Ensemble Forecasting

CBRFC
June 26-27, 2002

Ensemble Forecasting Workshop

Outline

- Ensemble Forecasting for Hydrology
  - What is an ensemble forecast?
  - Why do we do ensemble forecasting?
  - How do we generate ensembles?
  - What do we do with the ensemble once we have it?
- Review of Statistics for Ensemble Forecasting
  - General Review of Statistics
  - Getting a Sample from an Ensemble of Time Series
- Exercise

A Definition of Ensemble Forecasting

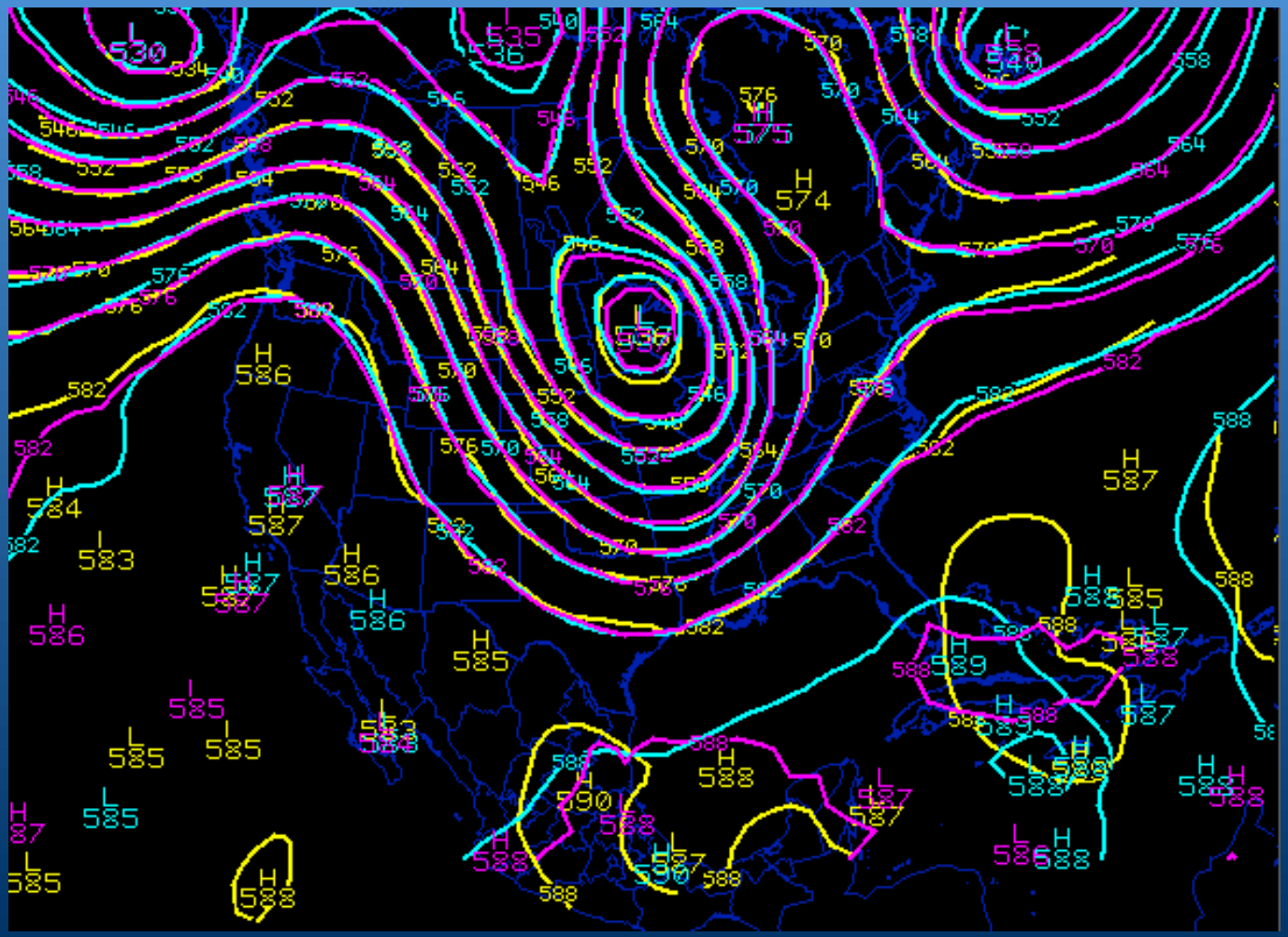
From www.hpc.ncep.noaa.gov/ensembletraining/

An ensemble forecast is a collection of two or more forecasts that verify at the same time.

A Met Ensemble

500 mb Heights from the 12z Cycle May 22, 2001 from the AVN (blue), Eta (Yellow), andNGM (purple).

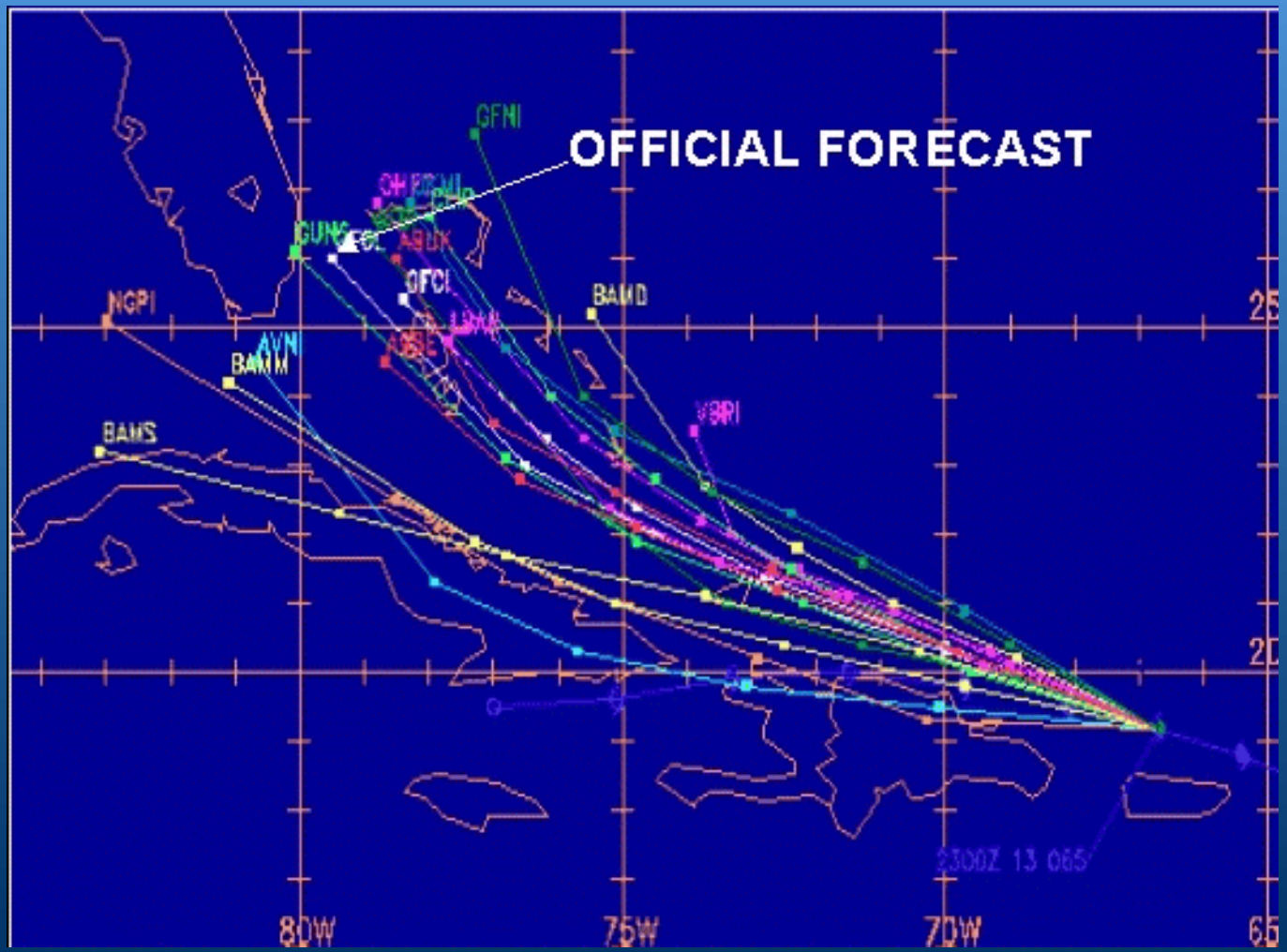
Image taken from http://www.hpc.ncep.noaa.gov/ensembletraining/



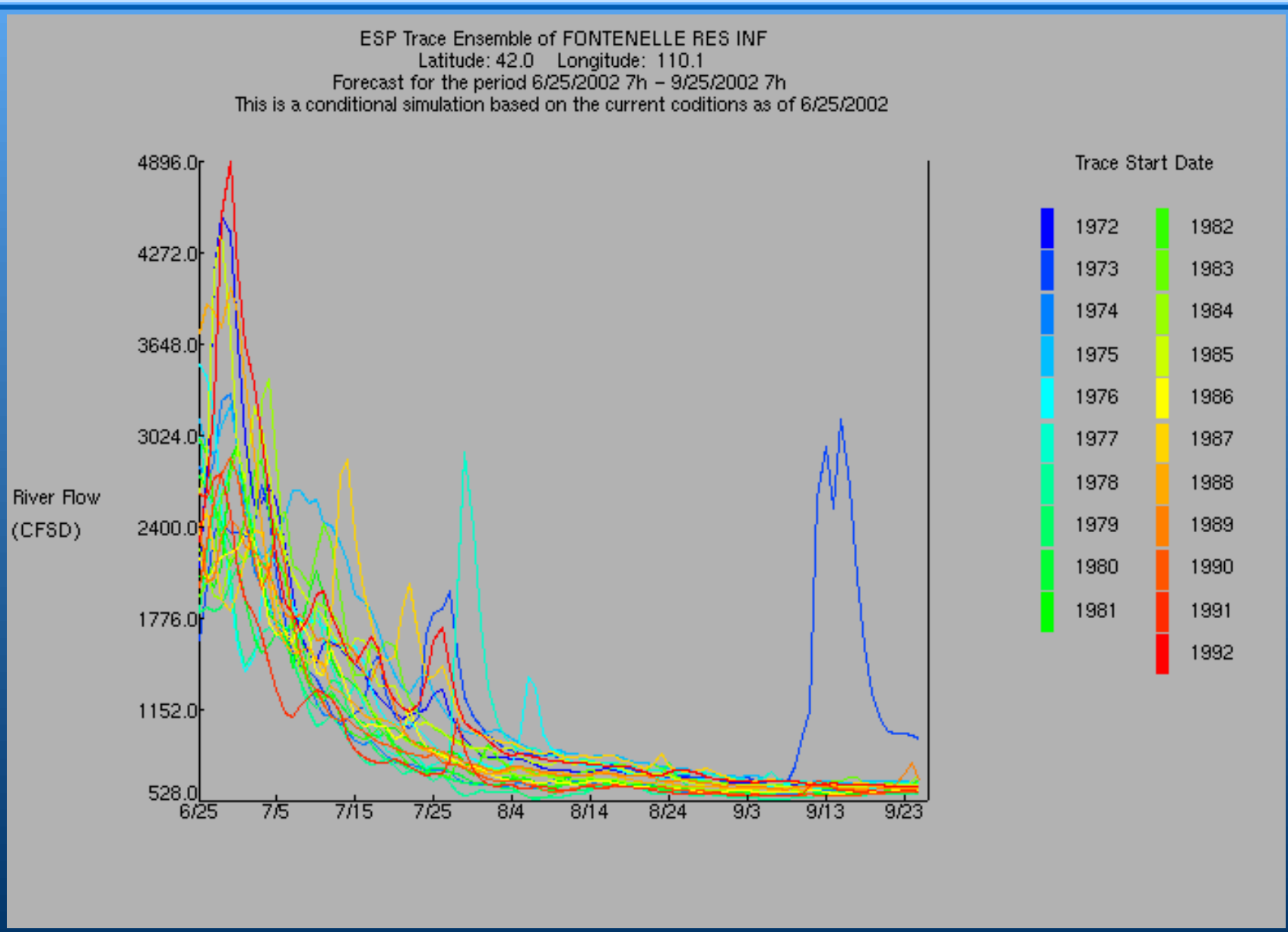
Huricane Tracks

Hurricane Debby
Track Model Guidance, 8/23/00 00z

Image taken from http://www.hpc.ncep.noaa.gov/ensembletraining/



Spaghetti Plot of Flows



Other Examples of Ensemble Forecasts

- Ad hoc scenarios like qpf/non-qpf runs
- Varying the QPF with mods in IFP
- Observed climate streamflows
- Agencies that use both NWS forecasts and their own models
- Review of like cases

Why do we do Ensemble Forecasting?

