Issues in Uncertainty Estimation for Hydrologic Modeling

(Use of Ensemble Approaches)

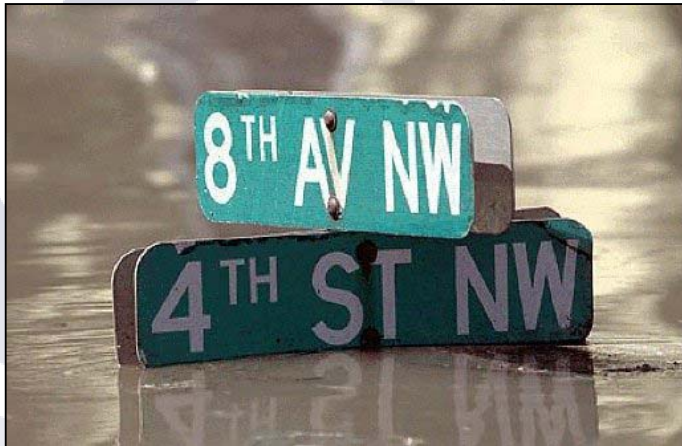
Hoshin Gupta
Adjunct Professor
Hydrology & Water Resources
The University of Arizona
Tucson, Arizona

Desirable Characteristics of a Hydrologic Model …

- State/Output Predictions are "*Accurate*" (unbiased)
- State/Output Predictions are "*Precise*" (minimal uncertainty)
- Input-State-Output behavior is "*Consistent*" with the available data
- Conceptual structure is "*Consistent*" with our perceptions (understanding) of the physical/behavioral structure of the system

… Operations Point of View

NWS/OHD/HL Hydrologic Modeling Priorities



… Verification of deterministic and probabilistic river forecasts.

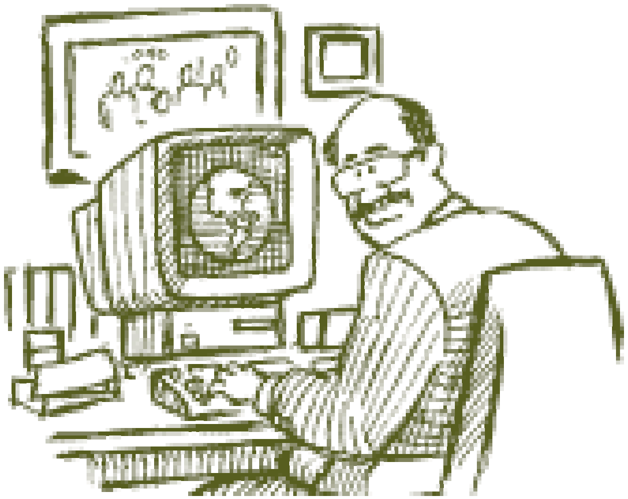
… Quantification of uncertainty in river forecasts including ensemble methods.

