EF5 Overview

University of Oklahoma/HyDROS

Module 2.1

Outline – Day 2

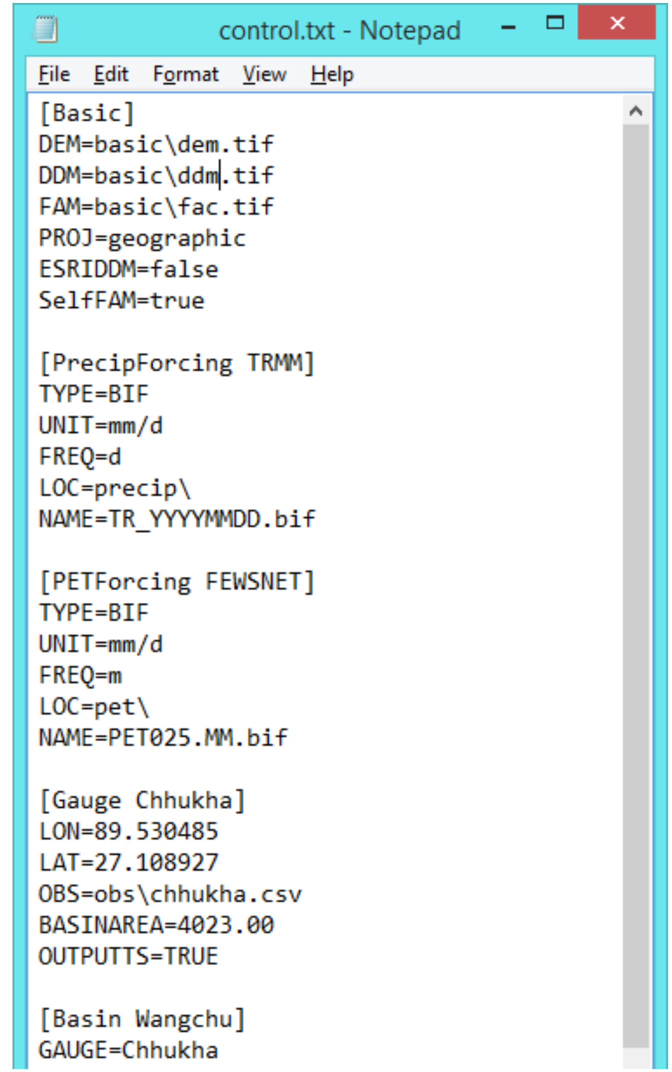
EF5 OVERVIEW

- Features of EF5
- Model structure
- Control file options
- Warm-up and model states
- Model evaluation indices

DEM DERIVATIVES

RAINFALL AND PET

AUTOMATIC CALIBRATION


control.txt - Notepad
File Edit Format View Help

```
[Basic]
DEM=basic\dem.tif
DDM=basic\ddm.tif
FAM=basic\fac.tif
PROJ=geographic
ESRIDDM=false
SelfFAM=true

[PrecipForcing TRMM]
TYPE=BIF
UNIT=mm/d
FREQ=d
LOC=precip\
NAME=TR_YYYYMMDD.bif

[PETForcing FEWSNET]
TYPE=BIF
UNIT=mm/d
FREQ=m
LOC=pet\
NAME=PET025.MM.bif

[Gauge Chhukha]
LON=89.530485
LAT=27.108927
OBS=obs\chhukha.csv
BASINAREA=4023.00
OUTPUTTS=TRUE

[Basin Wangchu]
GAUGE=Chhukha
```

Features of EF5

EF5 is the Ensemble Framework for Flash Flood Forecasting

- EF5 is a distributed hydrological model
- It takes advantage of multiple computing cores and parallel computing strategies for faster simulations

EF5 supports ensemble forecasting

- Multiple model cores exist in EF5 and more will be added in the future
- Each core contains different physics
- The user can thus see multiple model solutions for a particular scenario
- These multiple solutions are possible even with only *one* set of input data – this saves time!

Features of EF5

EF5 is cross platform

- This training is designed for Microsoft Windows
- EF5 was originally developed on Unix-like systems
- This includes Linux distributions and Mac OS X

EF5 is user-friendly

- As you saw in the Wang Chu example, if all the required data is provided, running EF5 is as simple as a double-click
- EF5 is flexible and can accept multiple input data formats
- If something goes wrong, EF5 includes helpful error messages to track down the problem

Features of EF5

EF5 includes coupling

Runoff generation and flow routing are coupled together which makes the simulation more realistic

EF5 output can be used to study atmospheric water, surface water, and subsurface water

EF5 is scalable

Water storage in the soil is modeled at the sub-grid scale, as are runoff generation processes, so you can think of the cells as “self-contained”, which means the number of cells and the size of the cells can be specified by the user

Suitable for flash floods at small scales, riverine floods at very large scales, or global forecasting

Model Structure

EF5 is broken down into two main sections

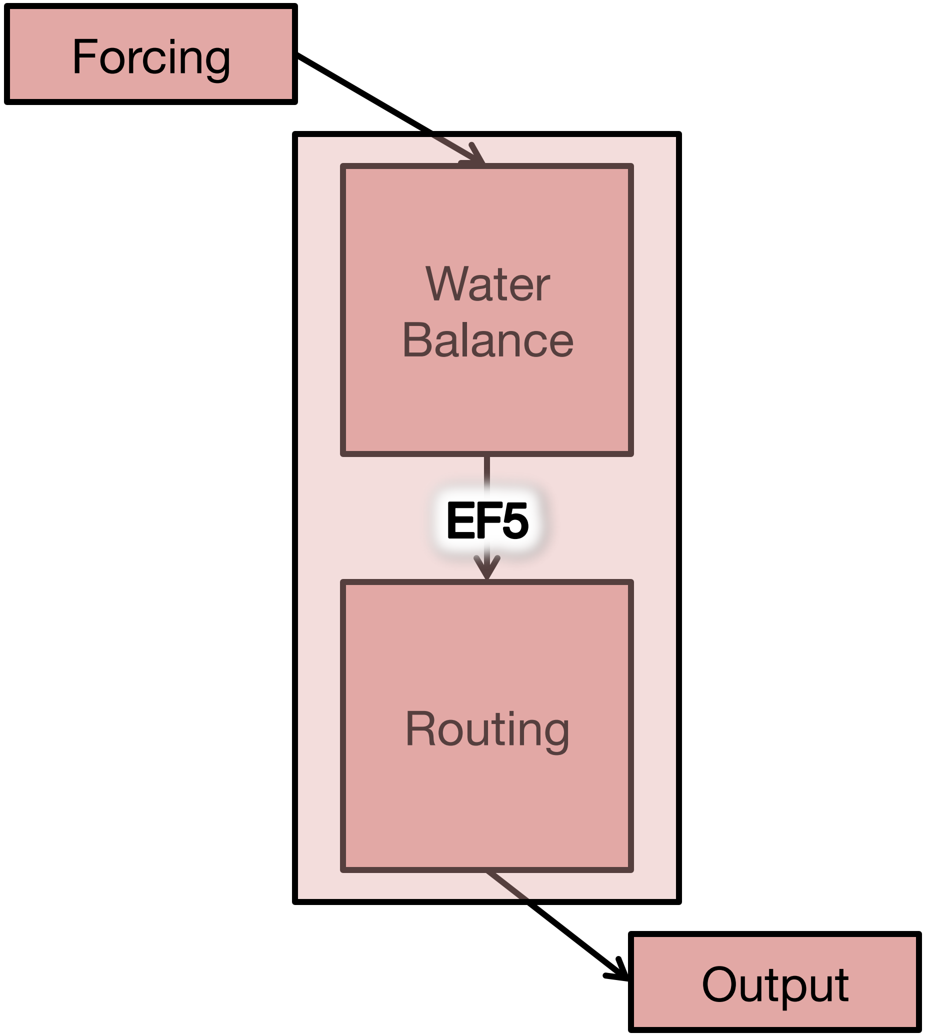
Water balance

Routing

The water balance concerns the water inputs and outputs for each model cell

Recall that the inputs (precipitation, upstream runoff, and interflow) must equal the outputs (runoff and interflow to be routed downstream)

Routing determines how quickly these outputs will travel downstream



Model Structure

Right now, EF5 has two options for the water balance with another under development

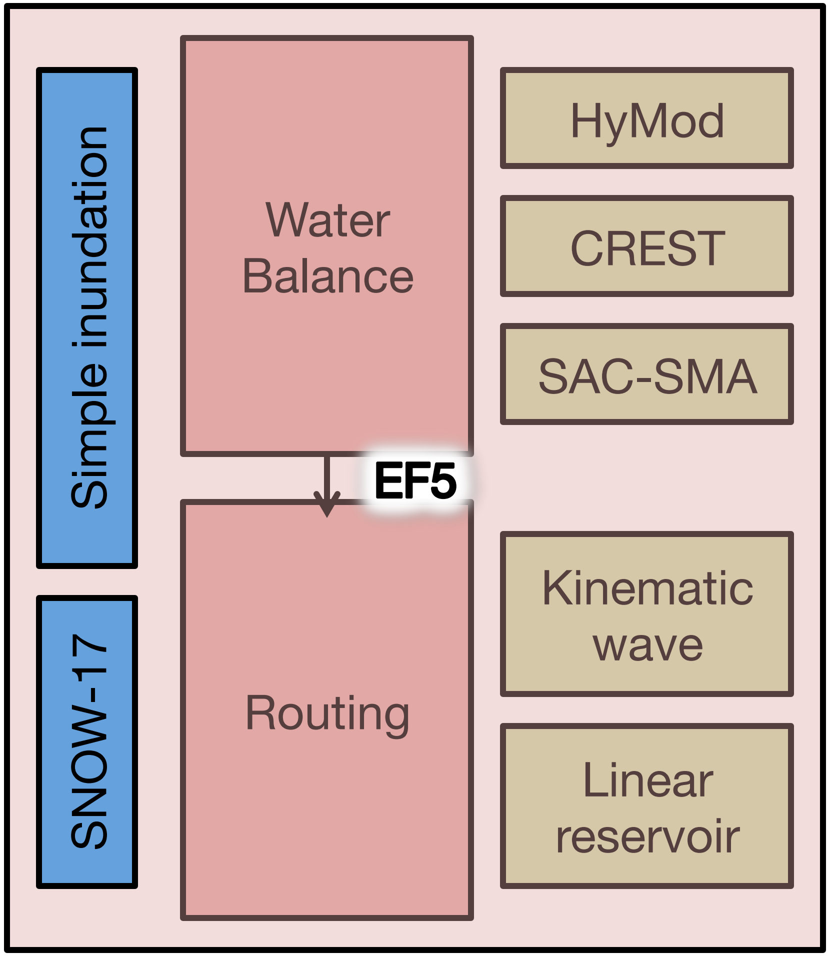
- CREST – used in all training examples
- SAC-SMA (Sacramento-Soil Moisture Accounting)
- HyMod (in development)

There are two choices for routing

- Kinematic wave – used in all training examples
- Linear reservoir (in development)

Other available model components

- SNOW-17
- Simple inundation

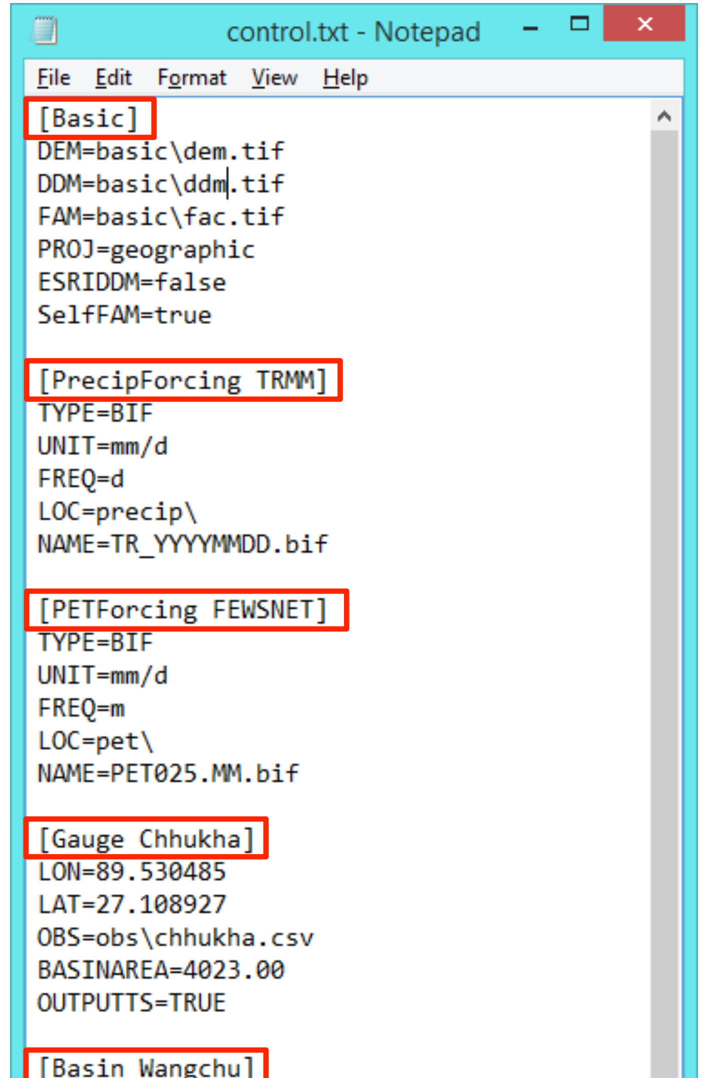


EF5 Control File

The EF5 control file tells the model what to do

It consists of several blocks, each of which starts with a [bracketed statement]. Types of blocks include…