UNCERTAINTY

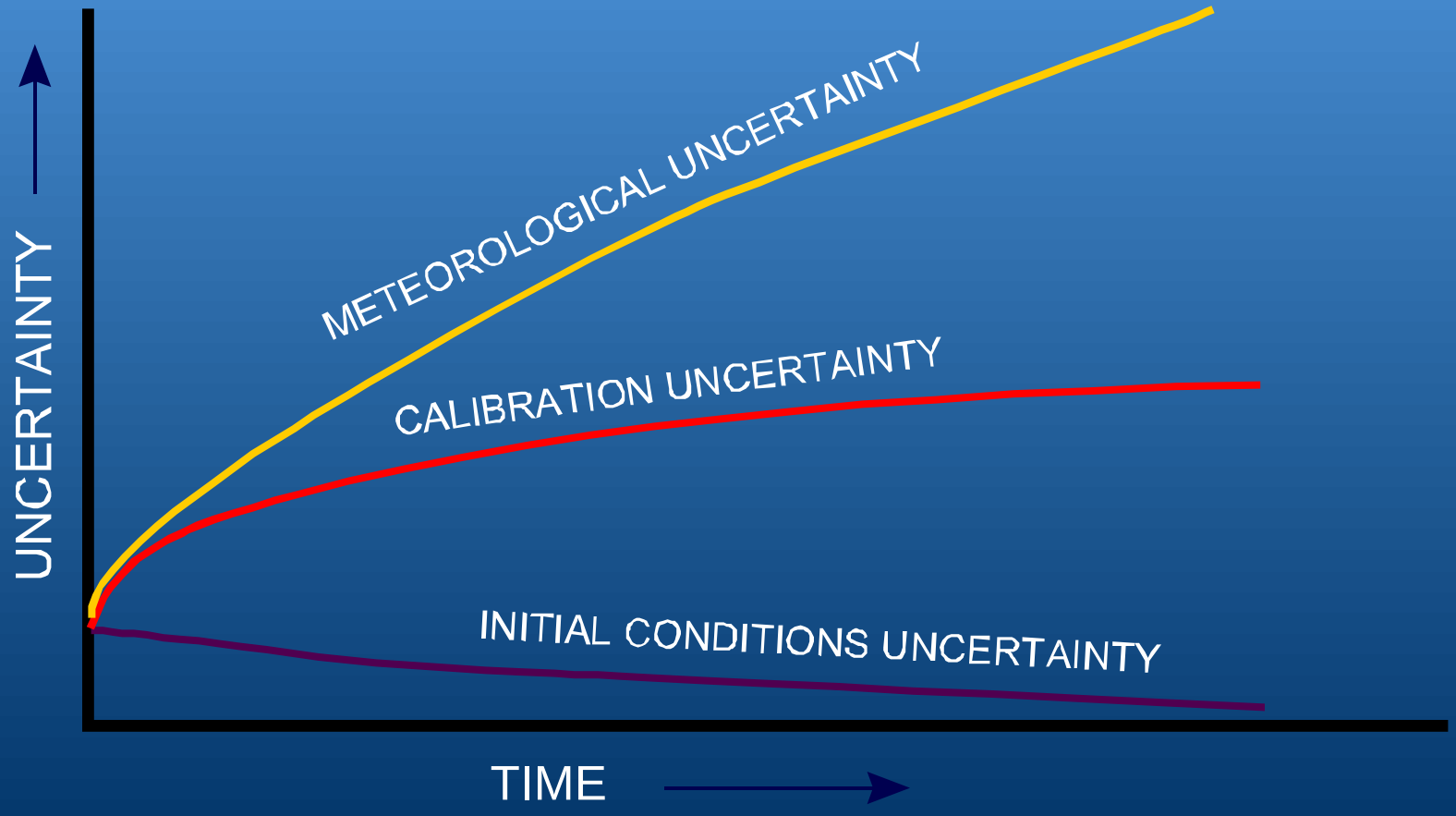
There are many sources of uncertatinty. The goal of ensemble forecasting is to quantify the forecast uncertatinty in an objecive manner.

Sources of Uncertainty @ CBRFC

- Observations
- Model initial conditions
- Input forecasts
- Model structure
- Model parameters
- Rating curves

Why do we do Ensemble Forecasting?

Major Sources of Uncertainty in Hydrologic Forecasts



Are Ensembles the only way?

Other Ways of Assessing Uncertainty

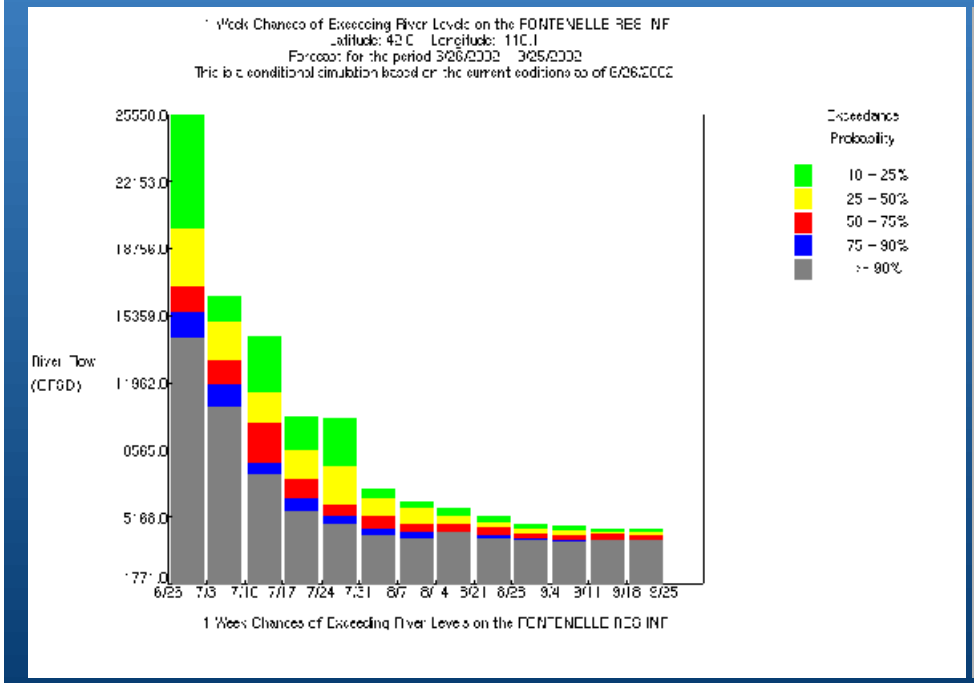
- Statistical techniques like Kalman filtering and error propagation
- Personal experience/opinion
- Modeling the error itself
- Review of verification statistics (conditional distributions especially)

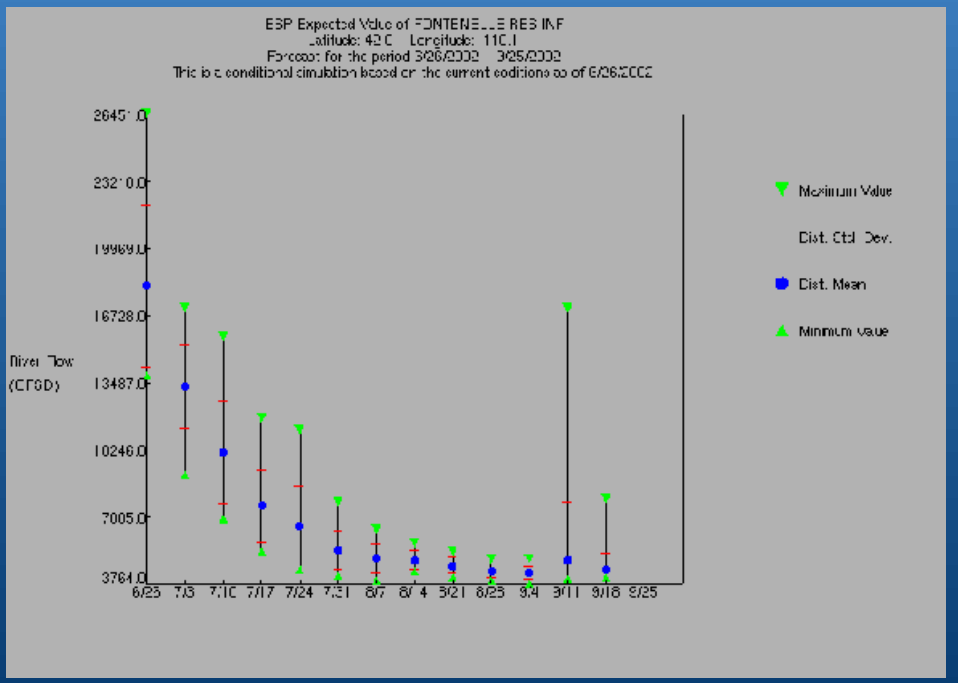
How Do We Generate Ensembles?

An ensemble method for each source of uncertatinty above

- Model structure – run two models
- Model parameters – run the same model with different parameter sets
- Model initial conditions including input observations – start the model different initial conditions (that's what the mets do)
- Input forecasts – vary the forecast input to the model (have distinguished between running the model up to the end of the observed period and the forecast period)

What do we do with the Ensembles?





What do we do with the Ensembles?

Example Ideas

- Pick one member you like the best
- Draw by hand what seems most likely to you
- Compute a mean hydrograph
- Pick out a range for some event
- Pass the ensemble on for input to another model
- Compute a distribution for a particular period or several periods
- We can draw pictures; we can make text statements

Introduction to Statistics for Ensemble Forecasting

More than anything else, AHPS is the application of statistical science to hydrology.

Probability

- "AHPS is probably about probability" - Dave Brandon, circa 2002
- In general, the probability of an event is the number of favorable outcomes divided by the total number of possible outcomes.
- e.g. the probability of drawing the queen of hearts from a deck of cards is 1/52 or 1.9%.