…

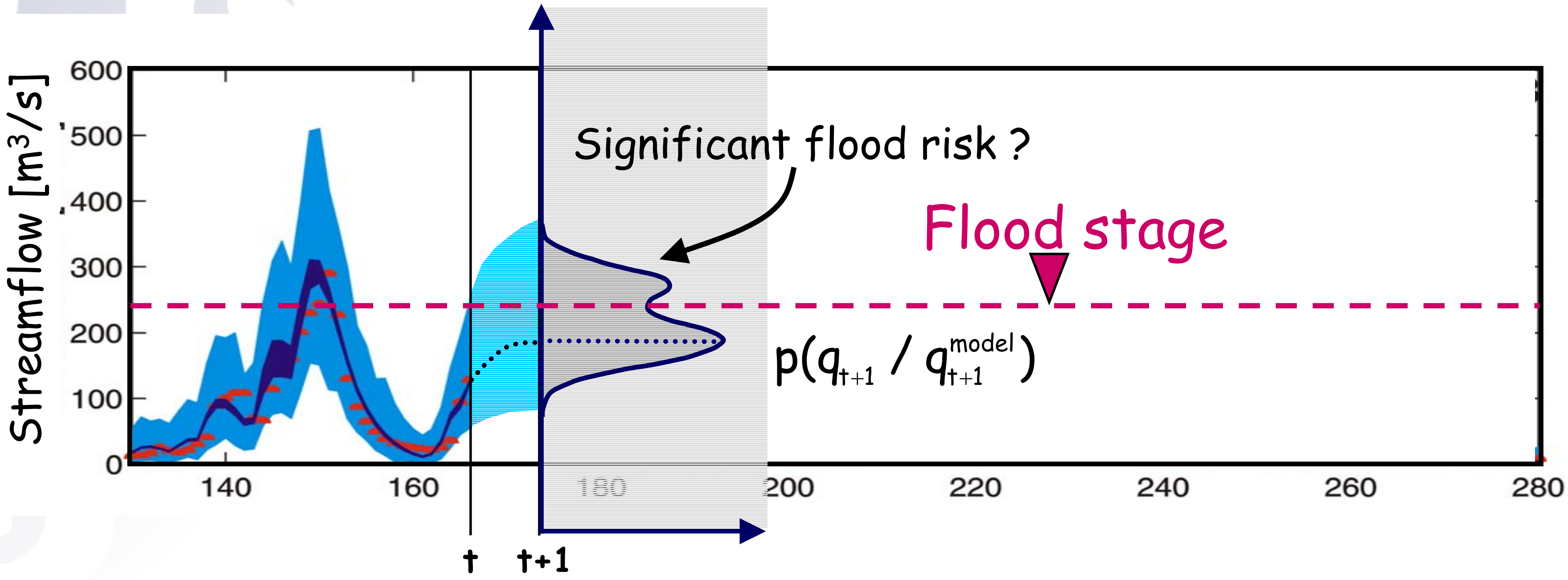
… Operational Needs

To be able to handle risk in decision making

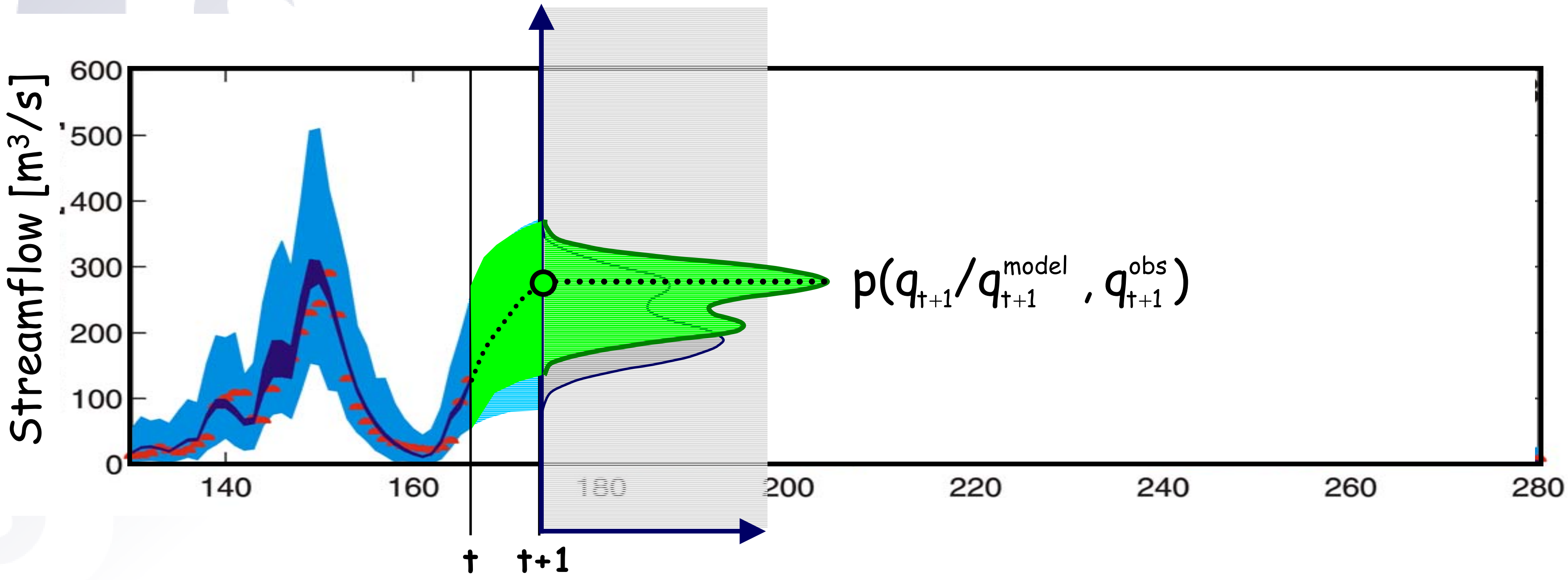
To be able to update forecasts as new information becomes available



Uncertainty estimate of streamflow prediction



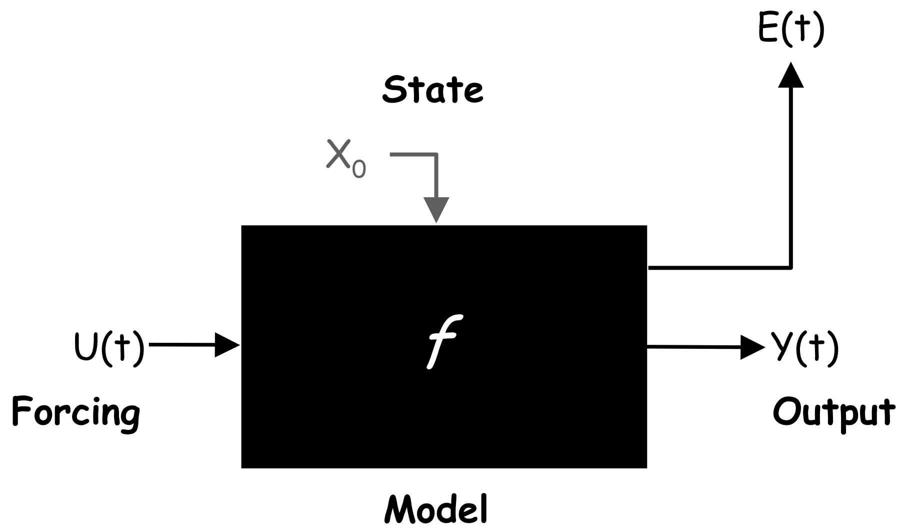
Updated estimate of streamflow uncertainty



Issues in Uncertainty Estimation … Systems Point of View …

Uncertainties exist in:

- Forcing
- Model Identification
- State
- Output Measurements



Merging Data with Models:

Multiple Sources and Types of Information
Data becomes available incrementally

Model Forcing (Precipitation) Uncertainty

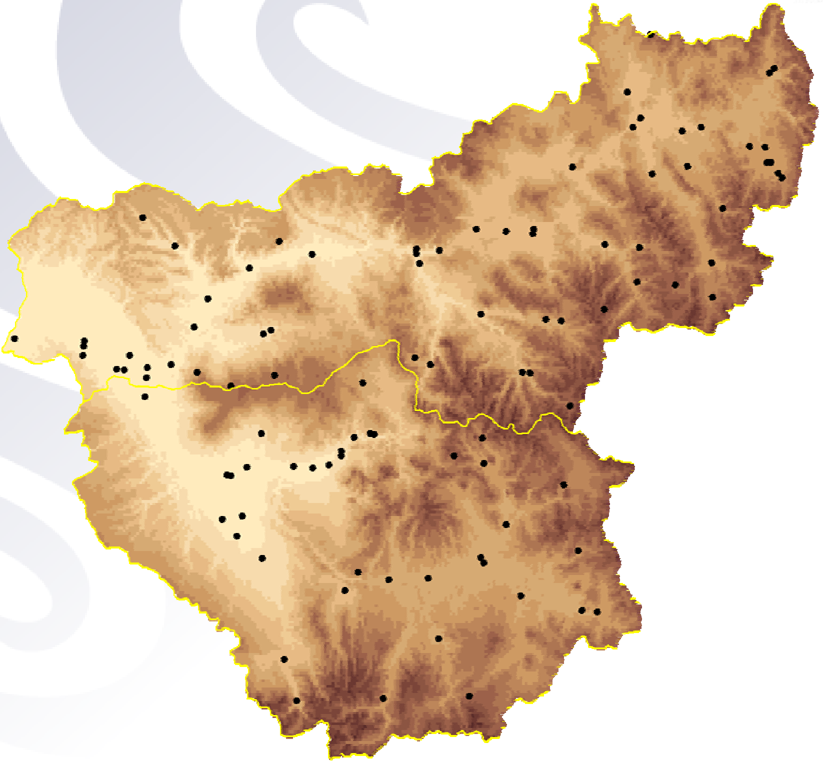
Sources of Uncertainty:

Detection
Measurement (Spacing, Support, Scale)
Coverage
Aggregation/disaggregation

Sources of Data:

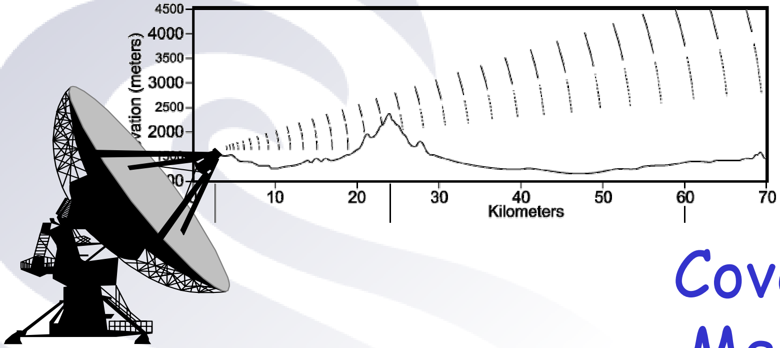
Gages
Radar
Satellite (indirect, time-space scale)
*Models**
Combinations / Other