- Basic
- PrecipForcing
- PETForcing
- Gauge
- Basin
- etc…


control.txt - Notepad

File Edit Format View Help

```
[Basic]
DEM=basic\dem.tif
DDM=basic\ddm.tif
FAM=basic\fac.tif
PROJ=geographic
ESRIDDM=false
SelfFAM=true

[PrecipForcing TRMM]
TYPE=BIF
UNIT=mm/d
FREQ=d
LOC=precip\
NAME=TR_YYYYMMDD.bif

[PETForcing FEWSNET]
TYPE=BIF
UNIT=mm/d
FREQ=m
LOC=pet\
NAME=PET025.MM.bif

[Gauge Chhukha]
LON=89.530485
LAT=27.108927
OBS=obs\chhukha.csv
BASINAREA=4023.00
OUTPUTTS=TRUE

[Basin Wangchu]
```

EF5 Control File

In each bracketed statement, the *type* of block should not be changed

But the user can change the *name* of each block

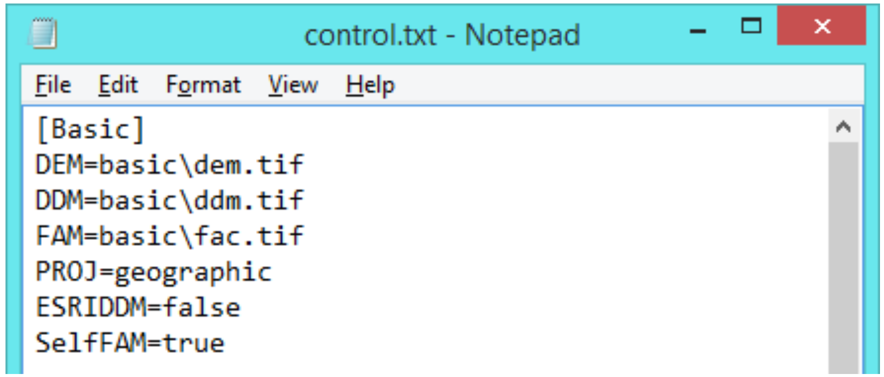
- TRRM
- FEWSNET
- Chhukha
- Wangchu
- etc…

control.txt - Notepad

File Edit Format View Help

```
[Basic]
DEM=basic\dem.tif
DDM=basic\ddm.tif
FAM=basic\fac.tif
PROJ=geographic
ESRIDDM=false
SelfFAM=true

[PrecipForcing TRMM]
TYPE=BIF
UNIT=mm/d
FREQ=d
LOC=precip\
NAME=TR_YYYYMMDD.bif

[PETForcing FEWSNET]
TYPE=BIF
UNIT=mm/d
FREQ=m
LOC=pet\
NAME=PET025.MM.bif

[Gauge Chhukha]
LON=89.530485
LAT=27.108927
OBS=obs\chhukha.csv
BASINAREA=4023.00
OUTPUTTS=TRUE

[Basin Wangchu]
```

Basic Block



control.txt - Notepad

File Edit Format View Help

```
[Basic]
DEM=basic\dem.tif
DDM=basic\ddm.tif
FAM=basic\fac.tif
PROJ=geographic
ESRIDDM=false
SelfFAM=true
```

The "Basic" block contains six pieces of information

- DEM, DDM, and FAM

  The file path to these topographical grids
- PROJ

  This is the map projection associated with the topographical information
- ESRIDDM and SelfFAM

  These are options related to how topographical information is encoded

More on all this in Module 2.2

PrecipForcing Block

```
[PrecipForcing TRMM]
TYPE=BIF
UNIT=mm/d
FREQ=d
LOC=precip\
NAME=TR_YYYYMMDD.bif
```

This "PrecipForcing" block is named "TRMM" in this example

- TYPE
  The file type of the precipitation
- UNIT
  The units of the precipitation
- FREQ
  The frequency at which EF5 will ingest precipitation files
- LOC
  The directory where the precipitation is located
- NAME
  How precipitation files are named (TR_20150101.bif, TR_20150102.bif, etc..)

More on all this in Module 2.3

PETForcing Block

```
[PETForcing FEWSNET]
TYPE=BIF
UNIT=mm/d
FREQ=m
LOC=pet\
NAME=PET025.MM.bif
```

This "PETForcing" block is named "FEWSNET" in this example

- TYPE
  The file type of the PET
- UNIT
  The units of the PET
- FREQ
  The frequency at which EF5 will ingest PET files
- LOC
  The directory where the PET is located
- NAME
  How PET files are named (PET025.01.bif, PET025.02.bif, etc.)

More on all this in Module 2.3

Gauge Block

```
[Gauge Chhukha]
LON=89.530485
LAT=27.108927
OBS=obs\chhukha.csv
BASINAREA=4023.00
OUTPUTTS=TRUE
```

Multiple "Gauge" blocks can be specified, one for each gauge you want to run the model at – this one is named "Chhukha"

- LON and LAT (or CELLX and CELLY)

  The longitude and latitude of the gauge (unprojected and in decimal degrees)

  (or the x- and y- grid coordinates of the gauge in the topographical basic files)
- OBS

  Location of a file with observed stream flow values (required for calibration)
- BASINAREA

  Contributing area of the gauge in square kilometers
- OUTPUTTS

  Output a time series file for the gauge

Basin Block

```
[Basin Wangchu]
GAUGE=Chhukha
```

Multiple "Basin" blocks can also be specified – this one is named "Wangchu"

- GAUGE

  The name of the gauge block to be included in this basin

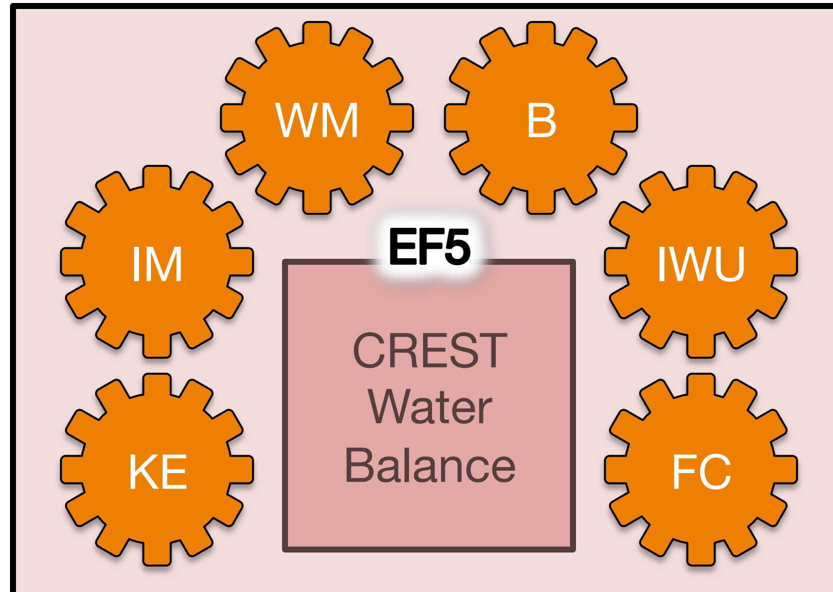
Note that the term basin here is not a physical basin

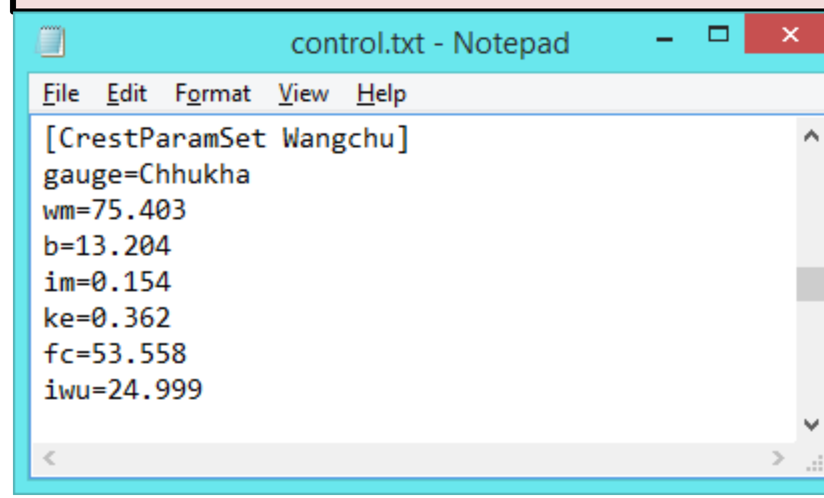
Basin here means a collection of gauges (or gauge blocks) to be modeled in a single run of EF5

Model Parameters

Each component of the model has its own set of parameters

- Parameters are like the knobs on a car radio
- They allow us to “tune” the model to get better simulations, like how you tune the radio to get better reception





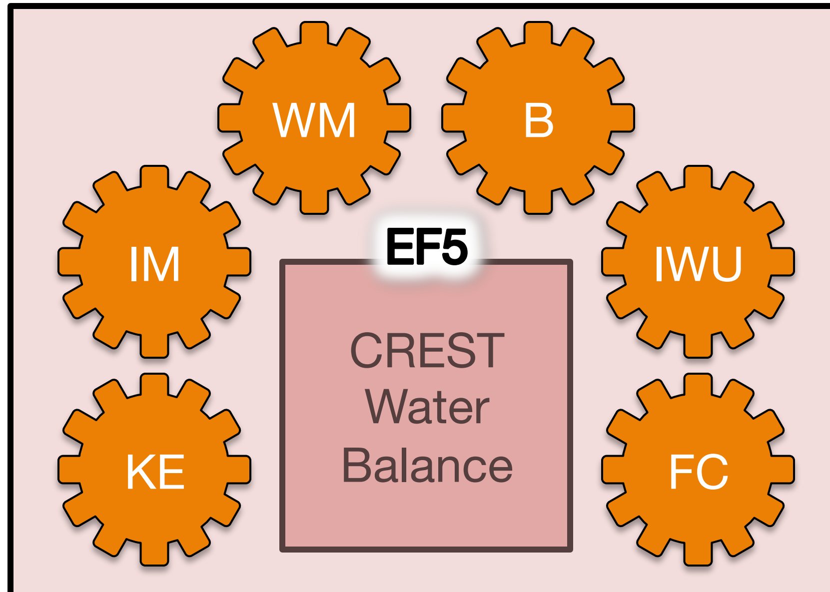
Let’s start with the CrestParamSet

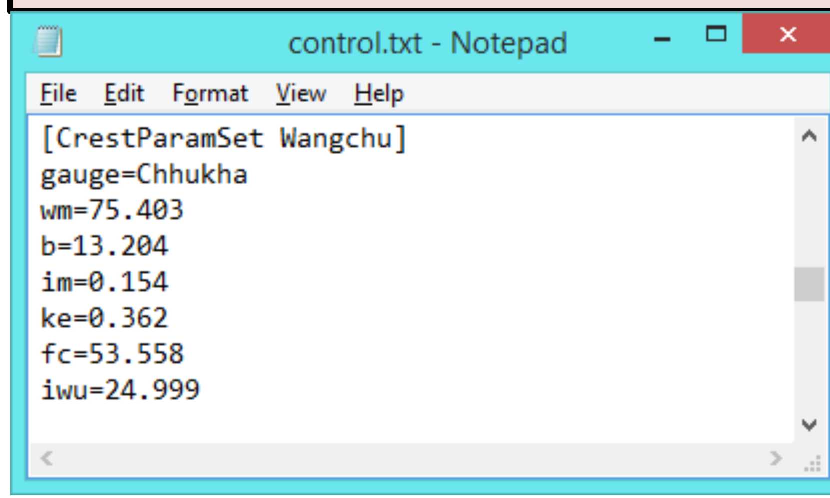
- 6 parameters

CrestParamSet Block

The CrestParamSet block in this example is named "Wangchu"