Weighted Probability

- What is the expectation (average value) resulting from tossing a pair of dice? (Prime totals excluded)

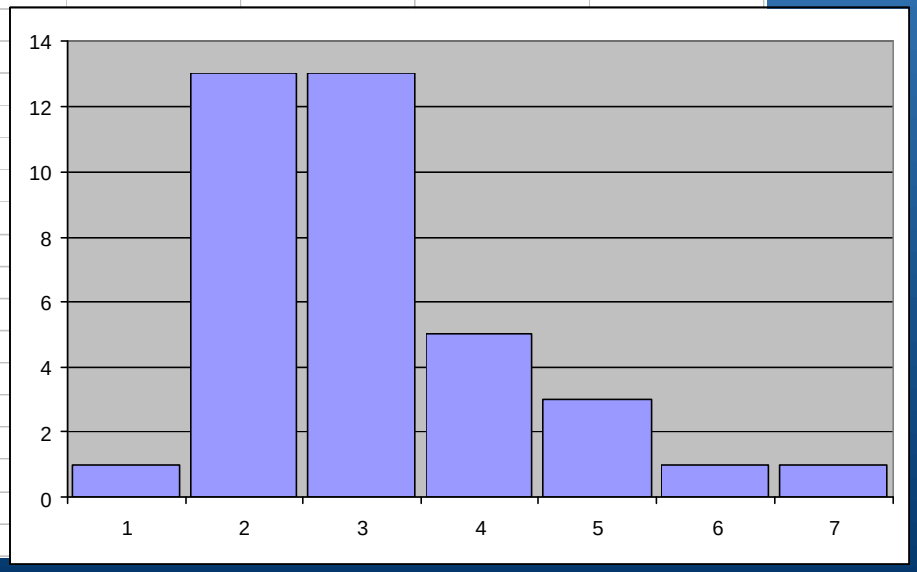
	Value	Possible Comb	Total Comb	Probability	Weight*Value	
	4	3	21	0.14	0.57	
	6	5	21	0.24	1.43	
	8	5	21	0.24	1.90	
	9	4	21	0.19	1.71	
	10	3	21	0.14	1.43	
	12	1	21	0.05	0.57	
Expectation	8.17				7.62	

Frequency Histogram

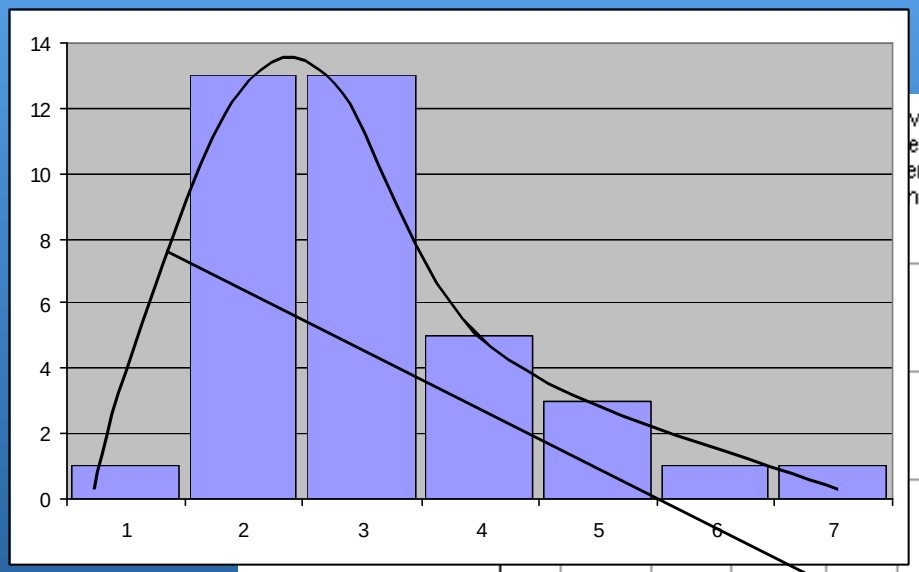
Apr-Jul Volumes Roaring Fk. Nr. Aspen

1977	19.47
1981	30.65
1976	30.68
1972	32.94
1989	34.29
1994	34.37
1967	35.09
1988	35.3
2000	35.46
1975	36.68
1998	36.83
1966	37.53
1974	38.92
1992	39.92
1969	40.3
1990	40.75
1968	42.87
1978	44.69
1982	46.47
1991	46.92
1986	48.86
1973	51.05
1971	51.61
2001	51.67

0-20	20-40	40-60	60-80	80-100	100-120
1	13	13	5	3	1



Cumulative Distribution Function



ver Levels on the ROARING FK – ASPEN at
e: 39.2 Longitude: 106.8
eriod 5/29/2002 24h – 5/30/2002 24h
n based on the current coditions as of 5/30/2002

River Flow
(CFSD)

788.0
591.0
394.0
197.0
0.0

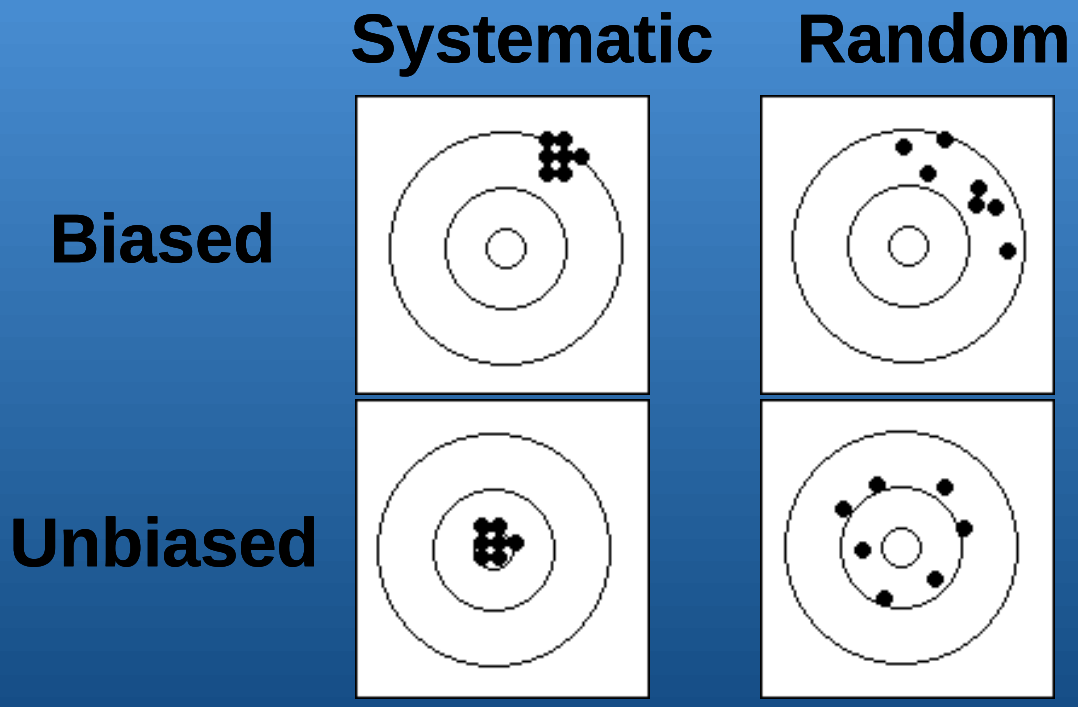
99% 98% 95% 90% 80% 70% 60% 50% 40% 30% 20% 10% 5% 2% 1%