Precipitation Gages



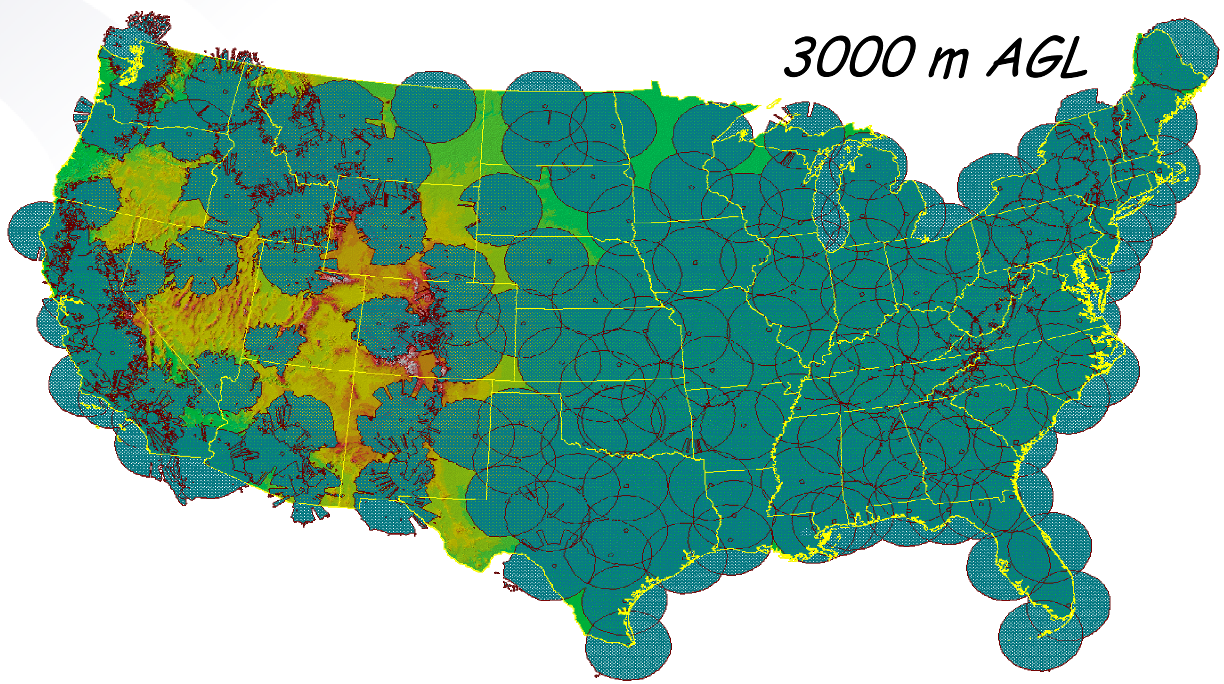
Accuracy of Catch
Sparsity of Coverage
Representativeness of Location

Basin-scale areal estimates obtained from "point" measurements by aggregation (e.g., Thiessen polygons)



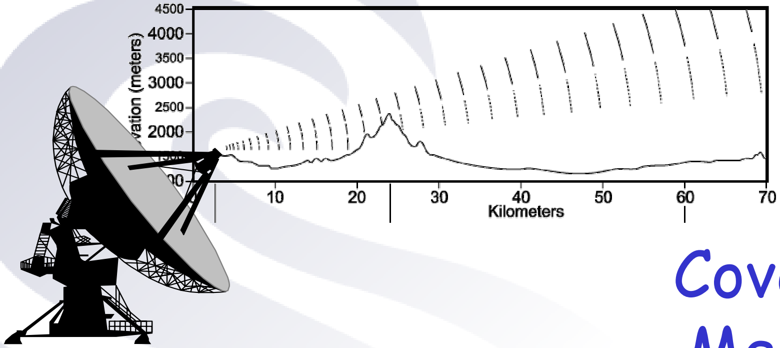
Ground-based Radar

Indirect measurement
Coverage blocked by mountains etc
Measures precipitation "in the air"



Coverage of the WSR-88D network over the US

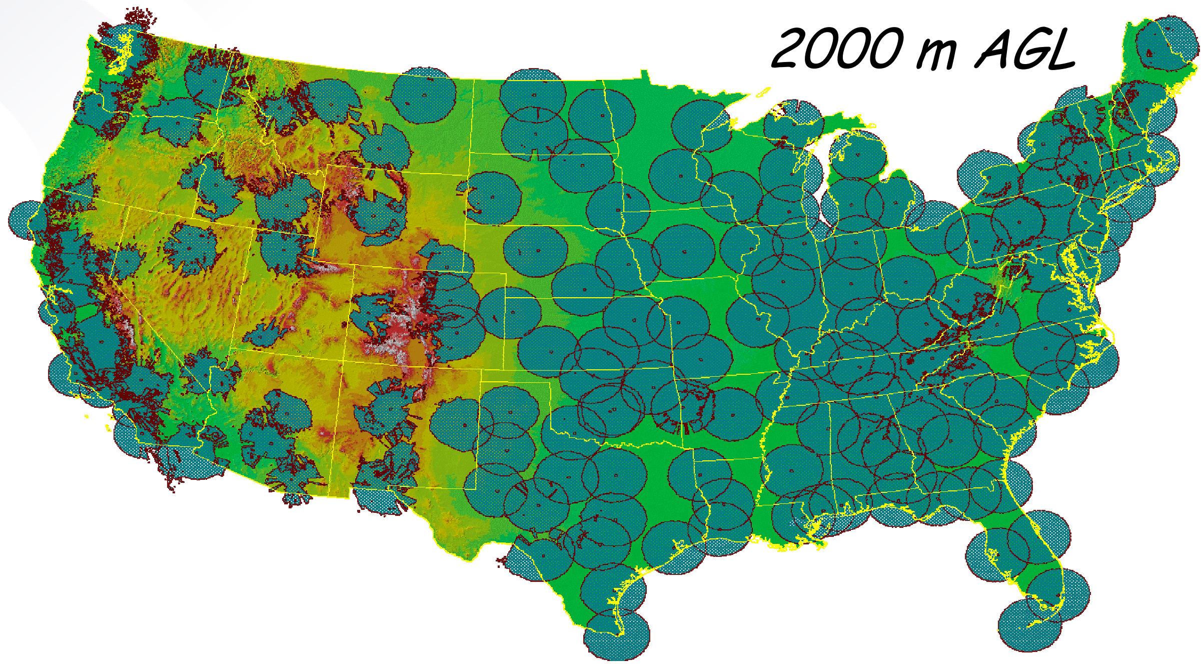
Ground-based Radar



Indirect measurement

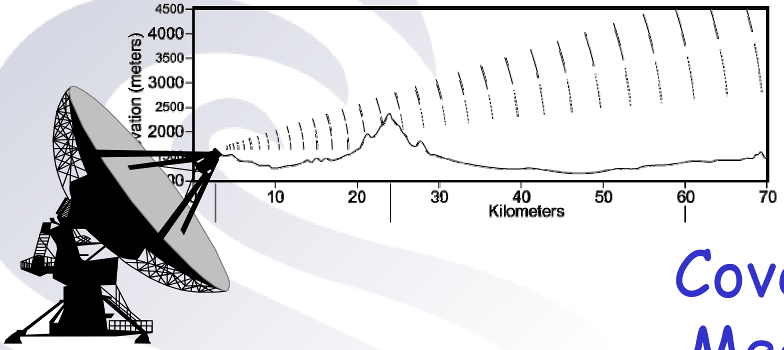
Coverage blocked by mountains etc

Measures precipitation "in the air"