- GAUGE
  The name of the gauge block to which these parameters apply
- WM
  Maximum soil water capacity
- B
  The exponent of the VIC
- IM
  Impervious area ratio
- KE
  Conversion factor from PET to actual ET
- FC
  Soil saturated hydraulic conductivity
- IWU
  Initial value of soil water, a % of WM



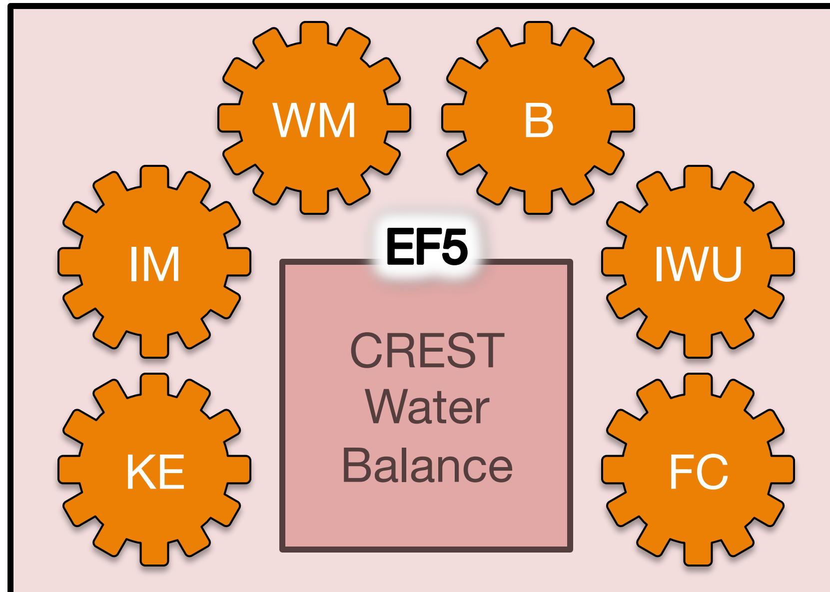


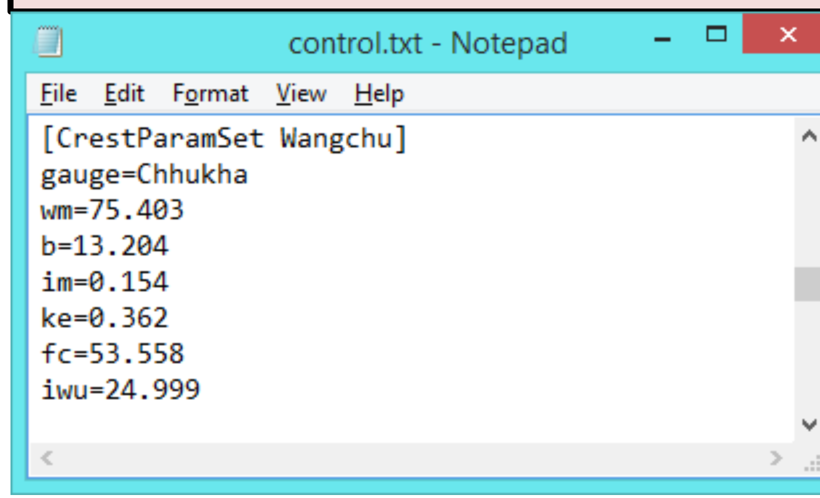

```
[CrestParamSet Wangchu]
gauge=Chhukha
wm=75.403
b=13.204
im=0.154
ke=0.362
fc=53.558
iwu=24.999
```

CrestParamSet Block

If you want to specify different parameters for different gauges, you can include additional CrestParamSet blocks

More on parameters in Module 2.4



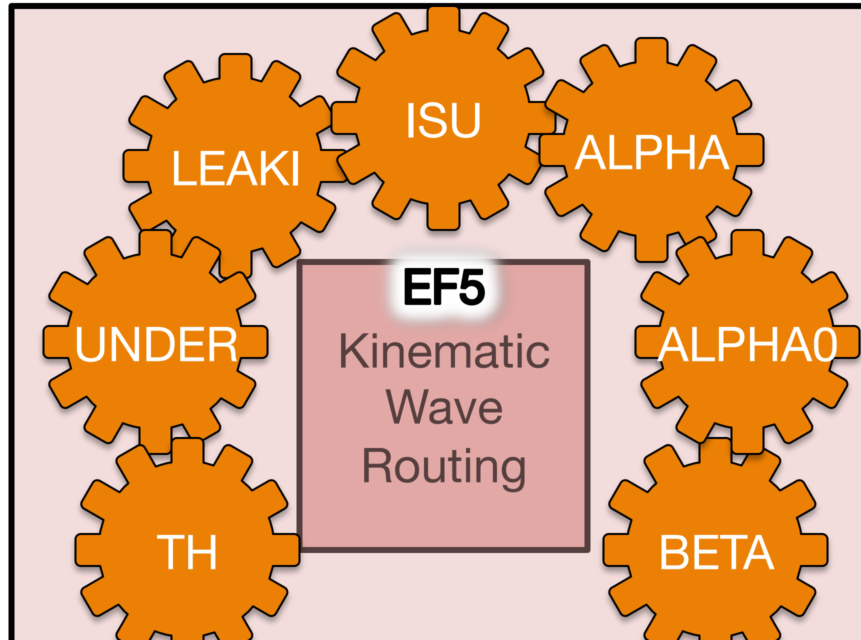



```
[CrestParamSet Wangchu]
gauge=Chhukha
wm=75.403
b=13.204
im=0.154
ke=0.362
fc=53.558
iwu=24.999
```

KWParamSet Block

The KWParamSet block in this example is named "Wangchu"

- GAUGE
  The name of the gauge block to which these parameters apply
- TH
  Number of grid cells needed to flow into a cell for it to be part of a channel
- UNDER
  The interflow flow speed multiplier
- LEAKI
  Amount of water leaked from interflow reservoir in each time step
- ISU
  Initial value of the interflow reservoir
- ALPHA and BETA
  Used in the equation Streamflow = alpha*(cross-sectional channel area)^{beta}
- ALPHA0
  The alpha value used for overland, not channel, routing



control.txt - Notepad

File Edit Format View Help

```
[kwparamset Wangchu]
gauge=Chhukha
under=2.976
leaki=0.042
th=4.031
isu=0.000000
alpha=2.847
beta=0.884
alpha0=1.174
```

Task Block

The Task block in this example is named "RunWangchu"

- STYLE

  This tells the model what time of task it will be completing (SIMU is used for basic simulations)

- MODEL

  Which water balance component will be used (CREST, SAC, or HyMOD)

- ROUTING

  Whether to use kinematic wave or linear reservoir routing (KW; LR)

- BASIN

  Name of the basin block over which the model should run

```
[Task RunWangchu]
STYLE=SIMU
MODEL=CREST
ROUTING=KW
BASIN=Wangchu
PRECIP=TRMM
PET=FEWSNET
OUTPUT=output\
PARAM_SET=Wangchu
ROUTING_PARAM_Set=Wangchu
TIMESTEP=1d
TIME_BEGIN=200101010000
TIME_WARMEND=200102010000
TIME_END=200212310000
```

Task Block

The Task block in this example is named “RunWangchu”

- PRECIP

  PrecipForcing block name
- PET

  PETForcing block name
- OUTPUT

  The location where the model output will be stored
- PARAM_SET

  Name of the CrestParamSet block
- ROUTING_PARAM_SET

  Name of the KWParamSet block

```
[Task RunWangchu]
STYLE=SIMU
MODEL=CREST
ROUTING=KW
BASIN=Wangchu
PRECIP=TRMM
PET=FEWSNET
OUTPUT=output\
PARAM_SET=Wangchu
ROUTING_PARAM_Set=Wangchu
TIMESTEP=1d
TIME_BEGIN=200101010000
TIME_WARMEND=200102010000
TIME_END=200212310000
```

Task Block

The Task block in this example is named "RunWangchu"

- TIMESTEP

  Time step used in the simulation (can be in years, months, days, hours, minutes, or seconds)
- TIME_BEGIN

  Time at which the model starts, in YYYYMMDDHHUUSS
- TIME_WARMEND

  Optional, time when the warm-up period ends
- TIME_END

  Time when the model run ends

```
[Task RunWangchu]
STYLE=SIMU
MODEL=CREST
ROUTING=KW
BASIN=Wangchu
PRECIP=TRMM
PET=FEWSNET
OUTPUT=output\
PARAM_SET=Wangchu
ROUTING_PARAM_Set=Wangchu
TIMESTEP=1d
TIME_BEGIN=200101010000
TIME_WARMEND=200102010000
TIME_END=200212310000
```

EF5 Control File

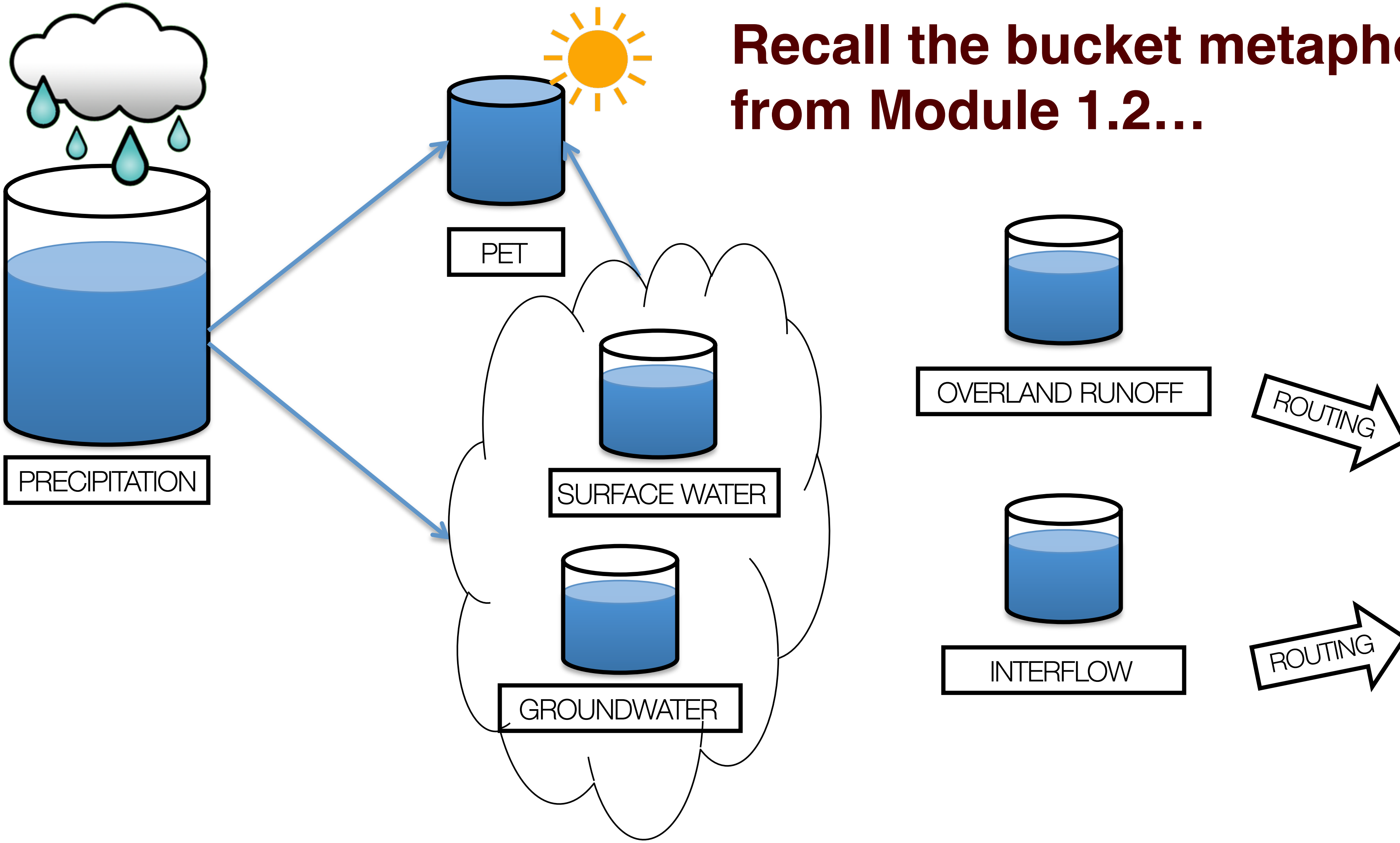
We've so far just described the basic options and requirements in the EF5 control file, but the model has many additional components that will be considered in the rest of the training

Now let's discuss two important modeling concepts:

- Model warm-up
- Model states

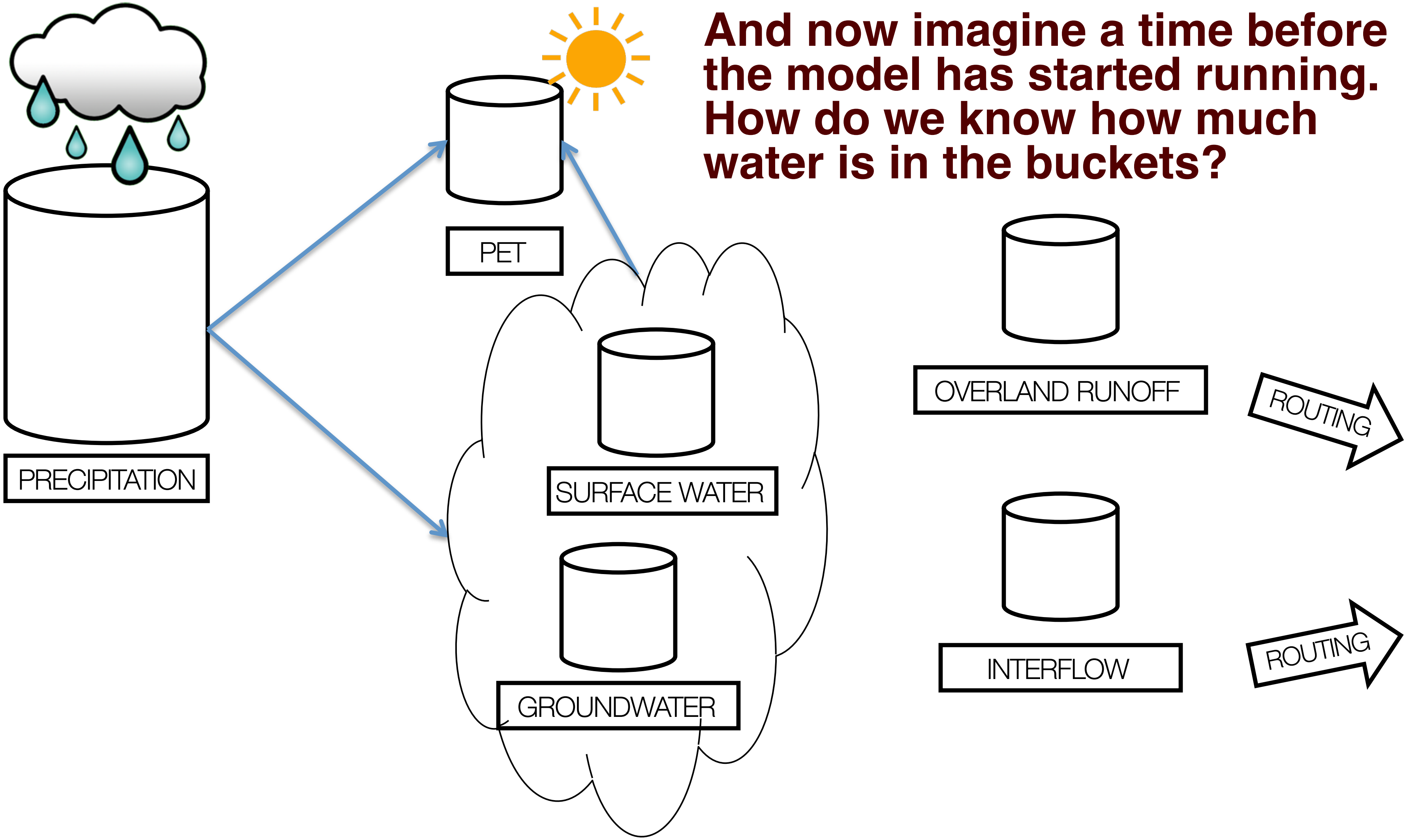
Warming up the Model

Recall the bucket metaphor from Module 1.2…



Warming up the Model

And now imagine a time before the model has started running. How do we know how much water is in the buckets?



Warming up the Model

The simple answer is we usually don't know

So we give the model some guess or estimate of the <u>initial conditions</u> or <u>initial model states</u>

These estimates aren't perfect, but that usually doesn't matter, thanks to the process of model warm-up

Warming up the Model

Here's how it works:

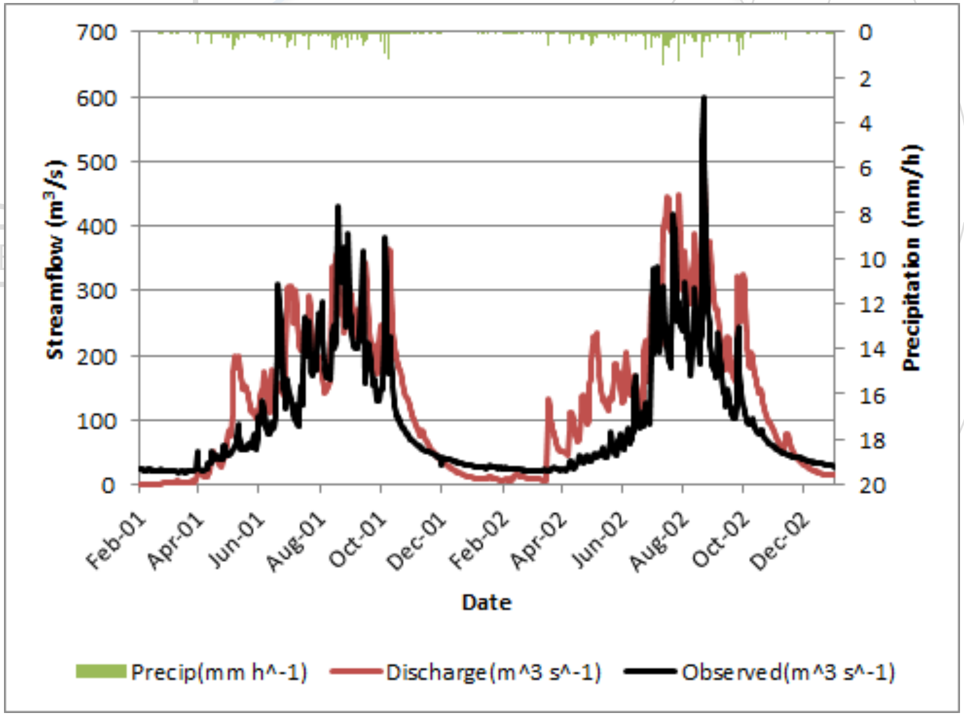
- You assign initial conditions to the model using parameters
- So, let's say you tell the model that your surface water, groundwater, runoff, and interflow buckets are half full
- Hopefully that's more accurate than starting the model completely dry (that is, with no water in any of the buckets)
- Now you let the model run for some specified <u>warm-up period</u> until the bucket levels get closer to reality

And then you only use the model results for the time period <u>after</u> the warm-up is complete

Warming up the Model

Without knowing it, we had a short (1 month) warm-up period in our first example

- The simulation began 1 Jan 2001
- The warm-up period ended 1 Feb 2001




```
[Task RunWangchu]
STYLE=SIMU
MODEL=CREST
ROUTING=KW
BASIN=Wangchu
PRECIP=TRMM
PET=FEWSNET
OUTPUT=output\
PARAM_SET=Wangchu
ROUTING_PARAM_Set=Wangchu
TIMESTEP=1d
TIME_BEGIN=200101010000
TIME_WARMEND=200102010000
TIME_END=200212310000
```

Warming up the Model

EF5 does not output the results from the model during the warm-up period, because they would not typically be used for any analysis

Let's see what happens without a warm-up period

In `\EF5_training\examples\wangchu\` **open** `control.txt` **and add a # symbol before "TIME_WARMEND"**

control.txt - Notepad

File Edit Format View Help

```
BASIN=Wangchu
PRECIP=TRMM
PET=FEWSNET
OUTPUT=output\
PARAM_SET=Wangchu
ROUTING_PARAM_Set=Wangchu
TIMESTEP=1d
TIME_BEGIN=200101010000
#TIME_WARMEND=200102010000
TIME_END=200212310000

[Execute]
TASK=RunWangchu
```

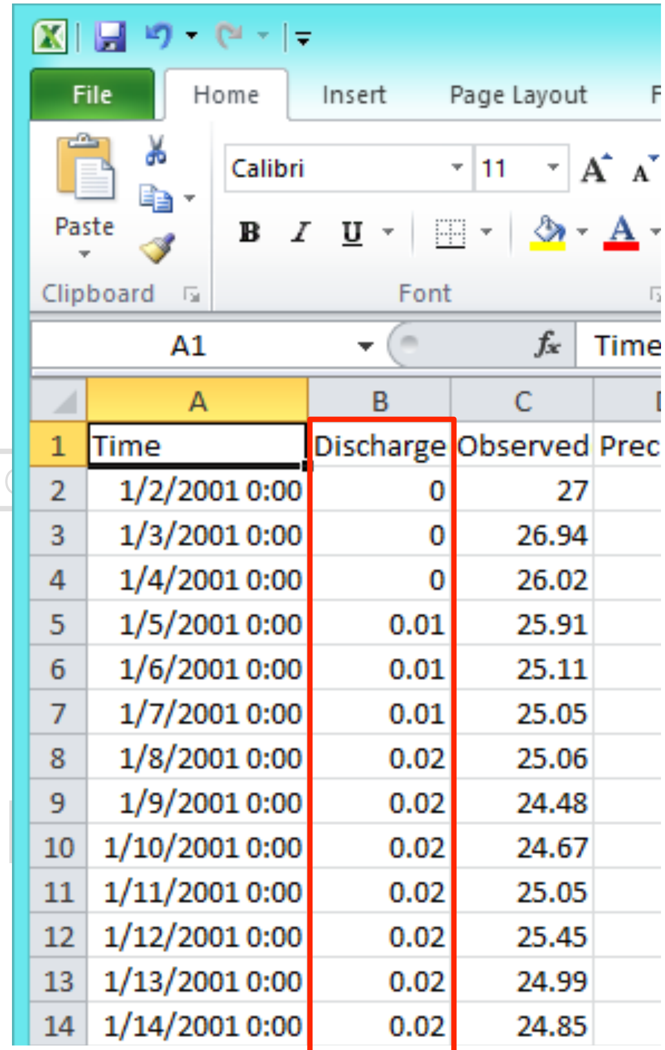
Warming up the Model

This turns that line into a “comment” so that EF5 doesn’t read it as part of the control file anymore

Save the control file and double-click `RunEF5.bat`

Open the results in your `output` **folder**

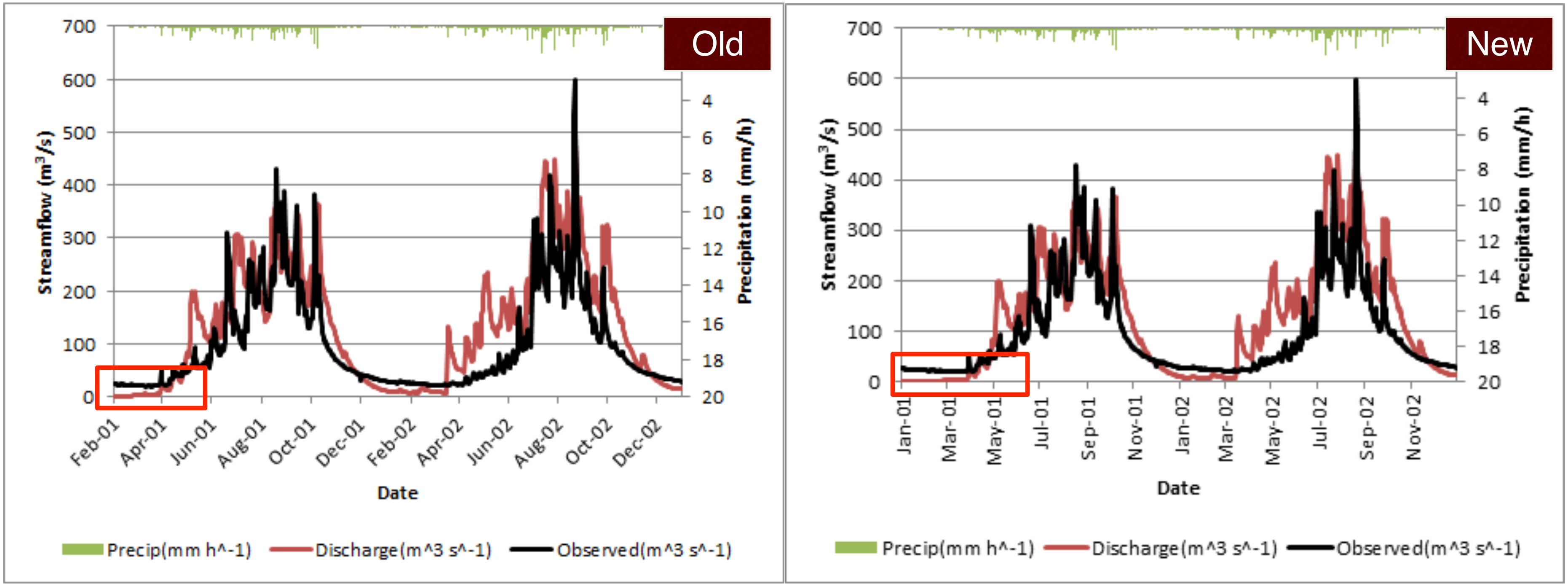
Notice how small the “Discharge” values are in January 2001 compared to the “Observed”