Chance of Exceedance

CS
HS
OBS

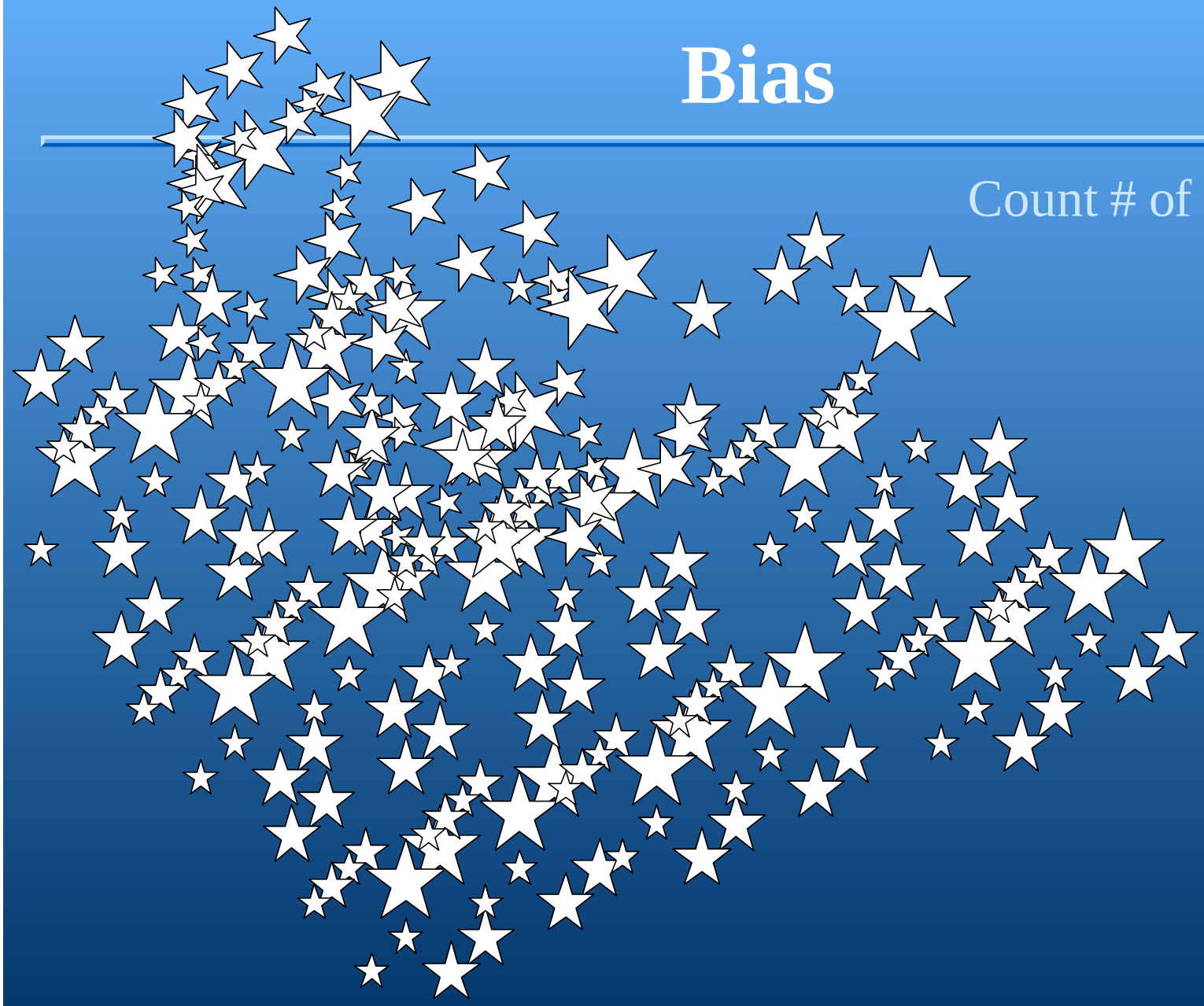
Bias

Two Kinds

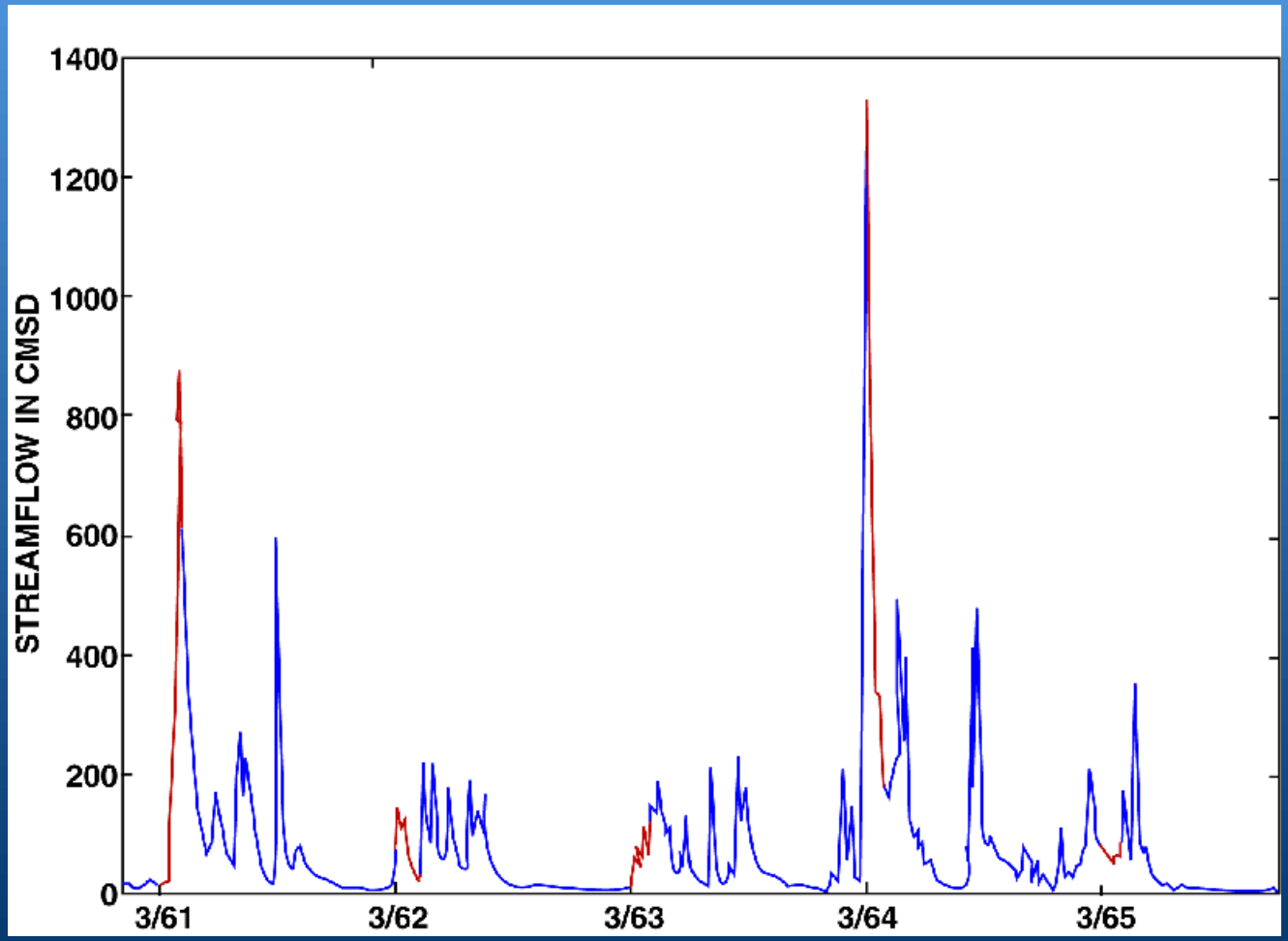


Bias

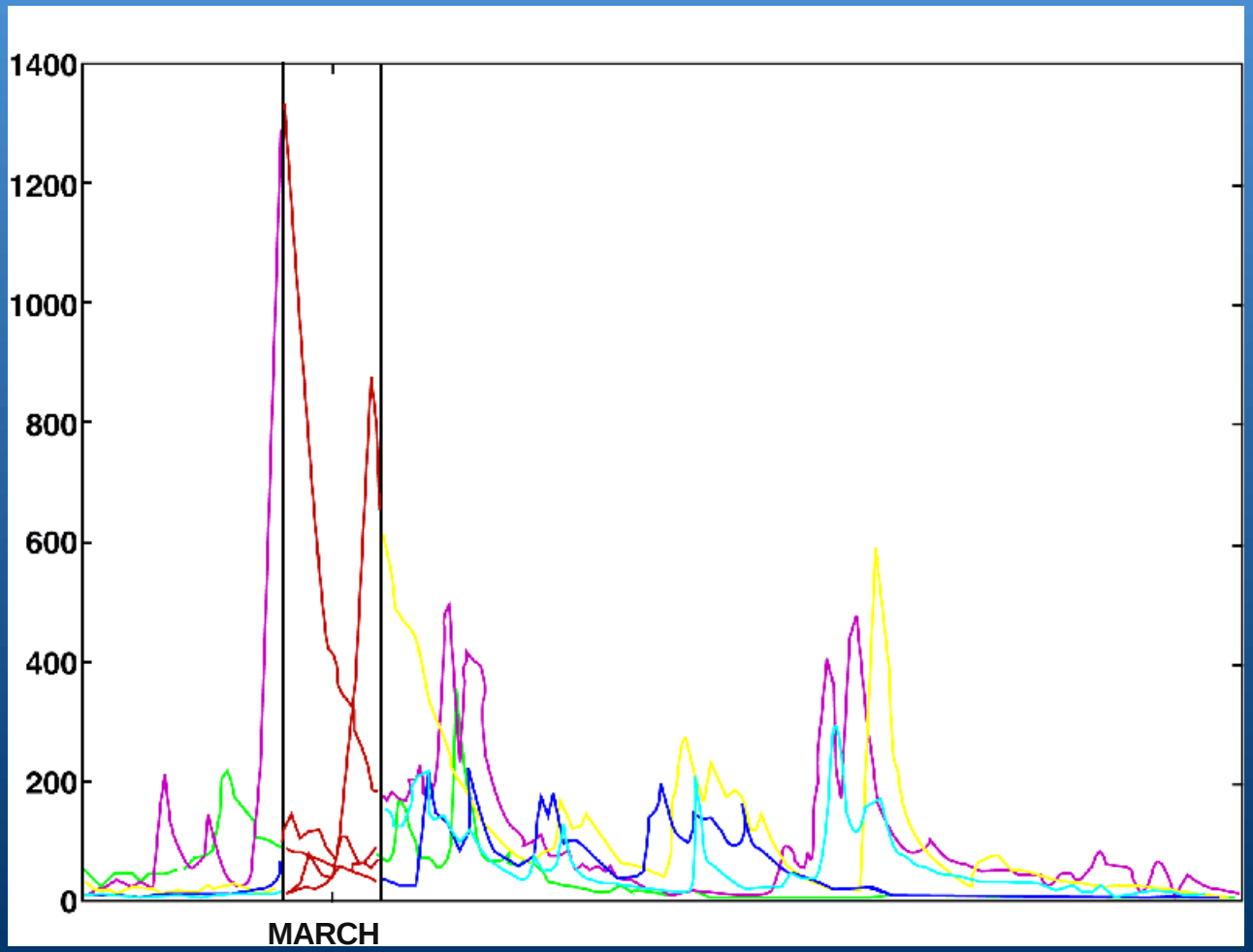
Count # of Stars – !!