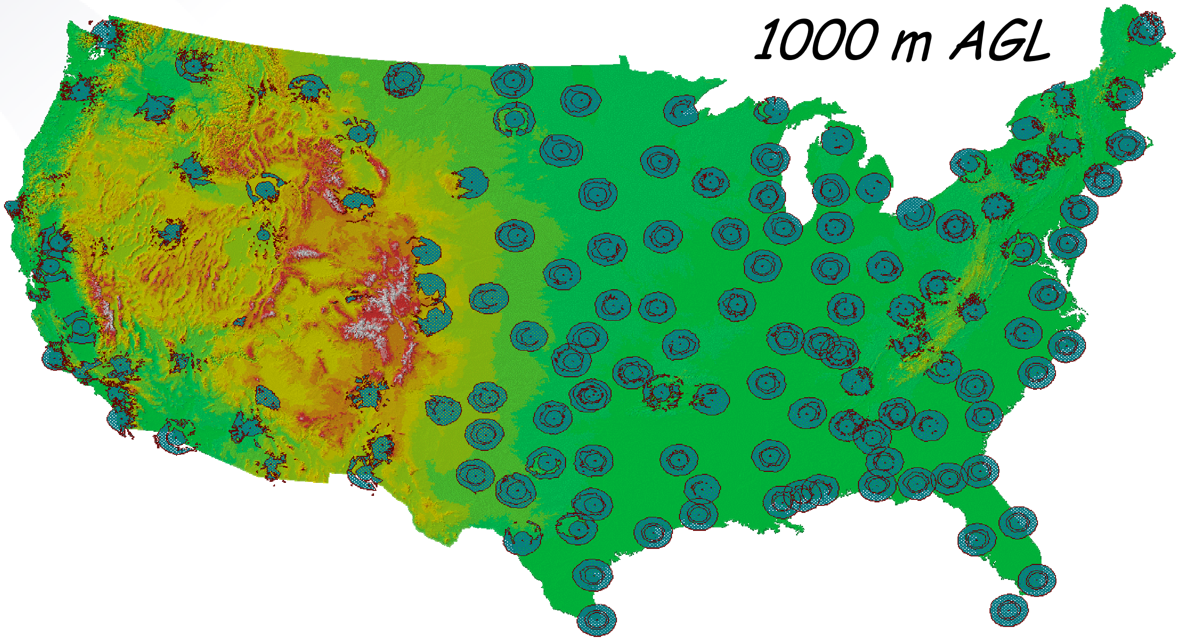


Coverage of the WSR-88D network over the US

Ground-based Radar

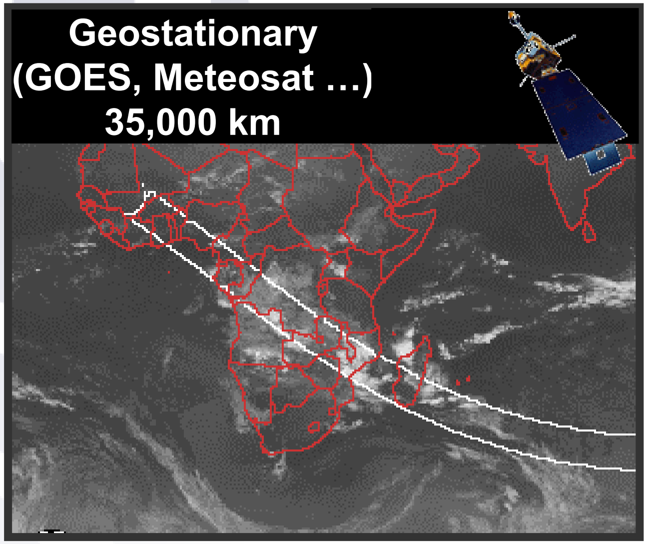


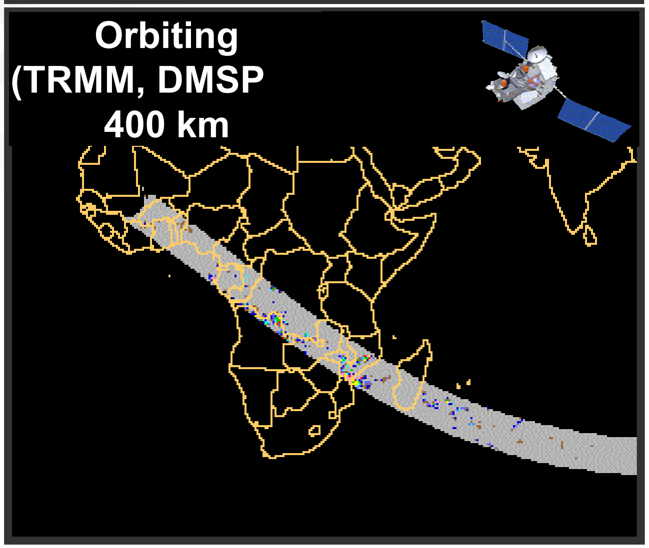
Indirect measurement
Coverage blocked by mountains etc
Measures precipitation "in the air"