	A	B	C	D
1	Time	Discharge	Observed	Prec
2	1/2/2001 0:00	0	27	
3	1/3/2001 0:00	0	26.94	
4	1/4/2001 0:00	0	26.02	
5	1/5/2001 0:00	0.01	25.91	
6	1/6/2001 0:00	0.01	25.11	
7	1/7/2001 0:00	0.01	25.05	
8	1/8/2001 0:00	0.02	25.06	
9	1/9/2001 0:00	0.02	24.48	
10	1/10/2001 0:00	0.02	24.67	
11	1/11/2001 0:00	0.02	25.05	
12	1/12/2001 0:00	0.02	25.45	
13	1/13/2001 0:00	0.02	24.99	
14	1/14/2001 0:00	0.02	24.85	

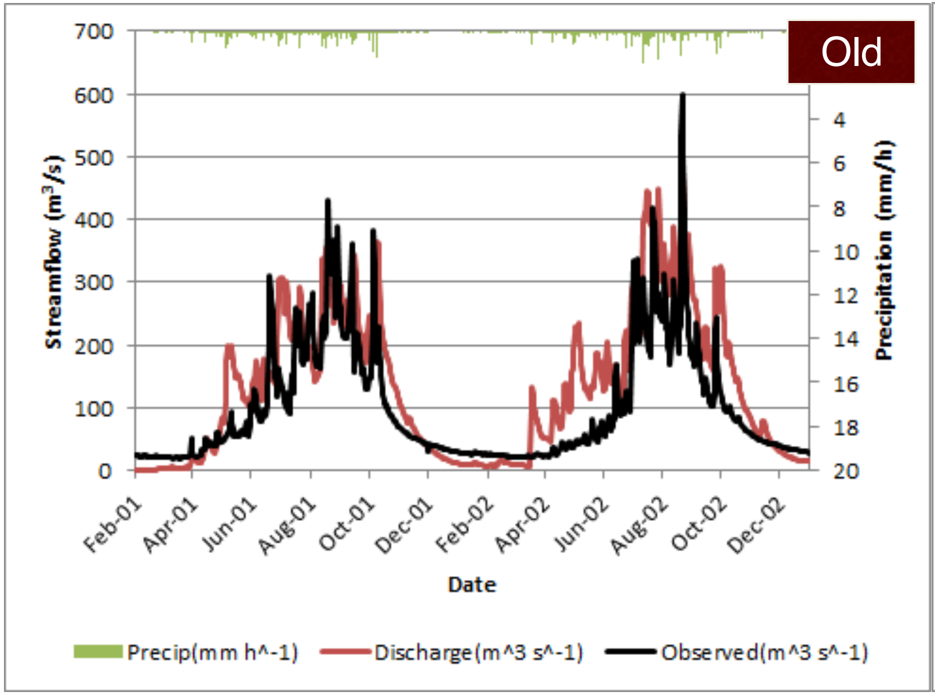
I went ahead and created a hydrograph using the instructions in Module 1.2

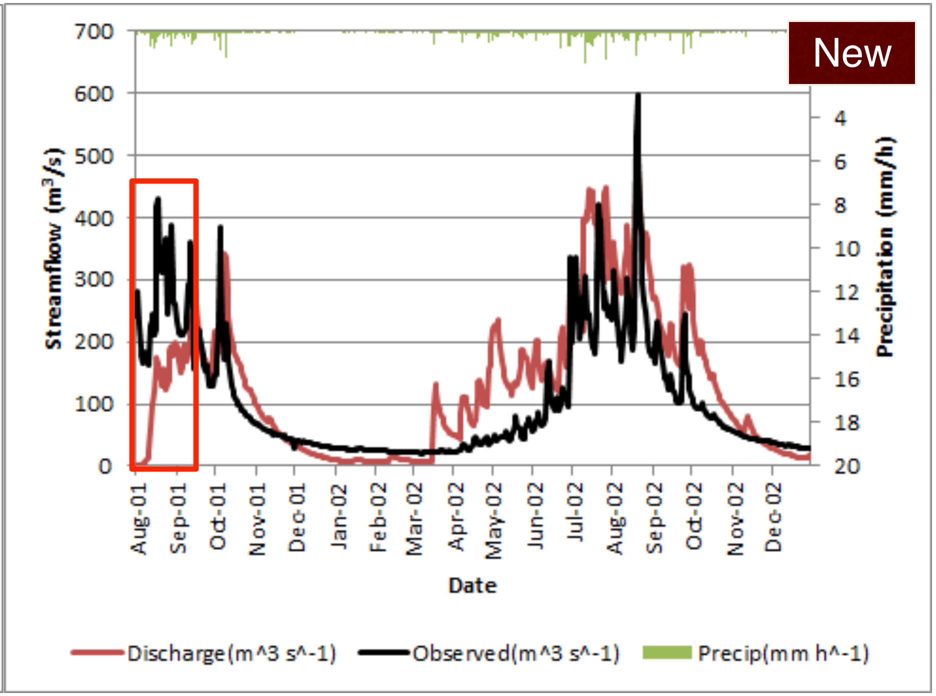
Our model results (blue) are too low at the beginning of each hydrograph



Warming up the Model

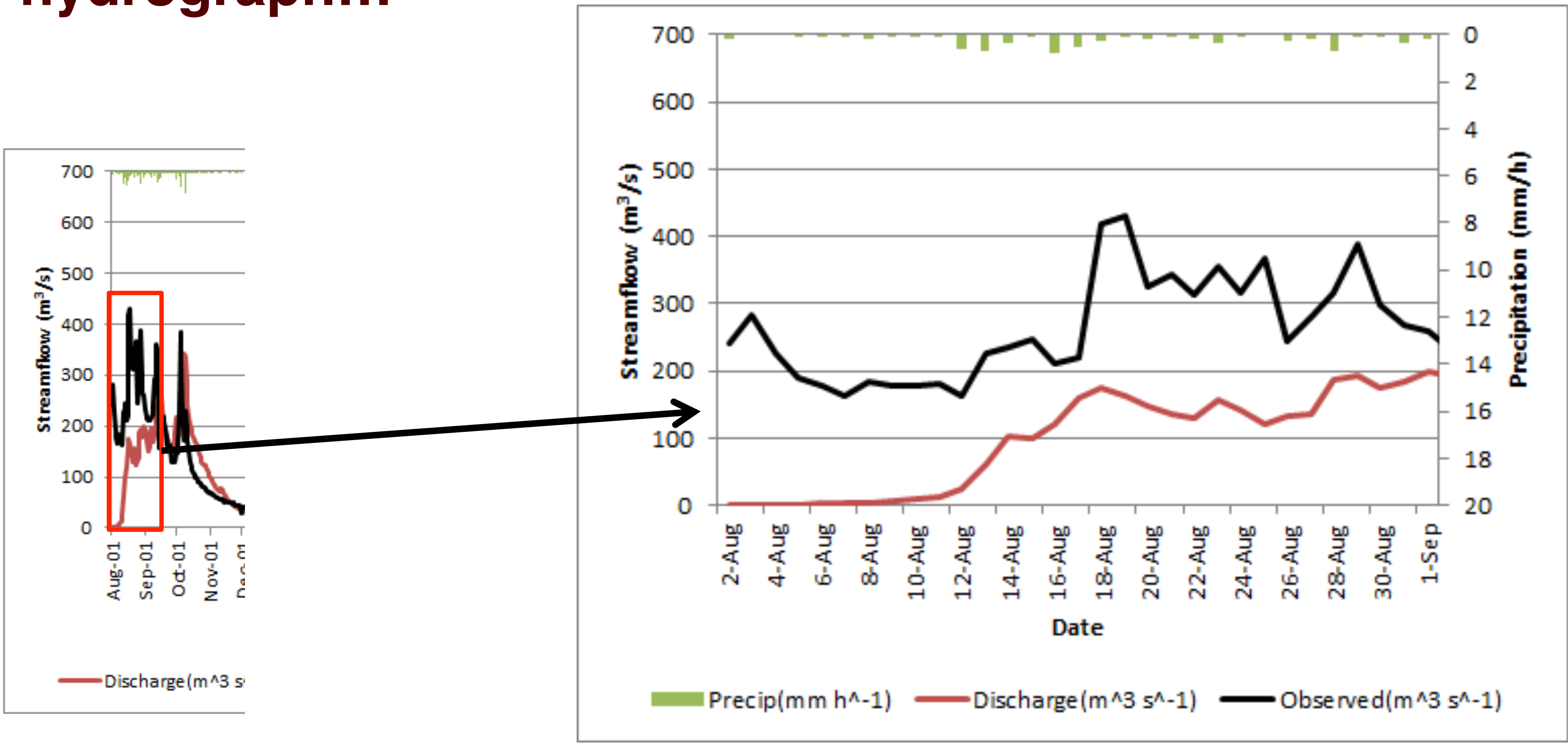
What if I start the simulation during the rainy season, with no warm-up period? Let's try a start date of August 1, 2001 (200108010000)





Warming up the Model

Then let's zoom in on the beginning of the hydrograph…

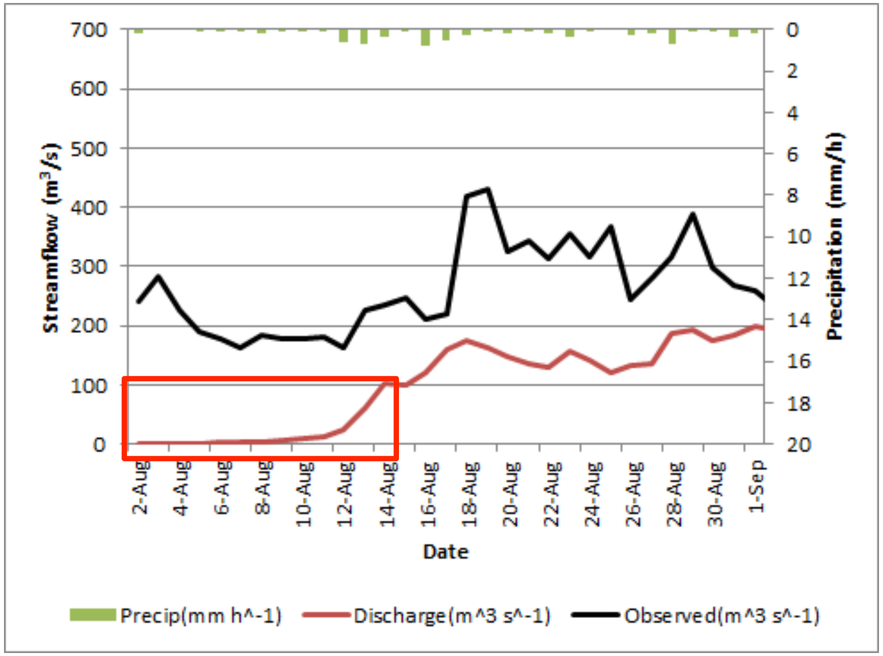


Warming up the Model

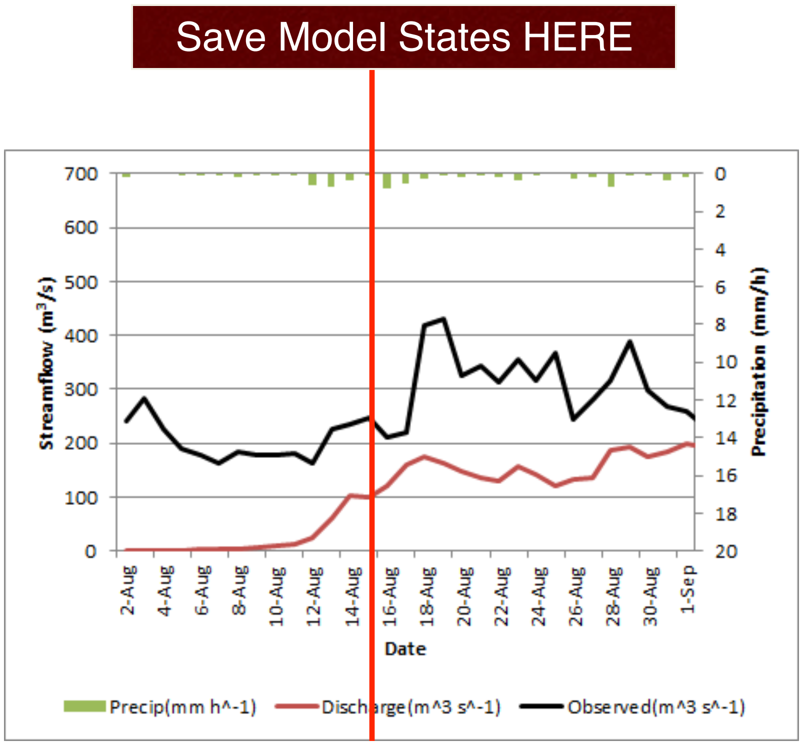
You can see the simulation start to pick up between August 12th and 18th, but before then it's far too low

