSBACH_V[!is.na(FLUXNET_V_FILTERED)]),
                use.na=false)[1])
        }
      }

      si=1
      for (time in aggregation_level ){

        if ( time != "annual" ) {

          variable_info <-
c(variable_names[3,V],variable_units[3,V],startyear,endyear)
          JSBACH_AGG <-
JSBACH_FLUXNET_COMPUTE_GROUPING(JSBACH_V,TIME, time,'mean')
          FLUXNET_AGG <-
JSBACH_FLUXNET_COMPUTE_GROUPING(FLUXNET_V,TIME, time,'mean')

          RMSE <- JSBACH_FLUXNET_EVALUATION("rmse", JSBACH_AGG$x,
FLUXNET_AGG$x)

          if (time == "daily" || time == "halfmonth" || time ==
"monthly") {
```

```
eval_stats[1,V,si]=sprintf("%.2f",coef(line(FLUXNET_AGG$x,
JSBACH_AGG$x),use.na=false)[1])
```

```
        }

      if (time=="perhour") {
              typestr <- c("Per Hour Monthly")
              gtype <- c("b")
             labs <-
month.abb[as.numeric(gsub("0","",substr(JSBACH_AGG$Group.1, 1,
2)))]
        }

    JSBACH_FLUXNET_GRAPHIC(JSBACH_AGG$x, FLUXNET_AGG$x,
JSBACH_AGG$Group.1,
         typestr, gtype, labs, variable_info, RMSE, NRMSE, RR,
NSE, PBIAS, RSR)