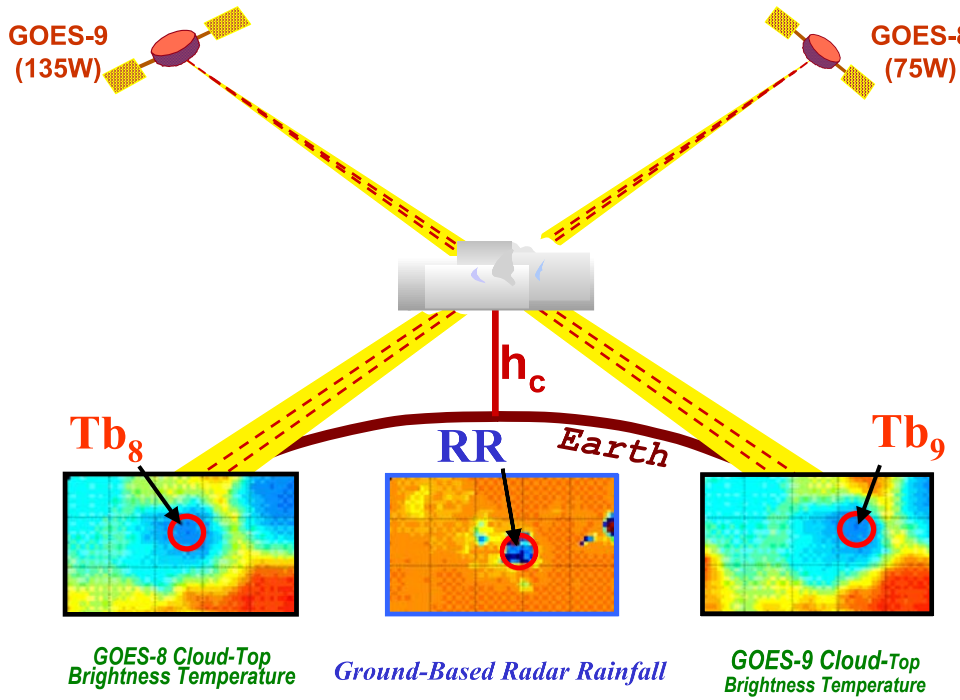
Coverage of the WSR-88D network over the US

Space-based Remote Sensing



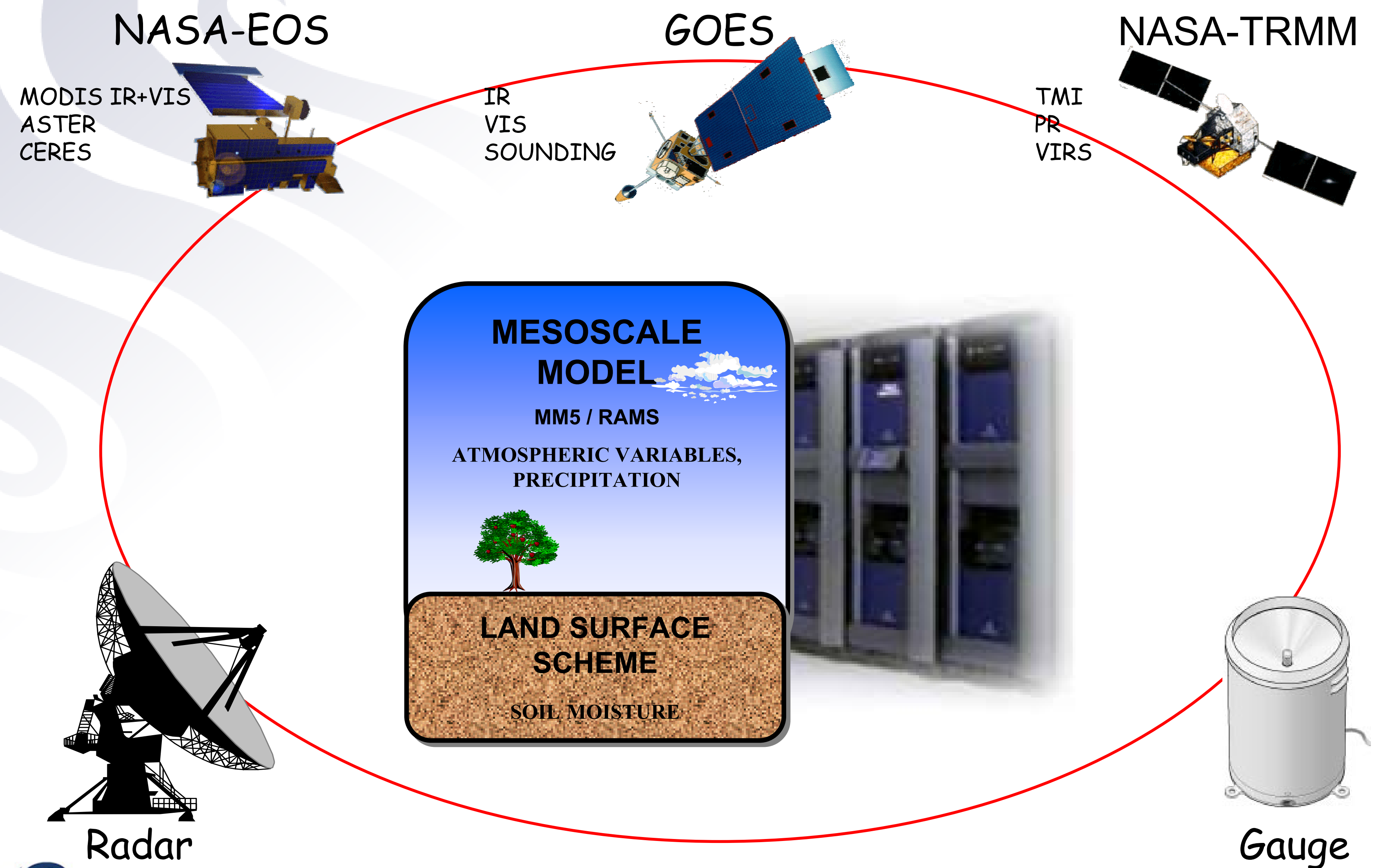


Indirect measurement
Time-space scale
Measures "in-the-air"
Parallax problem



(Pixel Size = 2 km)

Model & Combination Approaches



e.g., Sorooshian, Hsu, Gao etc (UC Irvine)

Model Identification Uncertainty

Multiple Plausible Descriptions:

Conceptualization
- Control Volume/Domain
- Inputs, State Variables, Outputs
- Feedbacks
- Components to be included/ignored

Mathematical Representation
- Structural Equations
- Deterministic / Stochastic

System Invariants
- Parameter Values
- Constants

… and the related "State" Uncertainty

System "Wetness":

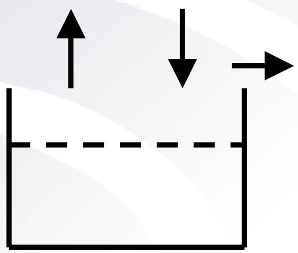
Conceptualization / Definition

- What is soil moisture anyway? [$dS/dt = I - O$]
- Dimensionality (low-D representation of infinite-D)
- Scale

Observability

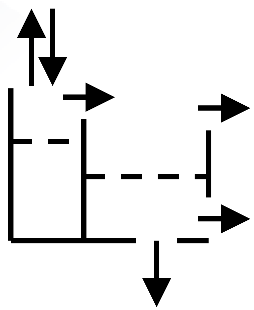
- Is there a high enough correlation between the "modeled state variable" and the observable quantity?