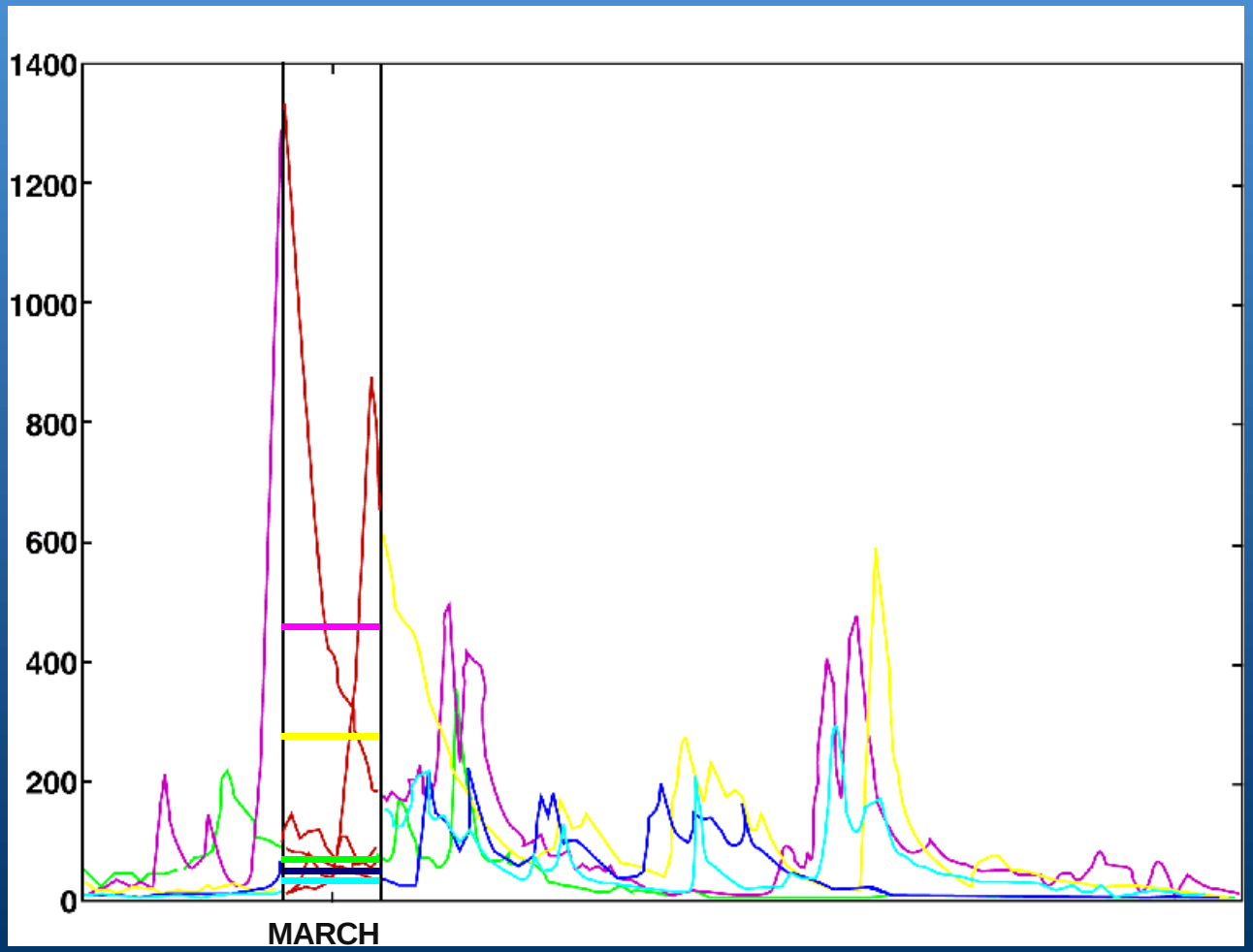
FORECAST POINT SAYI4



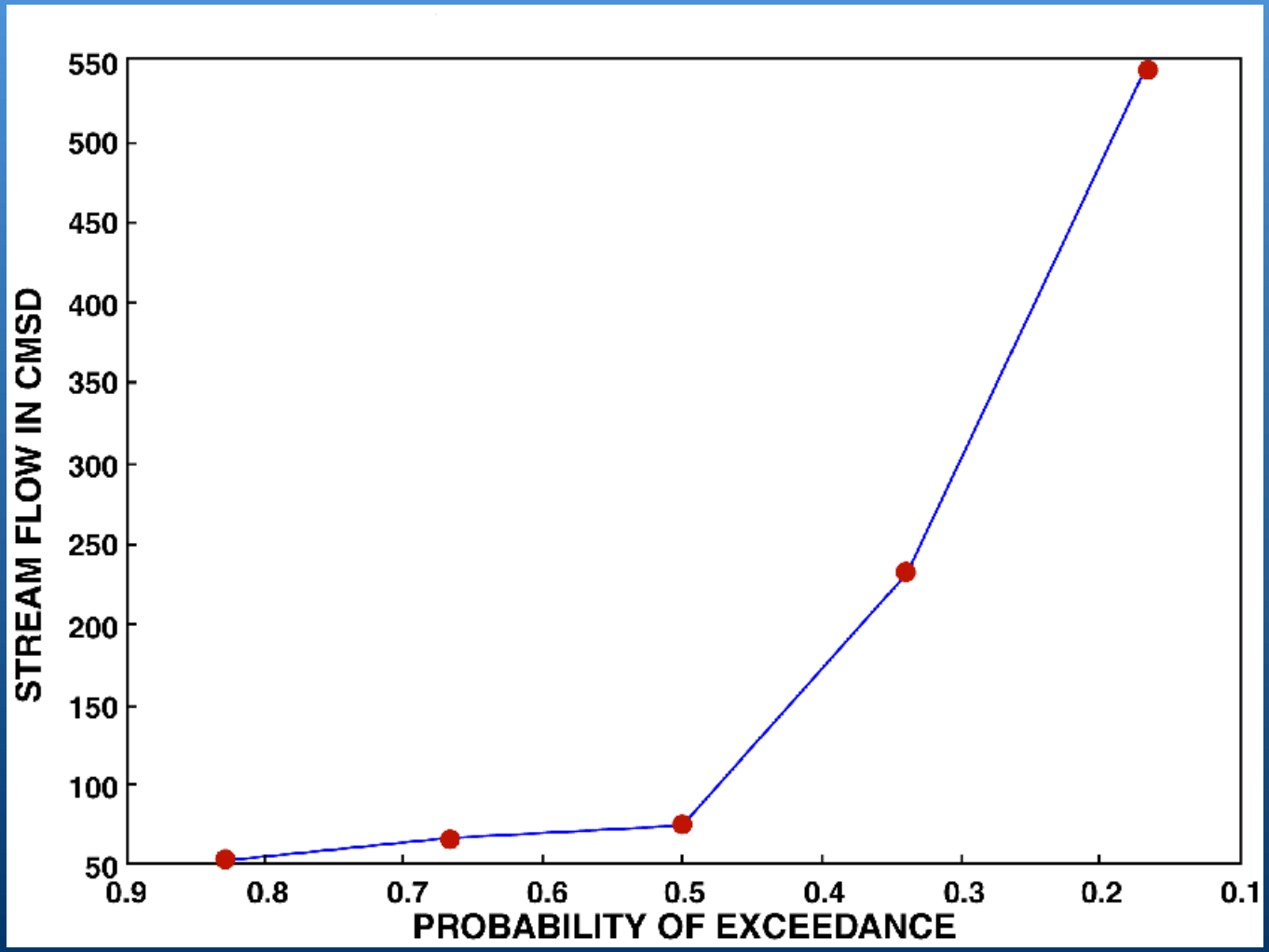
SPAGHETTI PLOT FOR FORECAST POINT SAYI4



SPAGHETTI PLOT FOR FORECAST POINT SAYI4



MEAN MONTHLY FLOWS AT SAYI4



Important Terms

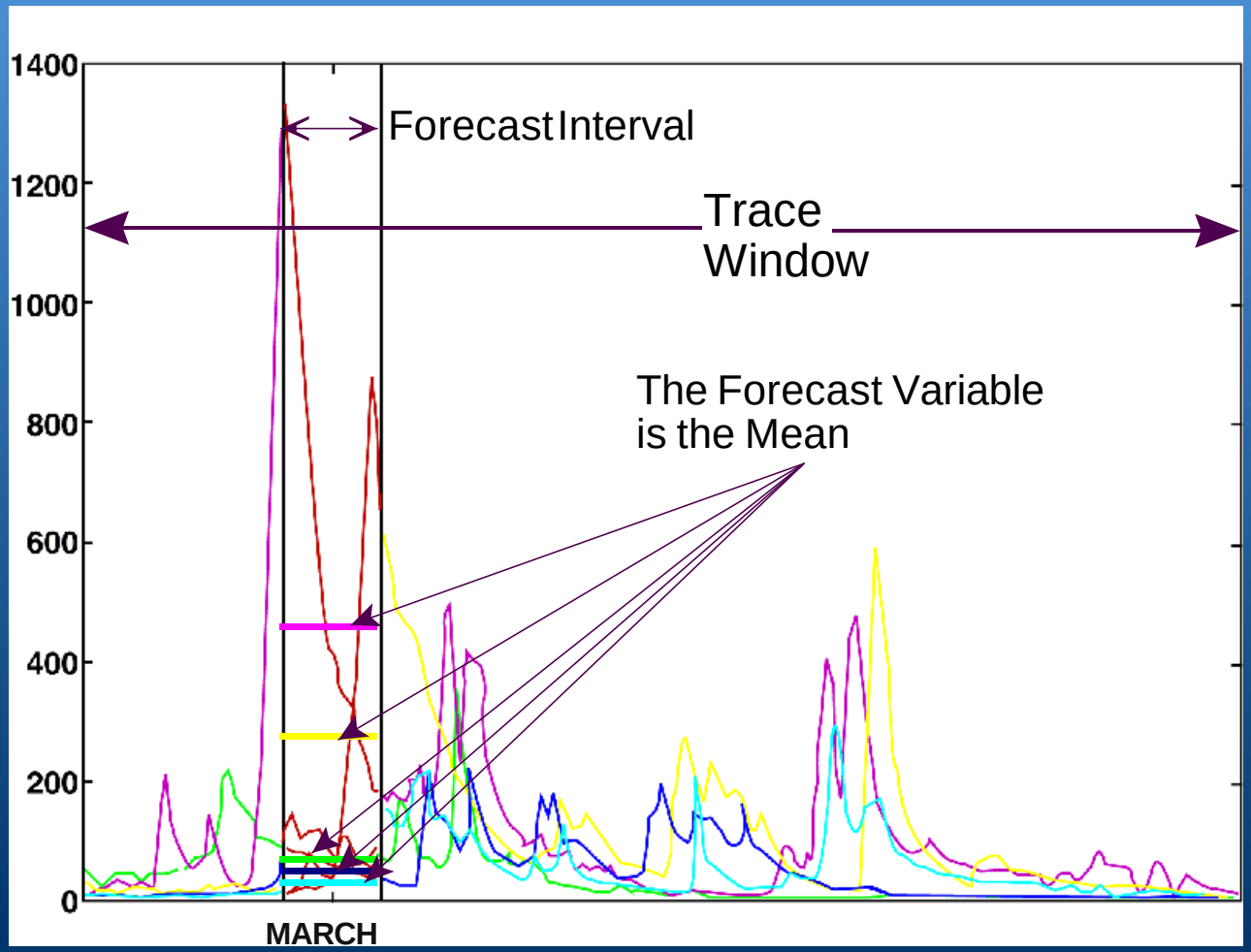
- Ensemble – A set of time series that represent possible outcomes.
- Trace – One time series in an ensemble.
- Distribution – A function that describes the likelihood that some set of events will occur.
- Sample Set – A set of values that represents a distribution. The values may be observations or forecasts.

Important Terms

(cont)

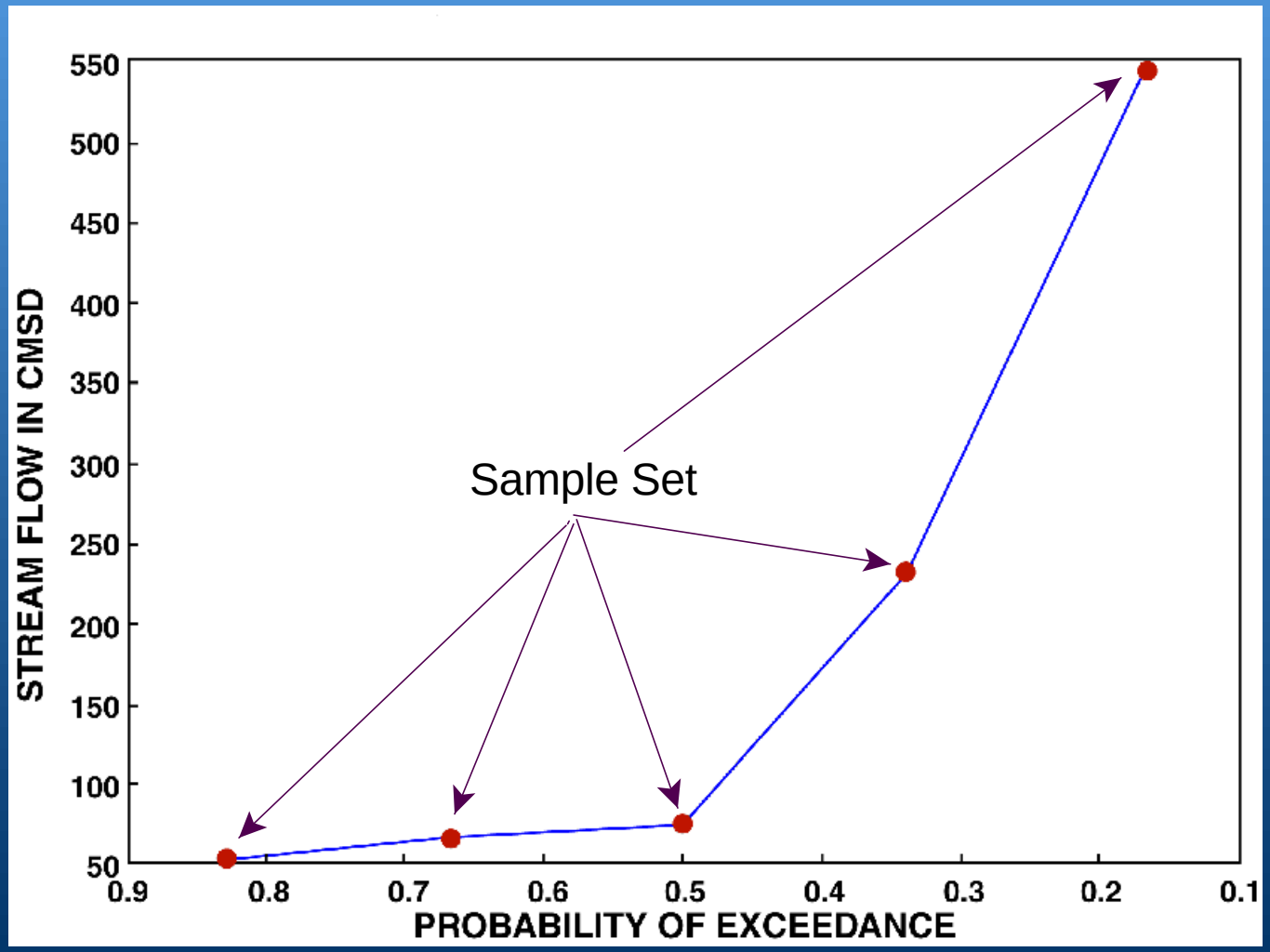
- Forecast Variable – The type of information to be extracted from an ensemble, maximum, minimum, mean, etc.
- Trace Window – The length of the forecast (ESP) run (also called analysis window).
- Forecast Interval – The time period over which the Forecast Variable is to be sampled. Not necessarily equal to the Trace Window. Multiple Forecasts may be extracted from a single ensemble. Also called the Forecast Window.