This suggests a warm-up period of 2 weeks is required, and a month would be pretty safe

Running without a warm-up period risks the inclusion of bad results at the beginning of your simulation



Model States



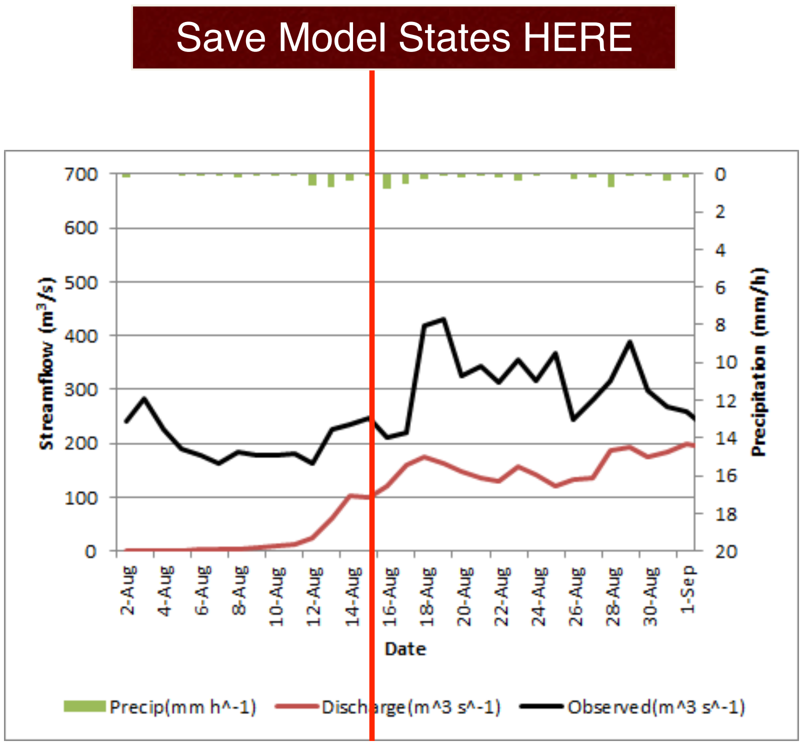
Now let's imagine a different situation

Say I'm running the model every day at the same time on Wang Chu Basin as new rainfall data becomes available

But I don't want to waste my computer time starting the model over at the beginning (including a warm-up period) each day just so I can get the latest day's results

The solution is to save the <u>model states</u> – the stream flow and the soil moisture at the end of the model run

Model States



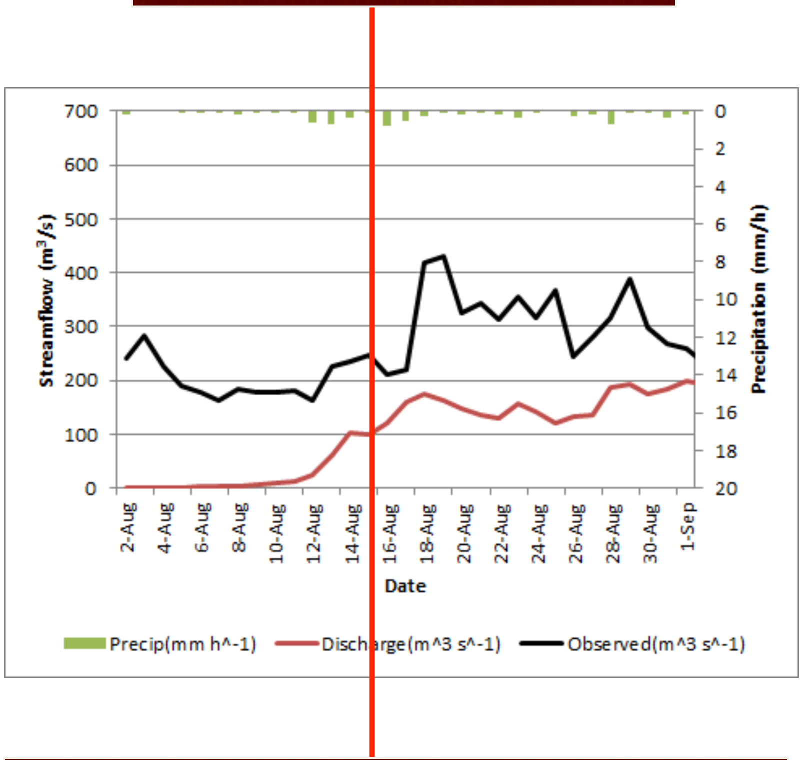
Saving model states is a powerful technique

A large basin with high-resolution information might require a 6 month warm-up period, which could take 6 hours to run

If I want twice-daily stream flow forecast updates, that means my computer might run 12 hours a day just to provide those two forecasts

Model States

Save Model States HERE



Now when I run the forecast for August 16, I can use the saved model states from August 15, instead of warming up from August 1 to August 15

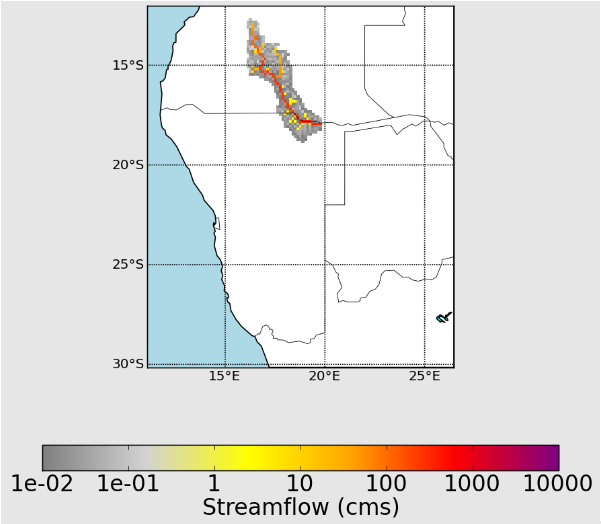
BUT – if I save model states at the end of each forecast I produce, I can run the model forward without a warm-up period, because I know how much water was in all the buckets at the end of the run

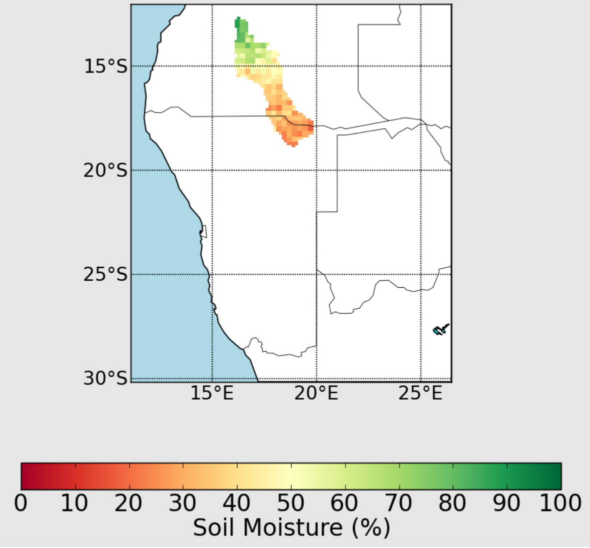
Now my forecast update might only take 5 minutes to run, instead of 6 hours

Model States

This technique allows the experimental OU/NASA forecasts of streamflow and soil moisture in Namibia, over the Okavango River Basin upstream of Rundu

flash.ou.edu/namibia

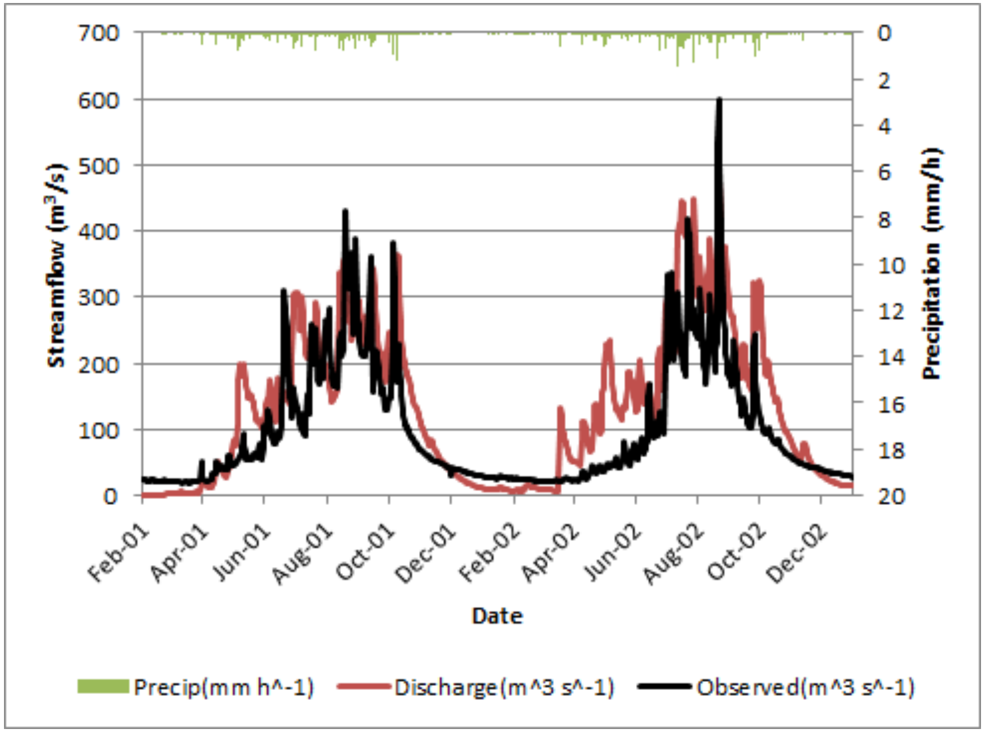




Each new set of forecast rainfall or estimated satellite rainfall can be “added” to be previous model run without warm-up, so that the results are available quickly

Recall Example 1

We produced this hydrograph…



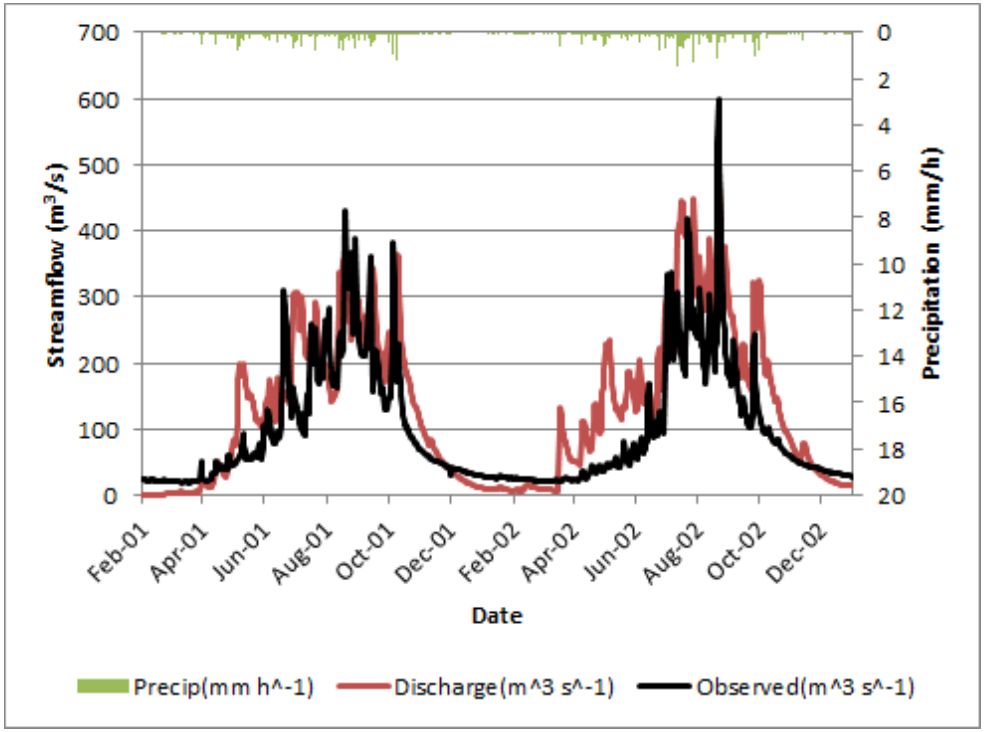
How do we know if this is a "good" result from our model?