Examples – Different vertical representations of "Soil Moisture"



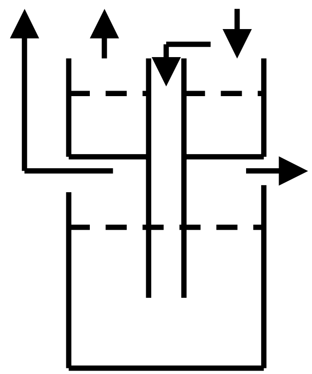
Bucket

1-D



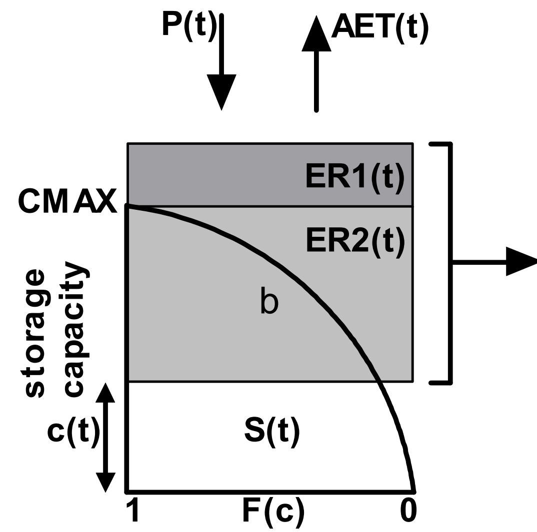
Tension/Free Store

2-D



Penman

2-D

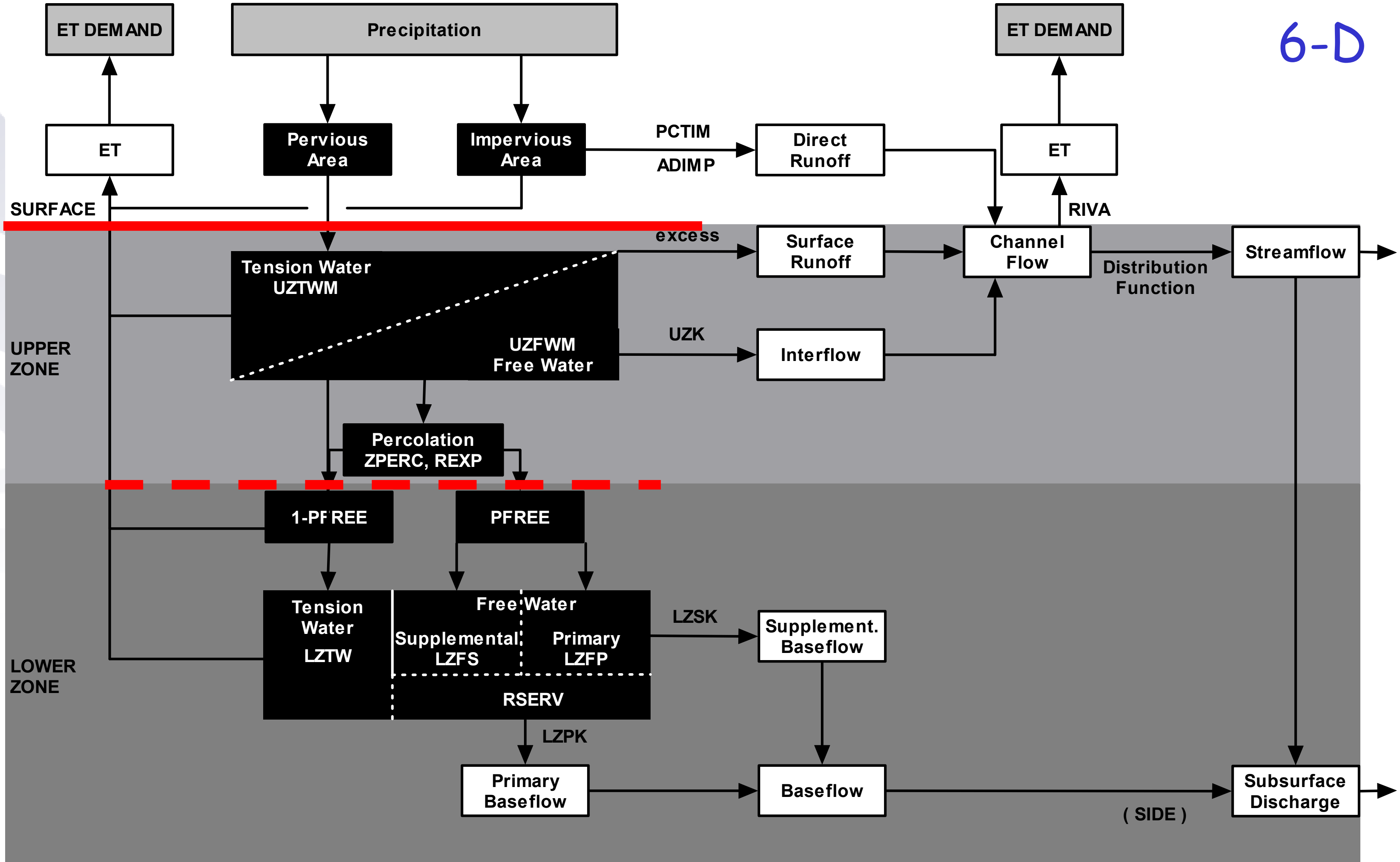


Probability Distributed Model (PDM)

1-D

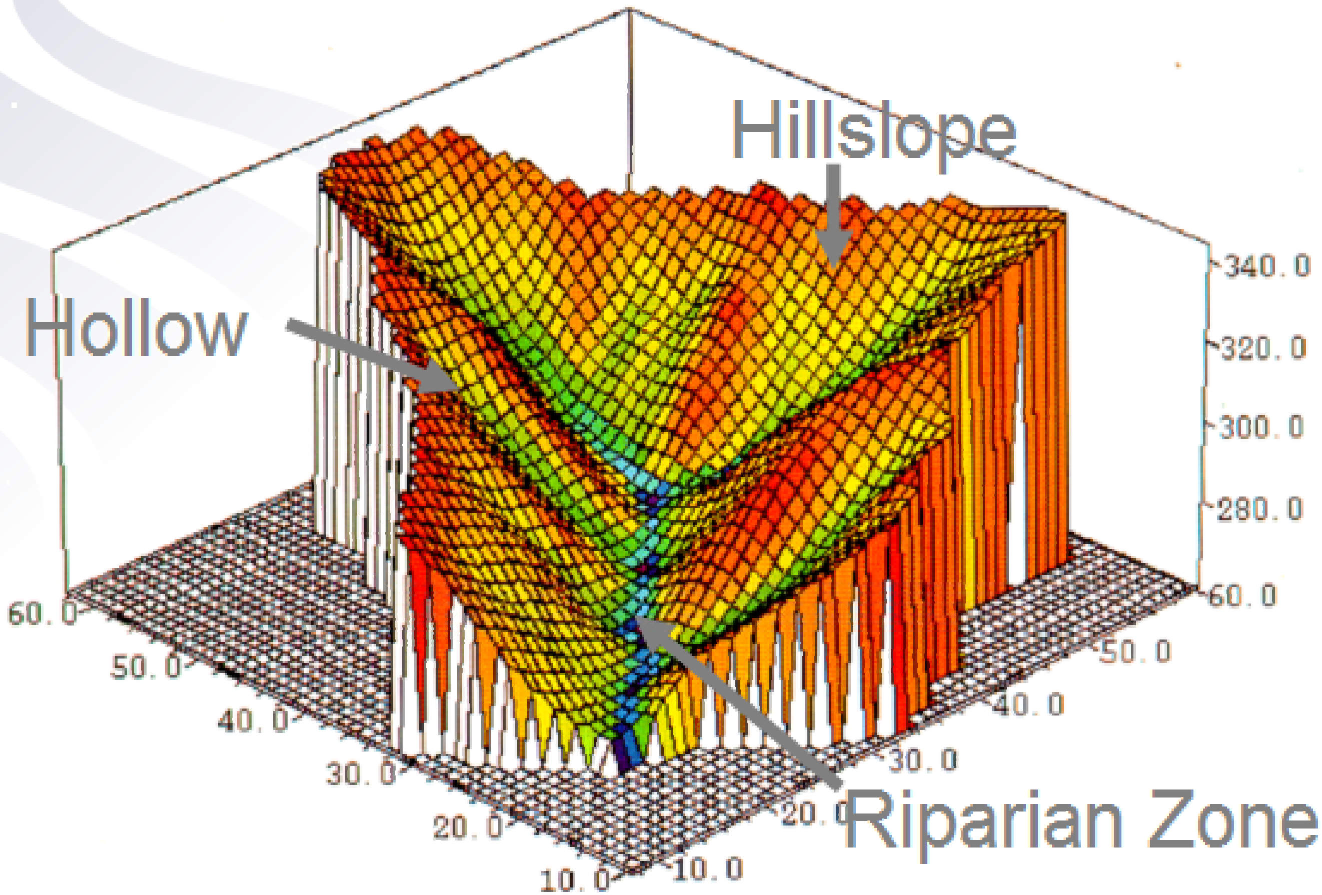
Example -- Sacramento Model (NWS) representation of "Soil Moisture"