      } else {
           annual_values[1,V]=mean(JSBACH_V,na.rm=TRUE)
          annual_values[2,V]=mean(FLUXNET_V,na.rm=TRUE)
      }
    }
  }
```

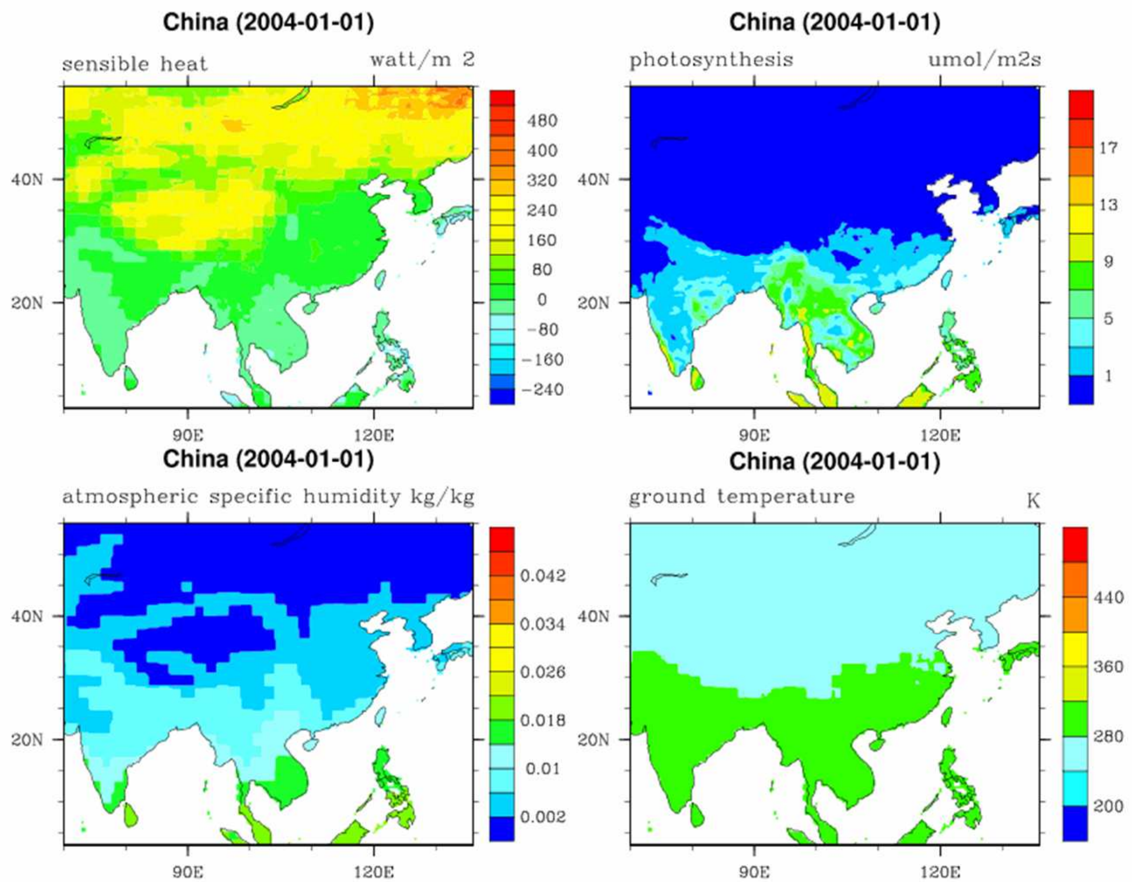
其它脚本语言集成

NCL功能函数

- 经纬度。经纬度与网格点的转换；
- 时间序列分析。时间序列数据的累计计算，包括求不同时间尺度的平均值、和、偏差等；
- 通量计算。针对净辐射、感热通量、潜热通量和地热通量，制定了其区域的2D的可视化功能；
- 值域函数。主要是针对几个主要的变量定制其值域范围；
- 数据读写。数据的读取和写入功能；
- 通用的图形函数。针对CLM的模拟定制了NCL图形绘画中的基本图形的设置。

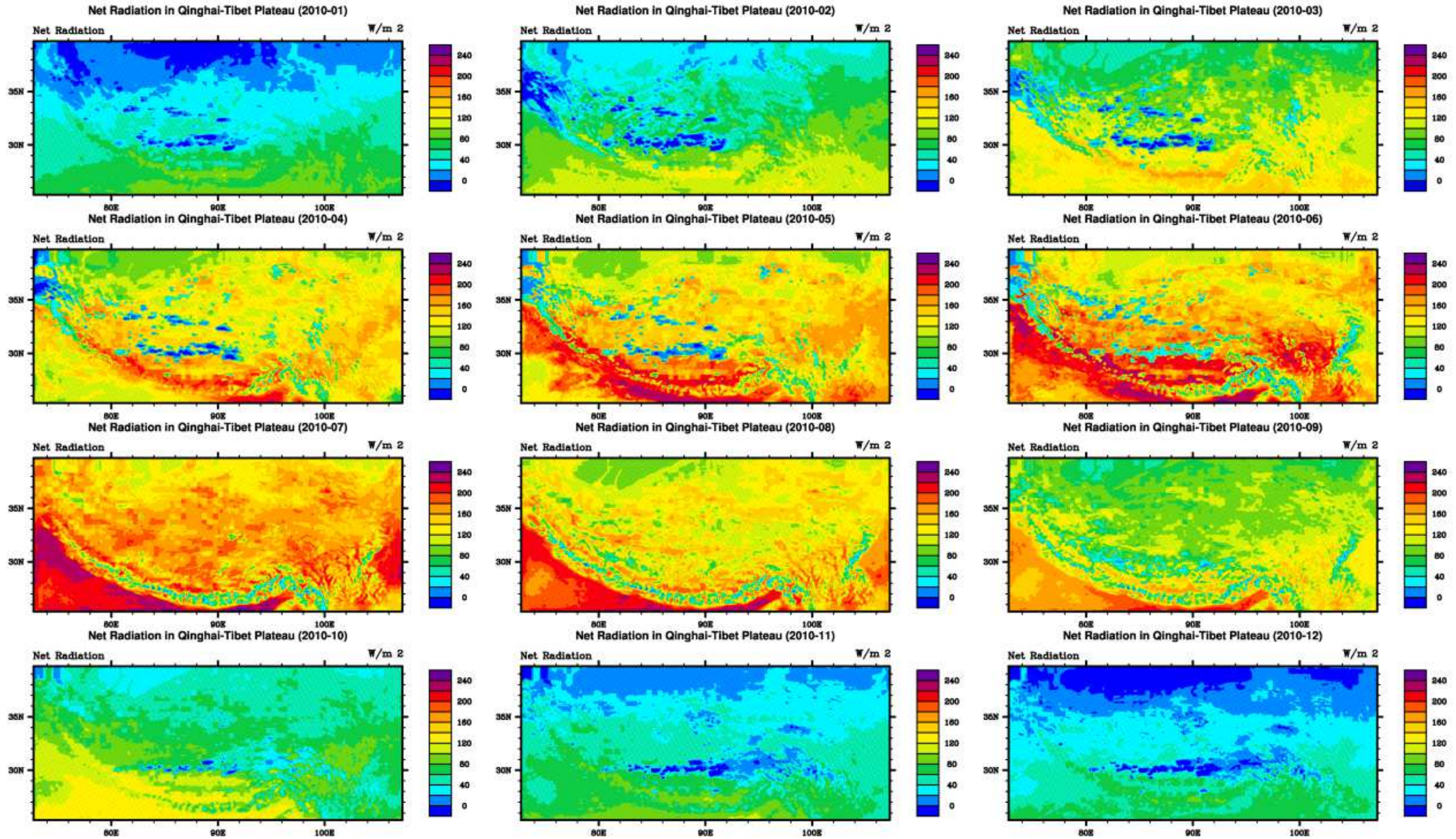
R更多的是针对功能的，而NCL更多的是针对变量的
R针对站点，而NCL针对区域的

NCL脚本语言

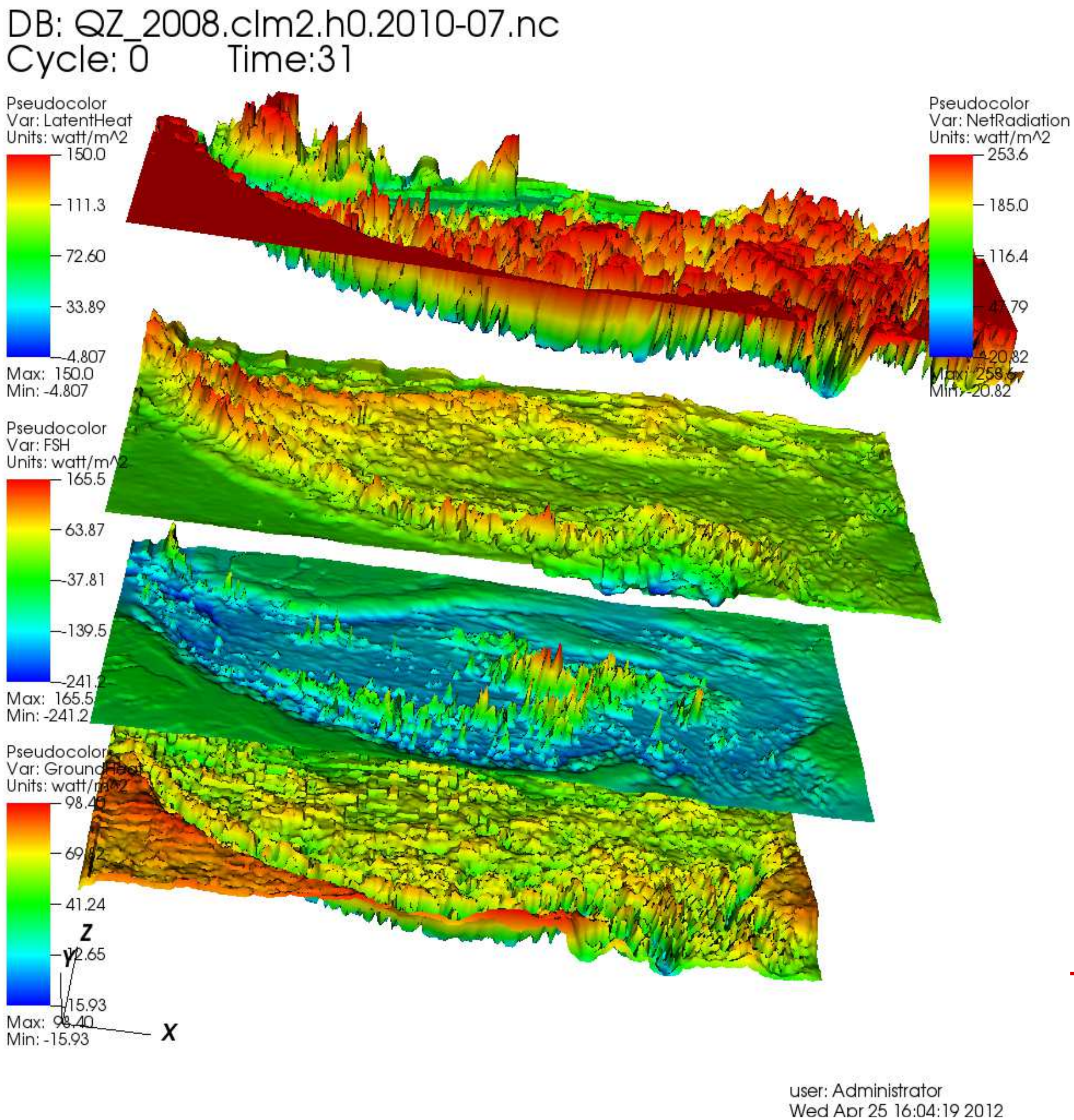


China (2004-01-01)
sensible heat
watt/m 2
480
400
320
240
160
80
0
-80
-160
-240
40N
20N
90E
120E
China (2004-01-01)
photosynthesis
umol/m2s