Model Evaluation Indices

In general, the result is better the closer the match between the observed streamflow (black) and the discharge (red)

But could you tell if the model was better in 2001 or 2002 just by looking at the hydrograph?

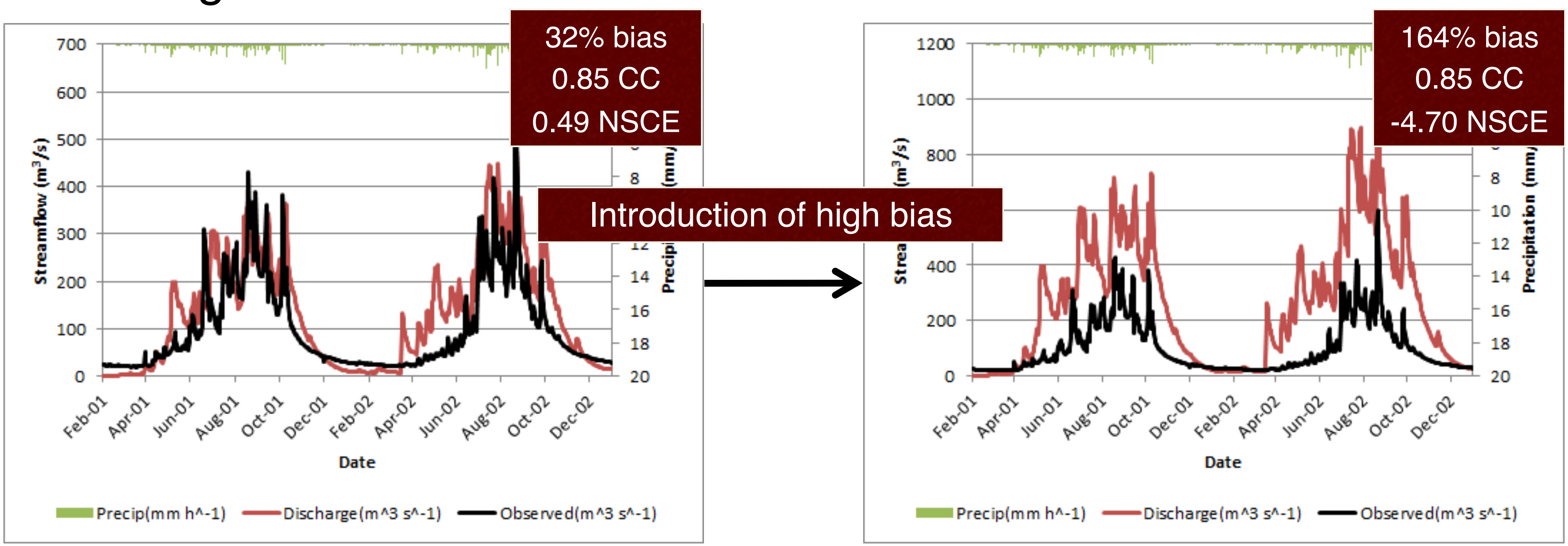
That's why several model evaluation indices exist



Model Evaluation Indices

Bias

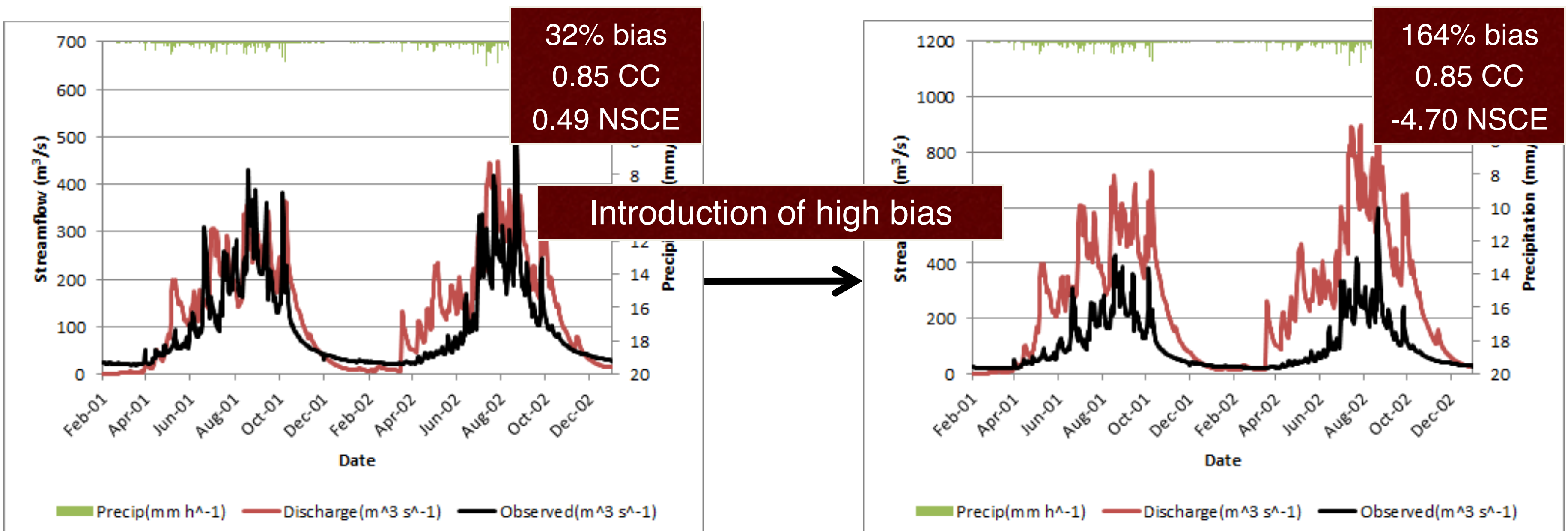
- Bias can be thought of as a shift up or down in a hydrograph
- The lower-right hydrograph now exhibits high bias – the red line has shifted up
- This is a conditional bias: the higher the original value, the greater the bias



Model Evaluation Indices

Bias

- Bias could be constant, or non-stationary
- Bias is expressed as a percent
- When bias > 0, the modeled values are too high, on average
- We like bias values close to zero



Model Evaluation Indices

How to calculate bias:

$$\text{Bias} = \left[\frac{\sum_{i=1}^{n} R_{sim,i} - \sum_{i=1}^{n} R_{obs,i}}{\sum_{i=1}^{n} R_{obs,i}} \right] \times 100$$

The first term $\sum_{i=1}^{n} R_{sim,i}$ **is the sum of all the simulated streamflow values**

The second term $\sum_{i=1}^{n} R_{obs,i}$ **is the sum of all the observed streamflow values**

In Microsoft Excel:

$$\sum_{i=1}^{n} R_{sim,i}$$

Use the SUM formula to add all the simulated values together

K3 | f_x =SUM(B2:B700)

Formula Bar

	A	B	C	D	E	F	G	H	I	J	K
1	Time	Discharge	Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow(mm*1000)			
2	#########	0.08	25.16	0	0	25.31	0	0.0001		Sum of Observed Values:	
3	#########	0.09	25.32	0	0	25.41	0	0.0003		Sum of Simulated Values:	91544.53

$$\sum_{i=1}^{n} R_{obs,i}$$

Use the SUM formula to add all the observed values together

K2 | f_x =SUM(C2:C700)

Formula Bar

	A	B	C	D	E	F	G	H	I	J	K
1	Time	Discharge	Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow(mm*1000)			
2	#########	0.08	25.16	0	0	25.31	0	0.0001		Sum of Observed Values:	69441.36
3	#########	0.09	25.32	0	0	25.41	0	0.0003		Sum of Simulated Values:	91544.53

$$\sum_{i=1}^{n} R_{sim,i} - \sum_{i=1}^{n} R_{obs,i}$$

Subtract the sum of observed values from the simulated sum

K4 | f_x =K3-K2

Formula Bar

	A	B	C	D	E	F	G	H	I	J	K
1	Time	Discharge	Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow(mm*1000)			
2	#########	0.08	25.16	0	0	25.31	0	0.0001		Sum of Observed Values:	69441.36
3	#########	0.09	25.32	0	0	25.41	0	0.0003		Sum of Simulated Values:	91544.53
4	#########	0.1	24.39	0	0	25.46	0	0.0001		Subtract:	22103.17

In Microsoft Excel:

Divide

$$\left[\frac{\sum_{i=1}^{n} R_{sim,i} - \sum_{i=1}^{n} R_{obs,i}}{\sum_{i=1}^{n} R_{obs,i}}\right]$$

K5 | f_x =K4/K2

Formula Bar

A	B	C	D			G	H	I	J	K
	Discharge	Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow(mm*1000)			
####	0.08	25.16	0	0	25.31	0	0.0001		Sum of Observed Values:	69441.36
####	0.09	25.32	0	0	25.41	0	0.0003		Sum of Simulated Values:	91544.53
####	0.1	24.39	0	0	25.46	0	0.0001		Subtract:	22103.17
####	0.12	23	0	0	25.48	0	0		Divide:	0.3183

And multiply by 100

$$\text{Bias} = \left[\frac{\sum_{i=1}^{n} R_{sim,i} - \sum_{i=1}^{n} R_{obs,i}}{\sum_{i=1}^{n} R_{obs,i}}\right] \times 100$$

f_x =K5*100

Formula Bar

D	E	F	G	H		J	K
Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow(mm*1000)			
0	0	25.31	0	0.0001		Sum of Observed Values:	69441.36
0	0	25.41	0	0.0003		Sum of Simulated Values:	91544.53
0	0	25.46	0	0.0001		Subtract:	22103.17
0	0	25.48	0	0		Divide:	0.3183
0	0	25.49	0	0		Multiply by 100:	31.82998

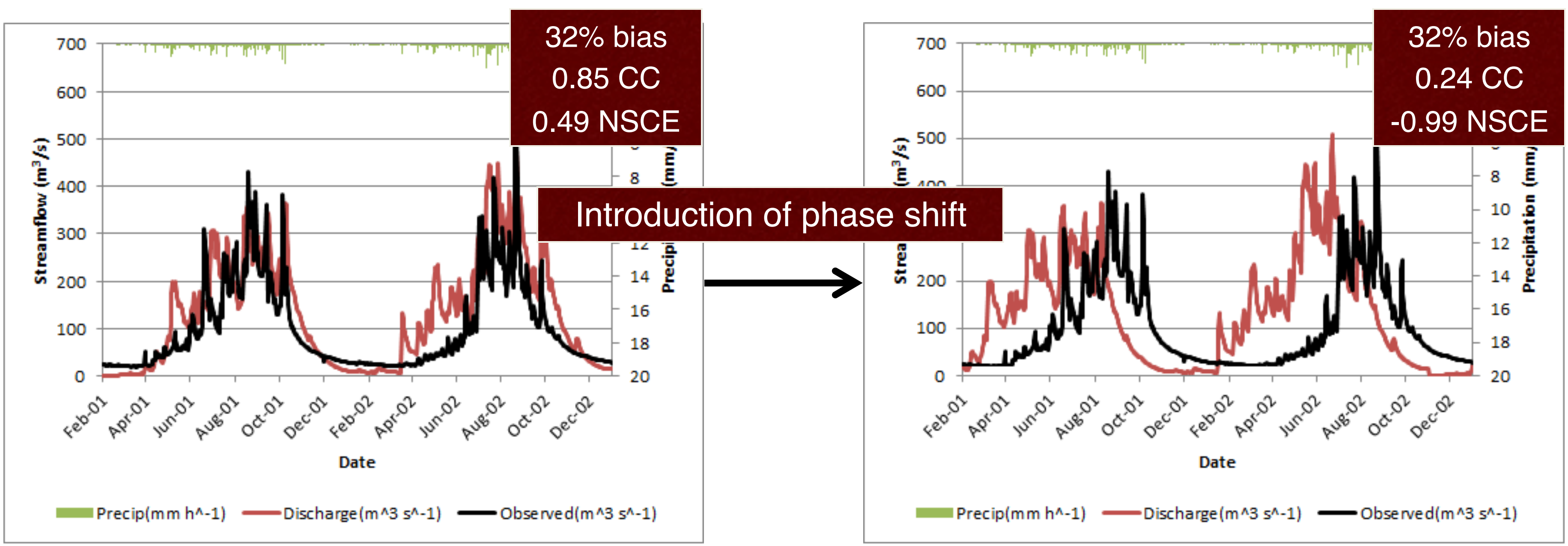
Thus, the bias in this example is 32%

Model Evaluation Indices

Correlation coefficient

This measures the agreement between the simulation and the observation time series