6-D



Example – Spatially Distributed Model representation of "Soil Moisture"

?-D



Output Measurement Uncertainty

Observations:

Evapotranspiration
Soil Moisture
Streamflow

Measurement Problems:

Detection
Representativeness
Scale
Measurement Error (Bias, Heteroscedasticity)

Thomas Bayes

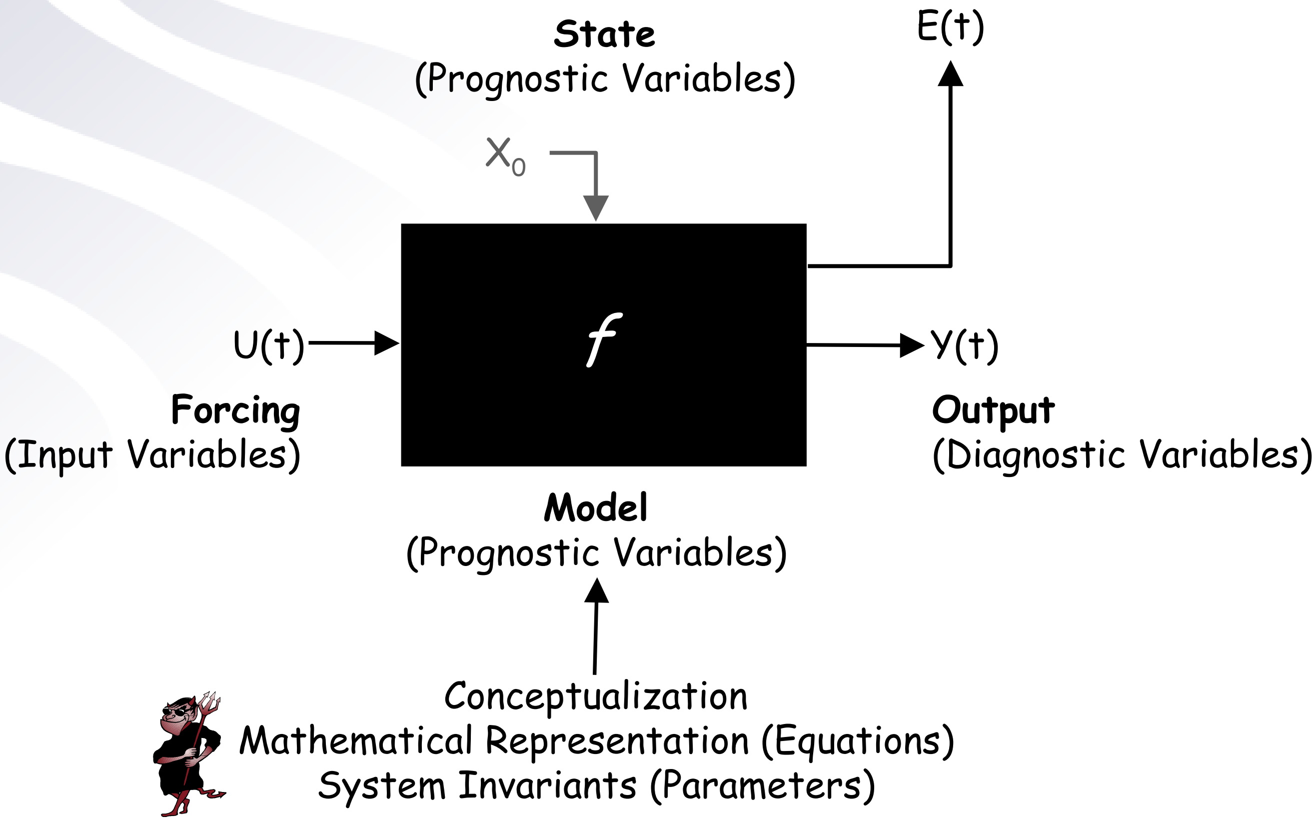
1702-1761, England

The fundamental basis for combining different types of <u>uncertain information</u> is given by Bayes Law