SPAGHETTI PLOT FOR FORECAST POINT SAYI4



Forecast Variable

Forecast Interval

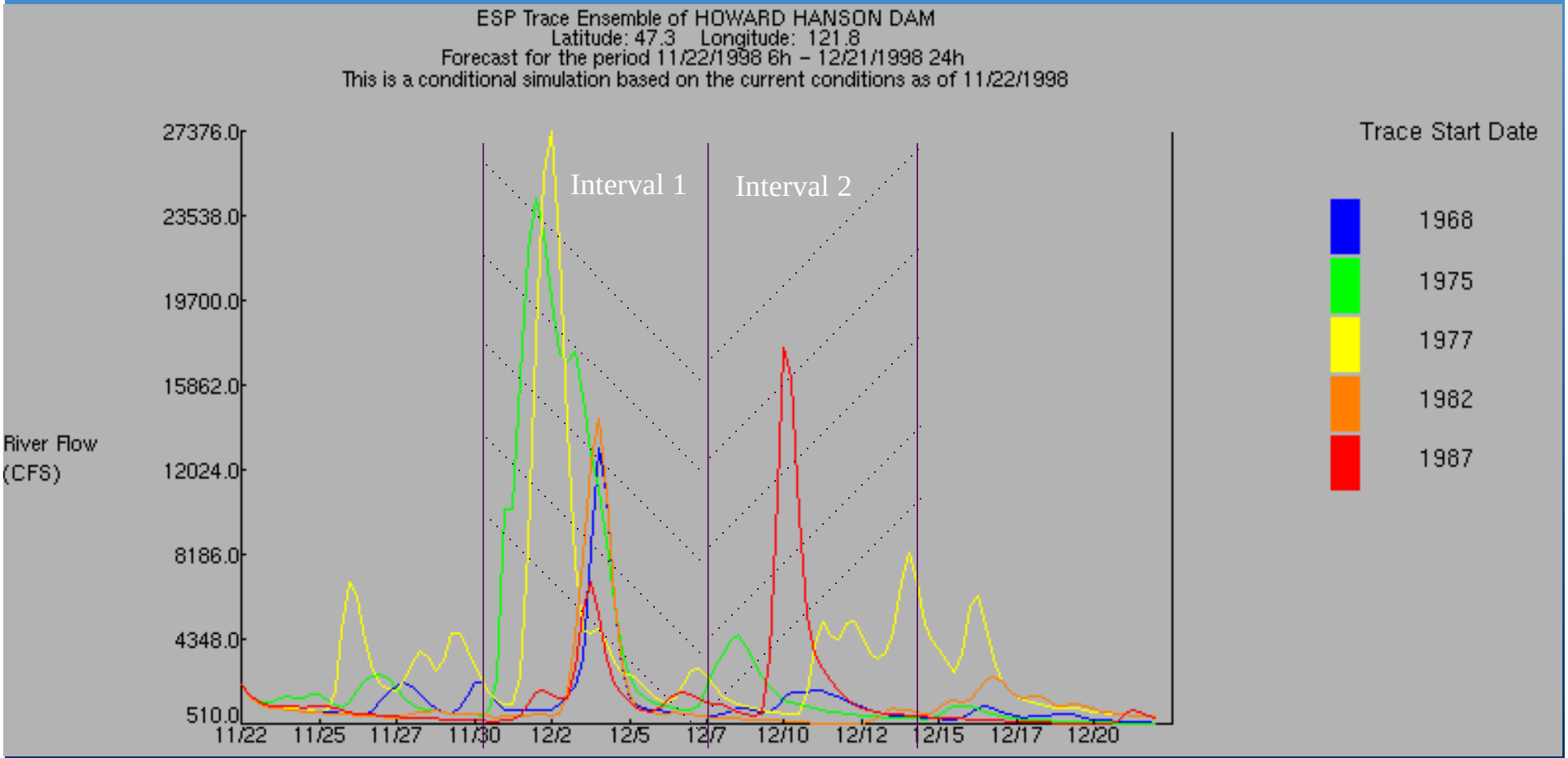


Steps for Deriving a Distribution

- Select an Interval and a Variable.
- Extract a SampleSet from the Ensemble.
- Fit a Distribution to the Sample Set.
  - For the empriical distribution: sort,rank and use plotting position (n/N+1).
  - For normal and others extract distribution parameters from the sample set.
  - For wakeby and others use numerical fitting algorithm.