17
13
9
5
1
40N
20N
90E
120E
China (2004-01-01)
atmospheric specific humidity kg/kg
0.042
0.034
0.026
0.018
0.01
0.002
40N
20N
90E
120E
China (2004-01-01)
ground temperature
K
440
360
280
200
40N
20N
90E
120E

NCL脚本语言


Net Radiation in Qinghai-Tibet Plateau (2010-01)
Net Radiation in Qinghai-Tibet Plateau (2010-02)
Net Radiation in Qinghai-Tibet Plateau (2010-03)
Net Radiation in Qinghai-Tibet Plateau (2010-04)
Net Radiation in Qinghai-Tibet Plateau (2010-05)
Net Radiation in Qinghai-Tibet Plateau (2010-06)
Net Radiation in Qinghai-Tibet Plateau (2010-07)
Net Radiation in Qinghai-Tibet Plateau (2010-08)
Net Radiation in Qinghai-Tibet Plateau (2010-09)
Net Radiation in Qinghai-Tibet Plateau (2010-10)
Net Radiation in Qinghai-Tibet Plateau (2010-11)
Net Radiation in Qinghai-Tibet Plateau (2010-12)
Net Radiation
W/m 2

三维可视化


DB: QZ_2008.clm2.h0.2010-07.nc
Cycle: 0 Time:31
Pseudocolor
Var: LatentHeat
Units: watt/m^2
150.0
111.3
72.60
33.89
-4.807
Max: 150.0
Min: -4.807
Pseudocolor
Var: NetRadiation
Units: watt/m^2
253.6
185.0
116.4
Pseudocolor
Var: FSH
165.5
63.87
-37.81
-139.5
Max: 165.5
Min: -241.2
Pseudocolor
41.24
15.93
Min: -15.93
Z
Y
X
user: Administrator
Wed Apr 25 16:04:19 2012

研究展望

研究展望

- 建模系统的GUI与命令行的取舍（web化可能是一个好的选择）
- R的并行化：Rmpi、pbdMPI、snow等
- R在地学上的大数据应用（卫星图像、雷达、气象站、各种观测仪器等）:: 寒旱区大数据中心
- R模型如何驱动google maps形成地学的空间决策支持系统