Measured on a 0 to 1 scale, where 1 is desirable and 0 is not desirable

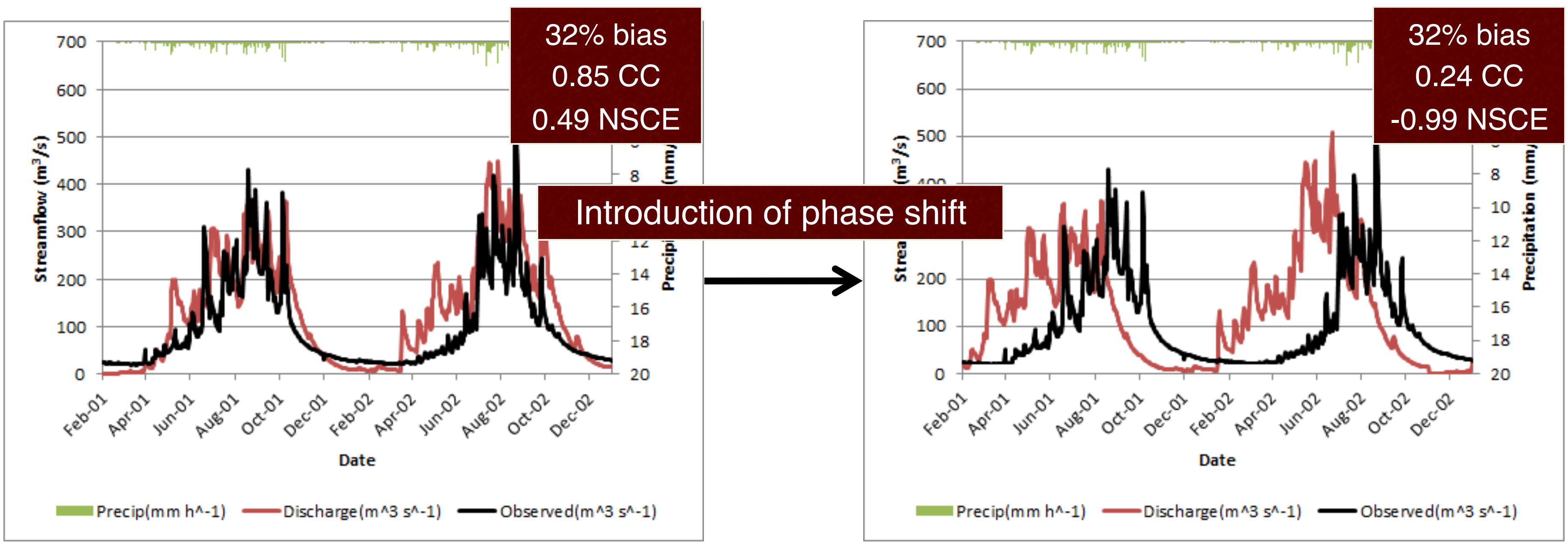


Model Evaluation Indices

Correlation coefficient

If bias identifies how far vertically displaced the simulation is from the observations, then correlation coefficient measures horizontal displacement between the two time series

This is highly simplified, but for our purposes it will work



How to calculate correlation coefficient

It looks complicated, but is actually extremely easy in Microsoft Excel

Just use the CORREL function – our answer is 0.85 for this example

$$CC = \frac{\sum_{i=1}^{n}\left(R_{obs,i}-\overline{R_{obs}}\right)\left(R_{sim,i}-\overline{R_{sim}}\right)}{\sqrt{\sum_{i=1}^{n}\left(R_{obs,i}-\overline{R_{obs}}\right)^2 \sum_{i=1}^{n}\left(R_{sim,i}-\overline{R_{sim}}\right)^2}}$$

K8 | =CORREL(B2:B700,C2:C700)

Formula Bar

	A	B	C	D	E	F	G	H	I	J	K
1	Time	Discharge	Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow(mm*1000)			
2	#########	0.08	25.16	0	0	25.31	0	0.0001		Sum of Observed Values:	69441.36
3	#########	0.09	25.32	0	0	25.41	0	0.0003		Sum of Simulated Values:	91544.53
4	#########	0.1	24.39	0	0	25.46	0	0.0001		Subtract:	22103.17
5	#########	0.12	23	0	0	25.48	0	0		Divide:	0.3183
6	#########	0.15	21.85	0	0	25.49	0	0		Multiply by 100:	31.82998
7	#########	0.16	21.48	0	0	25.53	0	0.0001			
8	#########	0.17	22.1	0	0	25.53	0	0		Correlation Coefficient:	0.853987
9	#########	0.17	22.48	0	0	25.53	0	0			

Nash-Sutcliffe Coefficient of Model Efficiency

Referred to as NSCE or "Nash"

Exists on a scale from -∞ to 1

When NSCE > 0, the model has more skill than if you just used the average stream flow over the period of the simulation

On a hydrograph, NSCE can be visualized as the amount of space or distance between the simulation and the observation, in a general sense

$$\mathrm{NSCE} = 1 - \frac{\sum_{i=1}^{n} \left(R_{obs,i} - R_{sim,i}\right)^2}{\sum_{i=1}^{n} \left(R_{obs,i} - \overline{R_{obs}}\right)^2}$$

How to calculate NSCE

$$\text{NSCE} = 1 - \frac{\sum_{i=1}^{n}\left(R_{obs,i} - R_{sim,i}\right)^2}{\sum_{i=1}^{n}\left(R_{obs,i} - \overline{R_{obs}}\right)^2}$$

Let's get $\overline{R_{obs}}$ first

SUM =AVERAGE(C$2:C$700)

	A	B	C	D	E	F	G	H	I
1	Time	Discharge	Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow	Average of observations
2	#########	0.08	25.16	0	0	25.31	0	0.0001	=AVERAGE(C$2:C$700)
3	#########	0.09	25.32	0	0	25.41	0	0.0003	99.34386266
4	#########	0.1	24.39	0	0	25.46	0	0.0001	99.34386266
5	#########	0.12	23	0	0	25.48	0	0	99.34386266
6	#########	0.15	21.85	0	0	25.49	0	0	99.34386266

What I've done here is take the average of the whole column of observations and then fill all of Column I with that same number

Model Evaluation Indices

How to calculate NSCE

- Now we'll calculate the numerator $\sum_{i=1}^{n}\left(R_{obs,i} - R_{sim,i}\right)^2$
- We need the square of the sum of the differences between each observation and simulation
- Luckily, Excel has a built-in formula called SUMXMY2 that will do this for us

=SUMXMY2(C2:C700,B2:B700)

Formula Bar

B	C	D	E	F		H	I	J	K
Discharge	Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow	Average of observations		
0.08	25.16	0	0	25.31	0	0.0001	99.34386266	Sum of Observed Values:	69441.36
0.09	25.32	0	0	25.41	0	0.0003	99.34386266	Sum of Simulated Values:	91544.53
0.1	24.39	0	0	25.46	0	0.0001	99.34386266	Subtract:	22103.17
0.12	23	0	0	25.48	0	0	99.34386266	Divide:	0.3183
0.15	21.85	0	0	25.49	0	0	99.34386266	Multiply by 100:	31.82998
0.16	21.48	0	0	25.53	0	0.0001	99.34386266		
0.17	22.1	0	0	25.53	0	0	99.34386266	Correlation Coefficient:	0.853987
0.17	22.48	0	0	25.53	0	0	99.34386266		
0.18	23.12	0	0	25.58	0	0.0001	99.34386266		
0.19	23.21	0	0	25.67	0	0.0003	99.34386266	NSCE Numerator:	3170080

How to calculate NSCE

- Now we'll calculate the denominator $\sum_{i=1}^{n}\left(R_{obs,i} - \overline{R_{obs}}\right)^2$
- We need the square of the sum of the differences between each observation and the mean observation
- Use SUMXMY2 again

=SUMXMY2(C2:C700,I2:I700)

Formula Bar

C	D	E	F	G	H	I	J	K
Observed	Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow	Average of observations		
25.16	0	0	25.31	0	0.0001	99.34386266	Sum of Observed Values:	69441.36
25.32	0	0	25.41	0	0.0003	99.34386266	Sum of Simulated Values:	91544.53
24.39	0	0	25.46	0	0.0001	99.34386266	Subtract:	22103.17
23	0	0	25.48	0	0	99.34386266	Divide:	0.3183
21.85	0	0	25.49	0	0	99.34386266	Multiply by 100:	31.82998
21.48	0	0	25.53	0	0.0001	99.34386266		
22.1	0	0	25.53	0	0	99.34386266	Correlation Coefficient:	0.853987
22.48	0	0	25.53	0	0	99.34386266		
23.12	0	0	25.58	0	0.0001	99.34386266		
23.21	0	0	25.67	0	0.0003	99.34386266	NSCE Numerator:	3170080
23	0	0	25.72	0	0.0001	99.34386266	NSCE Denominator:	6209129

Model Evaluation Indices

How to calculate NSCE

- Now we'll put it all together

$$NSCE = 1 - \frac{\sum_{i=1}^{n}\left(R_{obs,i} - R_{sim,i}\right)^2}{\sum_{i=1}^{n}\left(R_{obs,i} - \overline{R_{obs}}\right)^2}$$

- Take 1 minus the numerator over the denominator

fx =1-(K11/K12)

Formula Bar

D	E		G	H	I	J	K
Precip(mr	PET(mm h	SM(%)	Fast Flow(	Slow Flow	Average of observations		
0	0	25.31	0	0.0001	99.34386266	Sum of Observed Values:	69441.36
0	0	25.41	0	0.0003	99.34386266	Sum of Simulated Values:	91544.53
0	0	25.46	0	0.0001	99.34386266	Subtract:	22103.17
0	0	25.48	0	0	99.34386266	Divide:	0.3183
0	0	25.49	0	0	99.34386266	Multiply by 100:	31.82998
0	0	25.53	0	0.0001	99.34386266		
0	0	25.53	0	0	99.34386266	Correlation Coefficient:	0.853987
0	0	25.53	0	0	99.34386266		
0	0	25.58	0	0.0001	99.34386266		
0	0	25.67	0	0.0003	99.34386266	NSCE Numerator:	3170080
0	0	25.72	0	0.0001	99.34386266	NSCE Denominator:	6209129
0	0	25.76	0	0.0001	99.34386266	NSCE:	0.489449

Back to Our Question…

Was the model output "good"?

The correlation coefficient (0.85) is very good

Remember this is a 0-1 scale

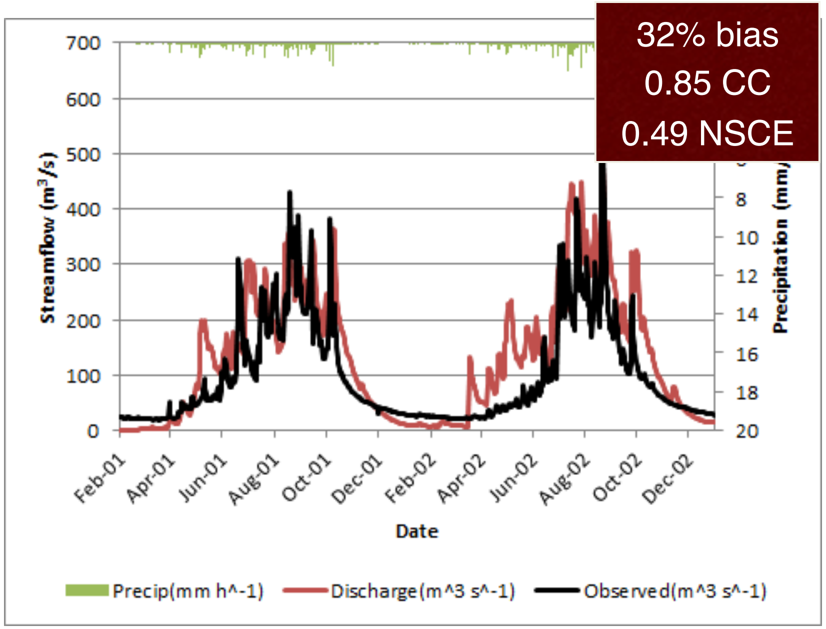
The bias (32%) is definitely higher than we'd prefer

We want values close to 0

The NSCE (0.49) is adequate

A really well-calibrated model should have a NSCE of 0.7 or more, if we are lucky

The verdict: more calibration could be definitely done, but this is a pretty good hydrograph



Coming Up….

The next module is

DEM Derivatives

You can find it in your `\EF5_training\presentations` **directory**

Module 2.1 References

EF5 v0.2 Readme, (March 2015).

EF5 Training Doc 2 – Hydrological Model Evaluation, (March 2015).

Nash, J. E. and J. V. Sutcliffe, (1970). River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology*, 10: 3, 282-290. (Paper – NSCE.pdf)