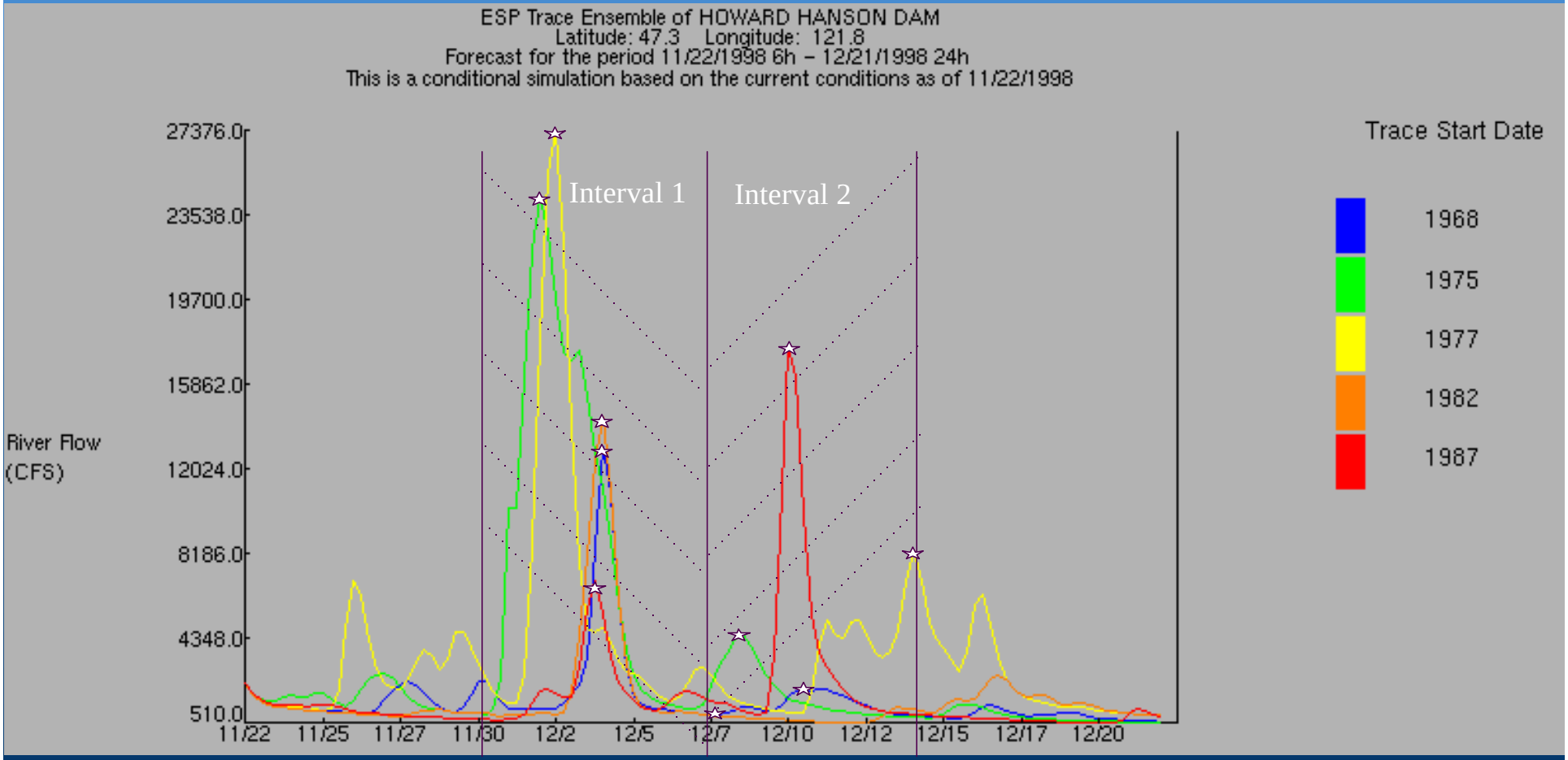
Collecting a Sample

Intervals



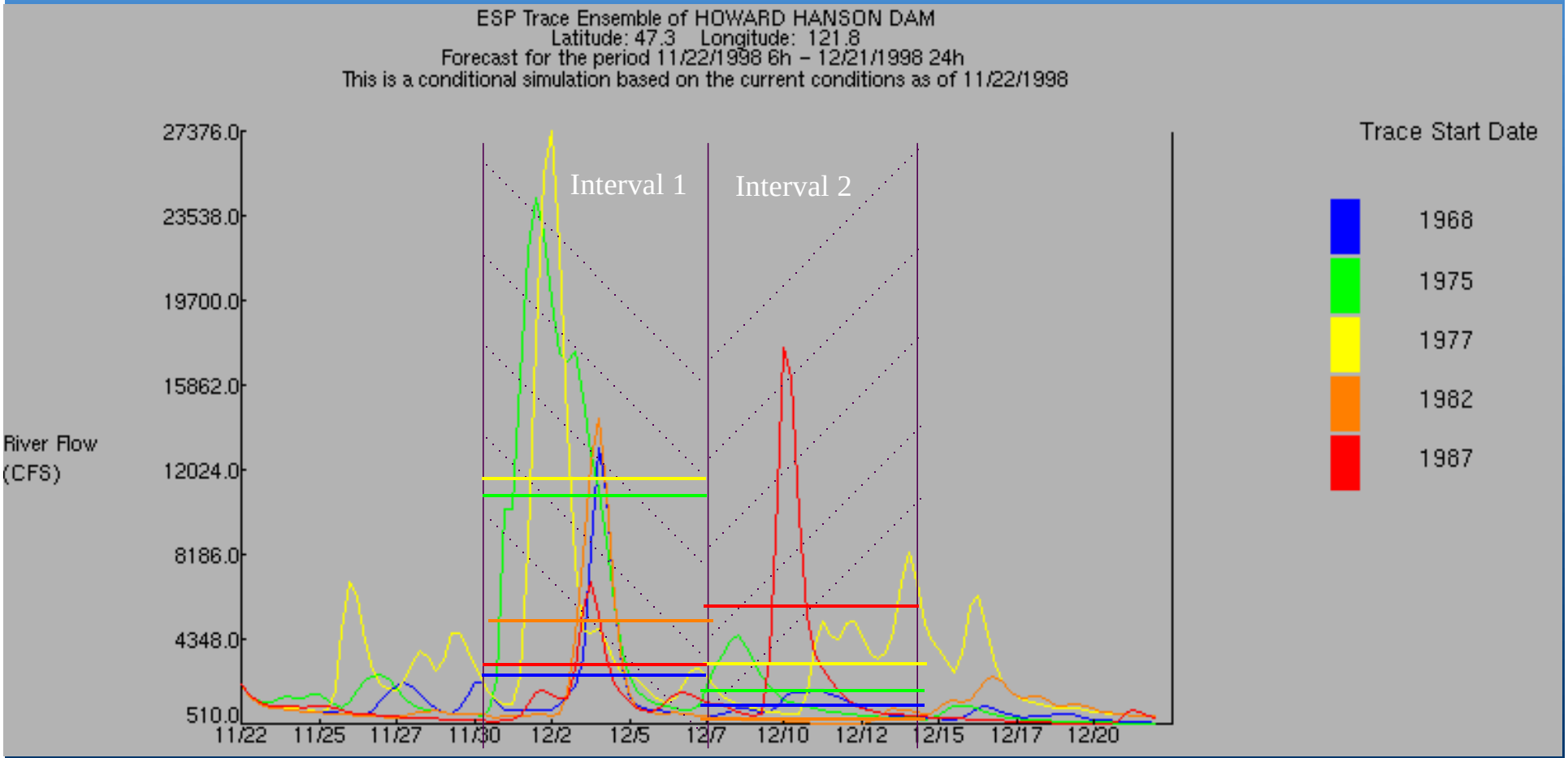
Collecting a Sample

Variable is max



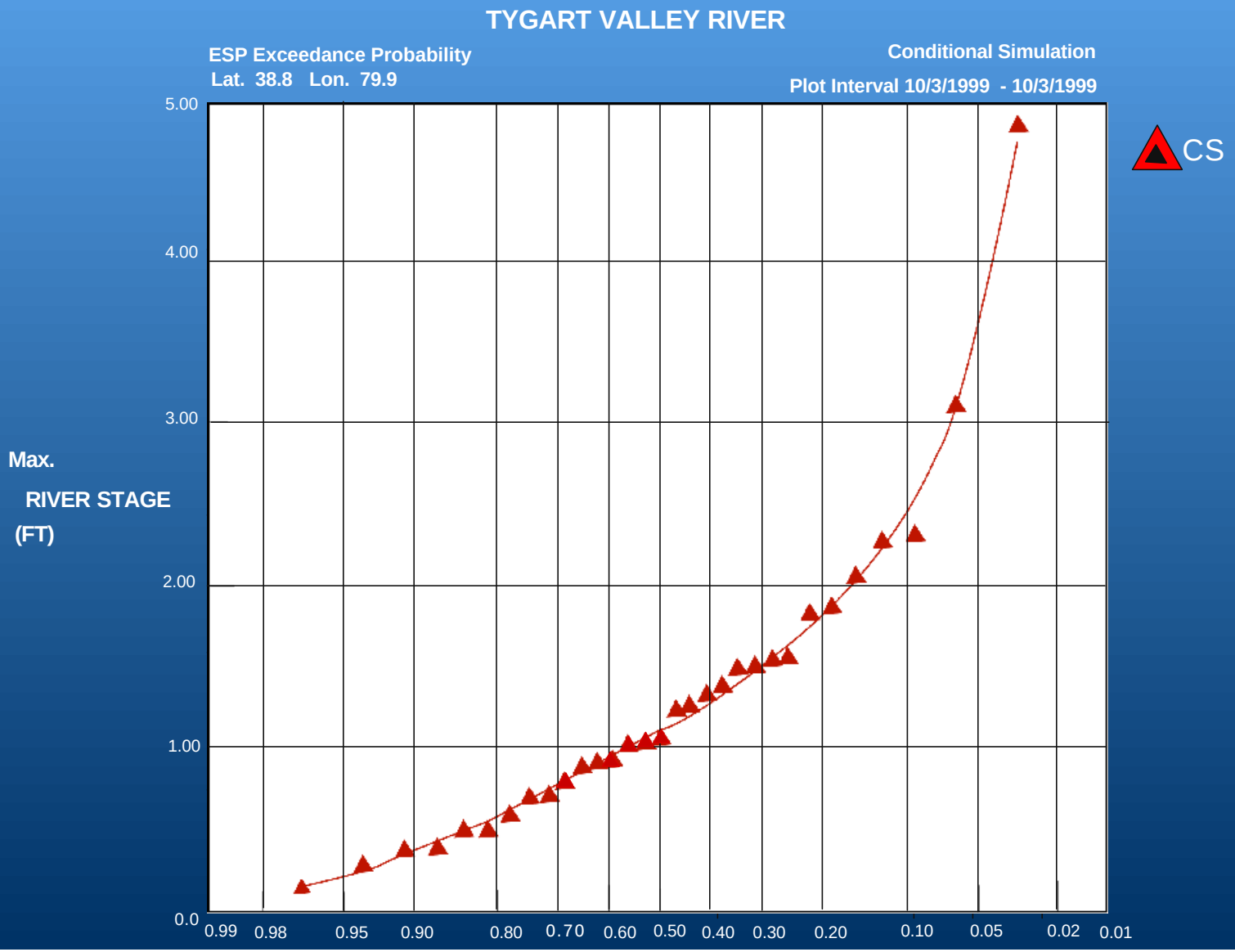
Collecting a Sample

Variable is Volume



Last Step: Fit a Distribution

Sometimes an Analytic Distibution Works



Last Step: Fit a Distribution

And sometimes it doesn't.

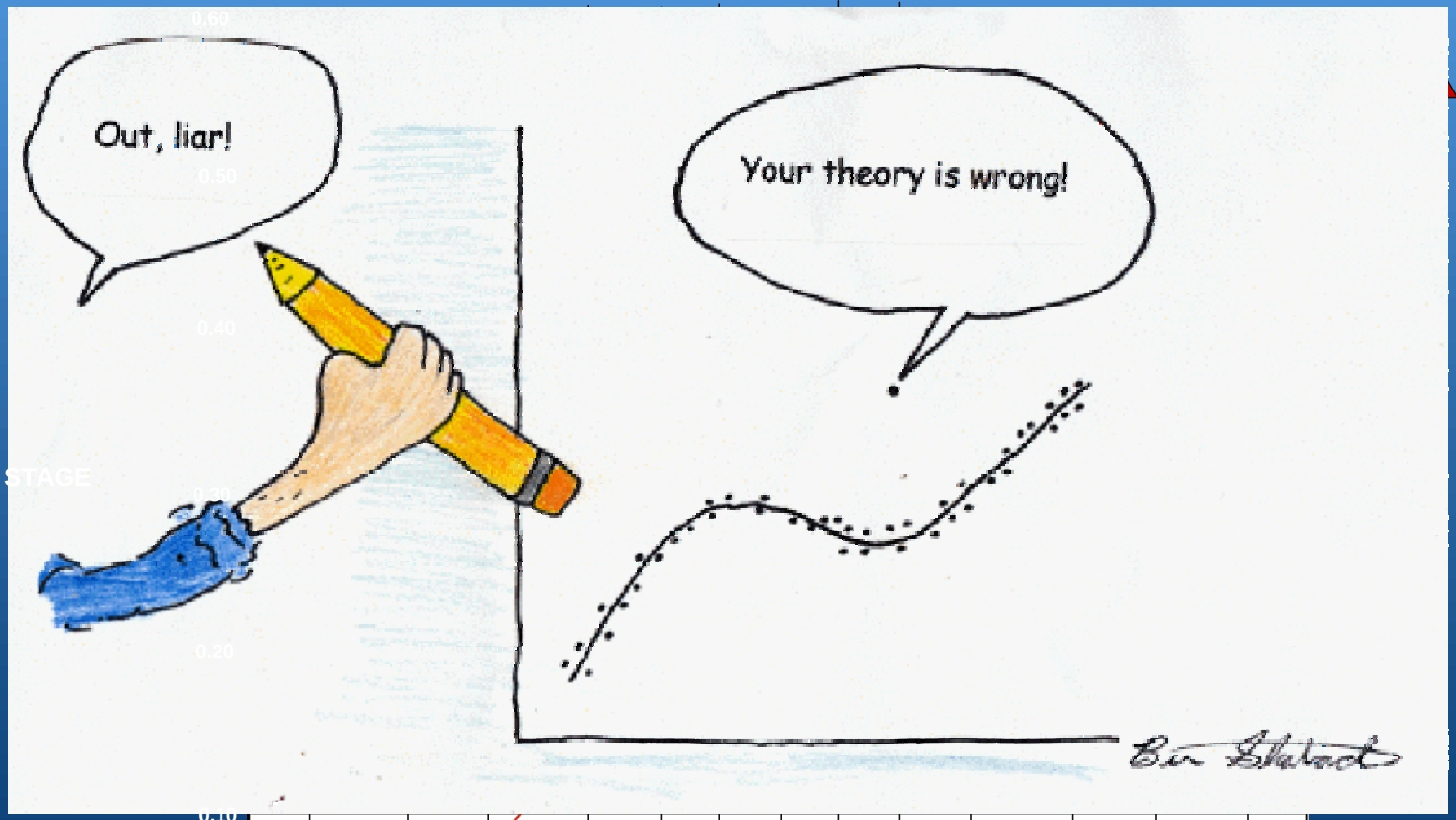
TYGART VALLEY RIVER

ESP Exceedance Probability
Lat. 38.8 Lon. 79.9

Conditional Simulation
Plot Interval 10/3/1999 18 -10/3/1999 24 INTL

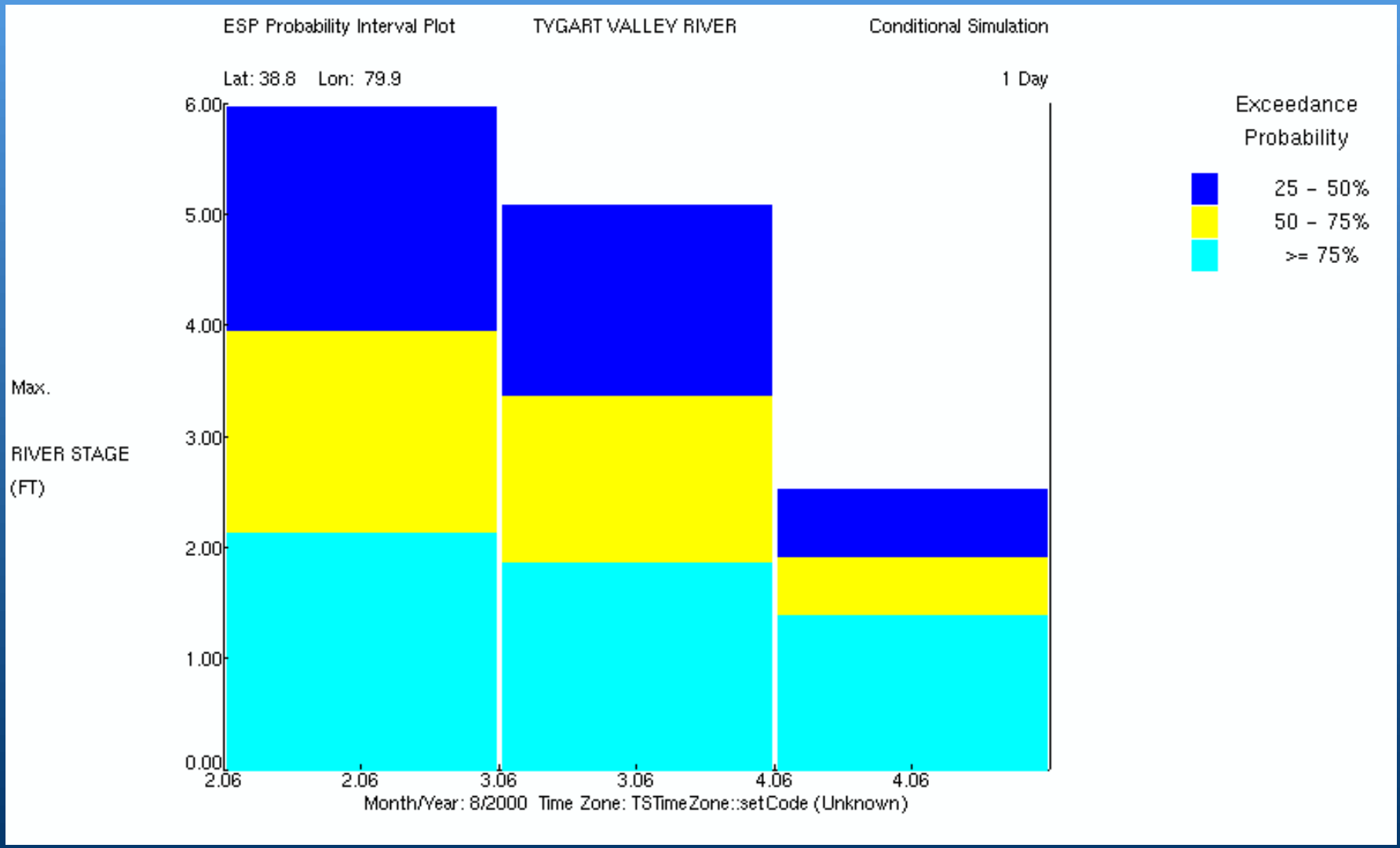
CS

RIVER STAGE
(FT)



0.60
0.50
0.40
0.20
0.10
0.0

0.99 0.98 0.95 0.90 0.80 0.70 0.60 0.50 0.40 0.30 0.20 0.10 0.05 0.02 0.01


ESP Probability Interval Plot
TYGART VALLEY RIVER
Conditional Simulation
Lat: 38.8 Lon: 79.9
1 Day
Exceedance Probability
25 - 50%
50 - 75%
>= 75%
Max.
RIVER STAGE
(FT)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
2.06
2.06
3.06
3.06
4.06
4.06
Month/Year: 8/2000 Time Zone: TSTimeZone::setCode (Unknown)

And Now an Exercise

Individual Exercise for NWRFC Ensemble Forecasting Workshop
Exercise 1

This is an exercise to help **you** identify what you understand and what **you** do not understand of the material we have covered. If **you** do not understand something, please ask for an explanation. In order that **you** determine if **you** understand the material, please work individually on this exercise.

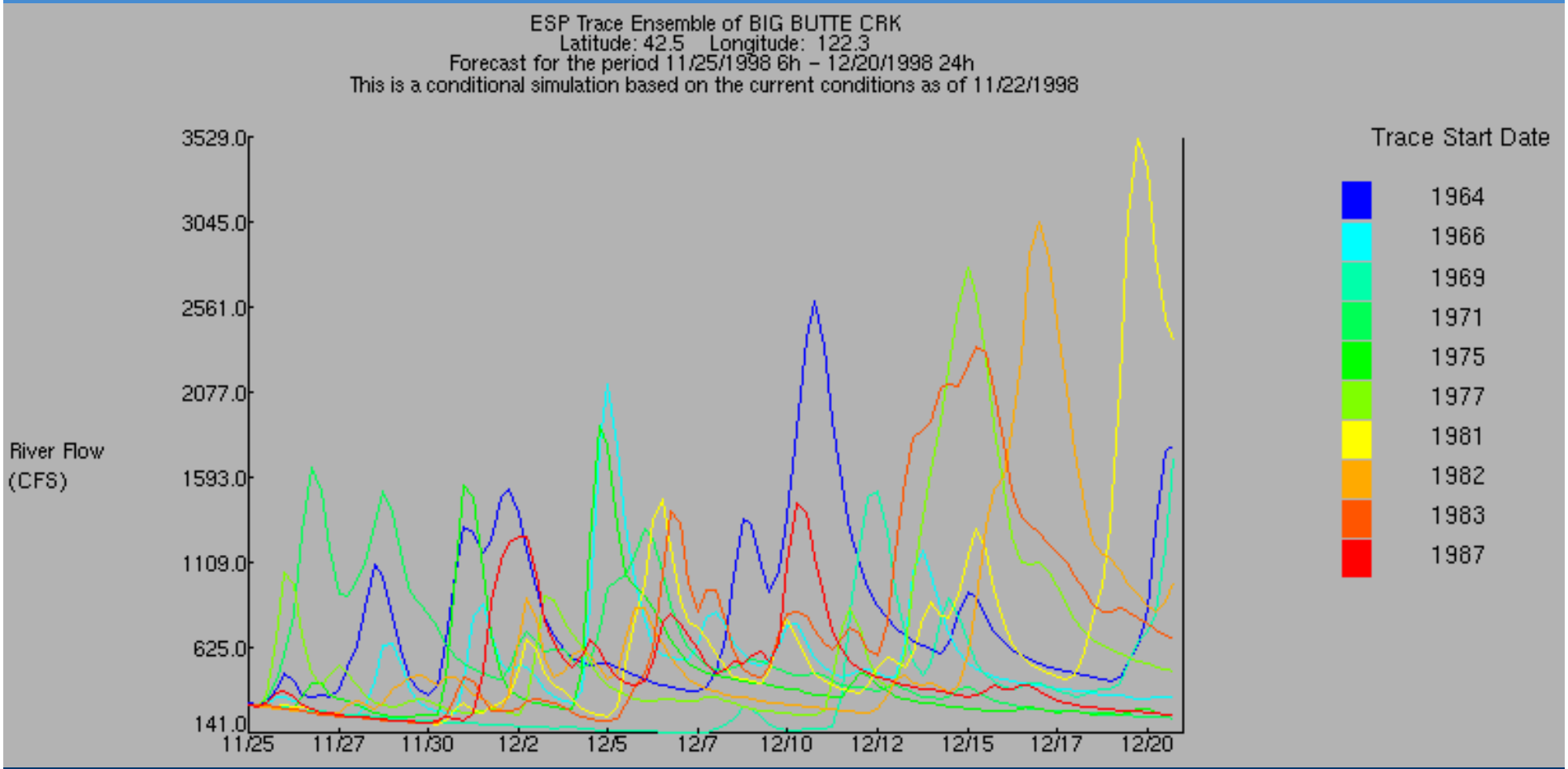
You have been given an ensemble spaghetti plot.

Your mission is to:

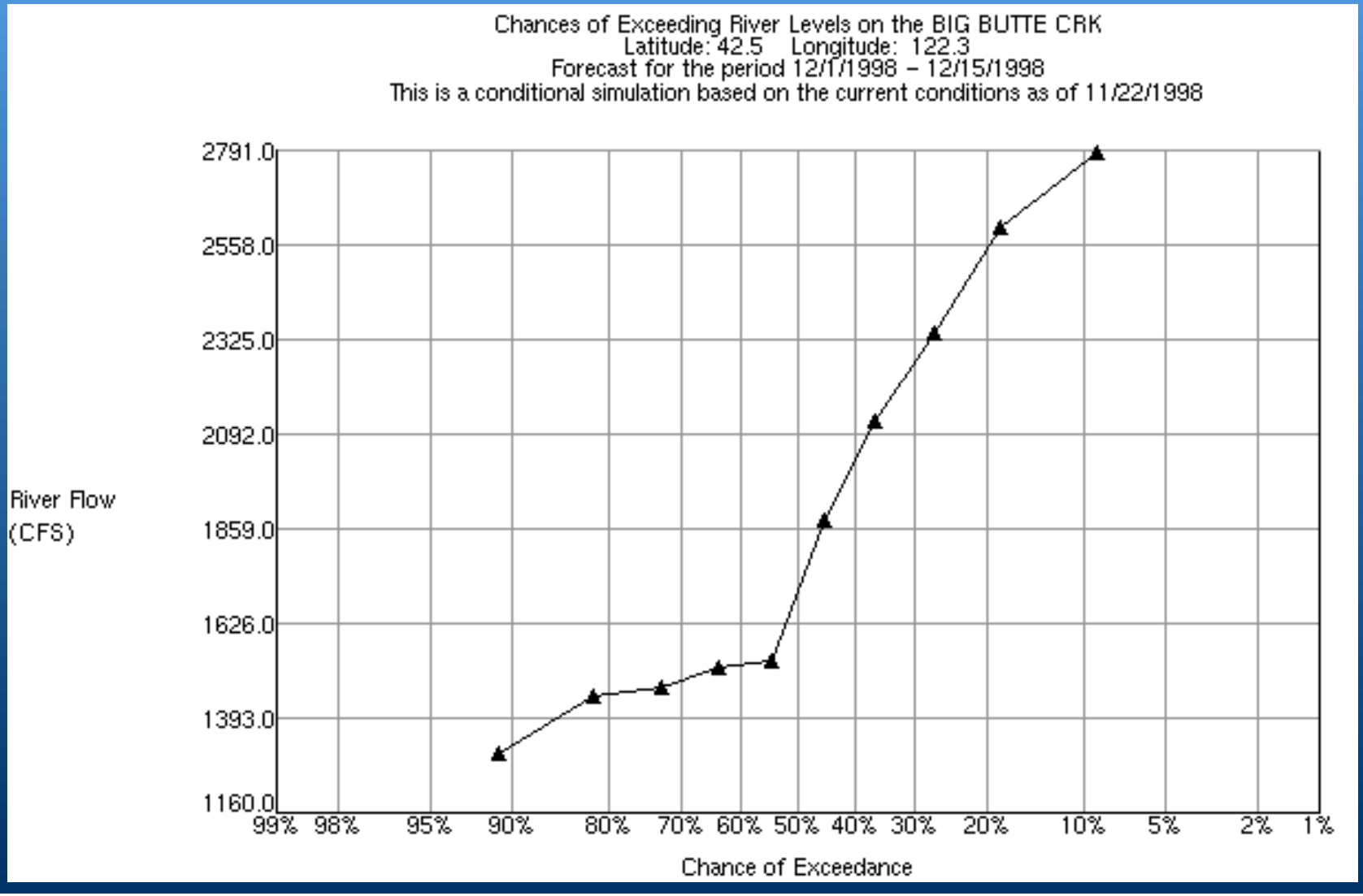
1) extract from the spaghetti plot the forecast distribution of maximum flows for the period 12/1 to 12/15;
2) create a product that you think would communicate to users the likelihood of flooding at this point;
3) identify the interval and the variable for this forecast.

When everyone has produced a product, we will attach them to the board in the front of the room for display/discussion.

Spaghetti Plot for Exercise 1



Exercise 1 ESPADP Forecast



Exercise 1 ESPADP Table

exr_fcst_tbl.txt

Current NWS ESP Method

ESP Trace Generation