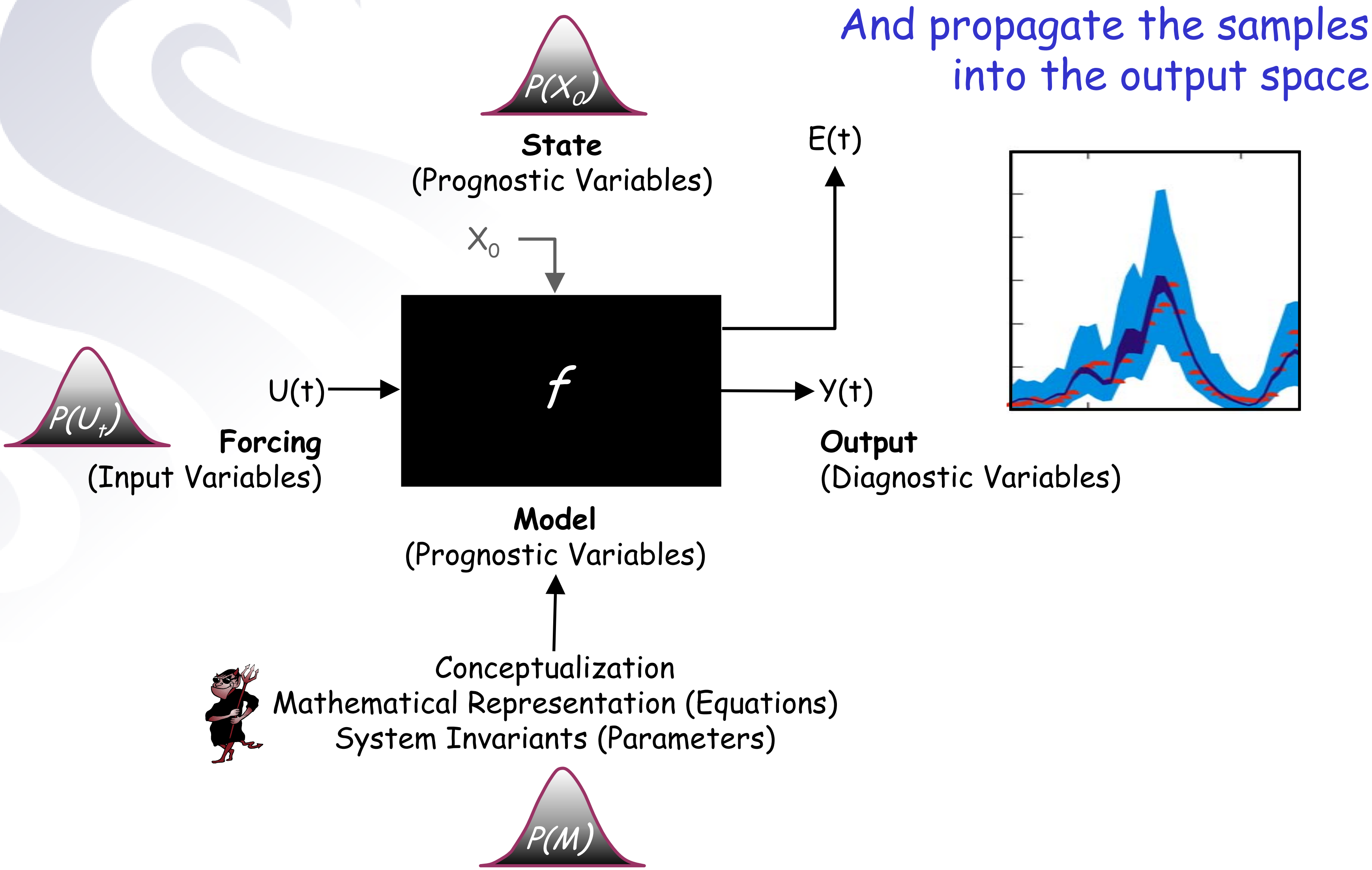
Implementation of Bayes Law By the Ensemble Approach

Prediction (Propagating the Uncertainty)
Data Assimilation (Updating the Uncertainty)

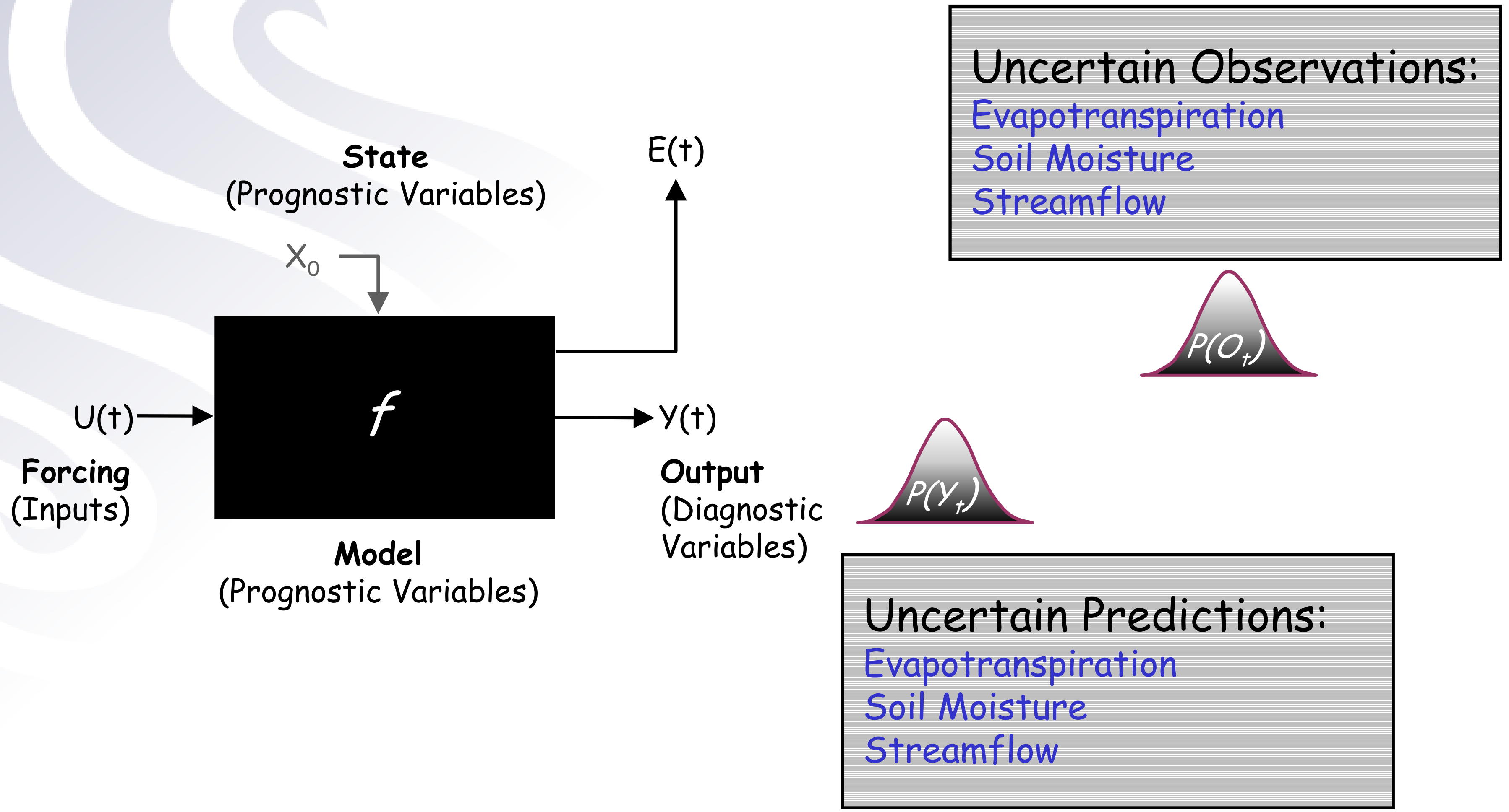
Prediction (Propagating the Uncertainty) …



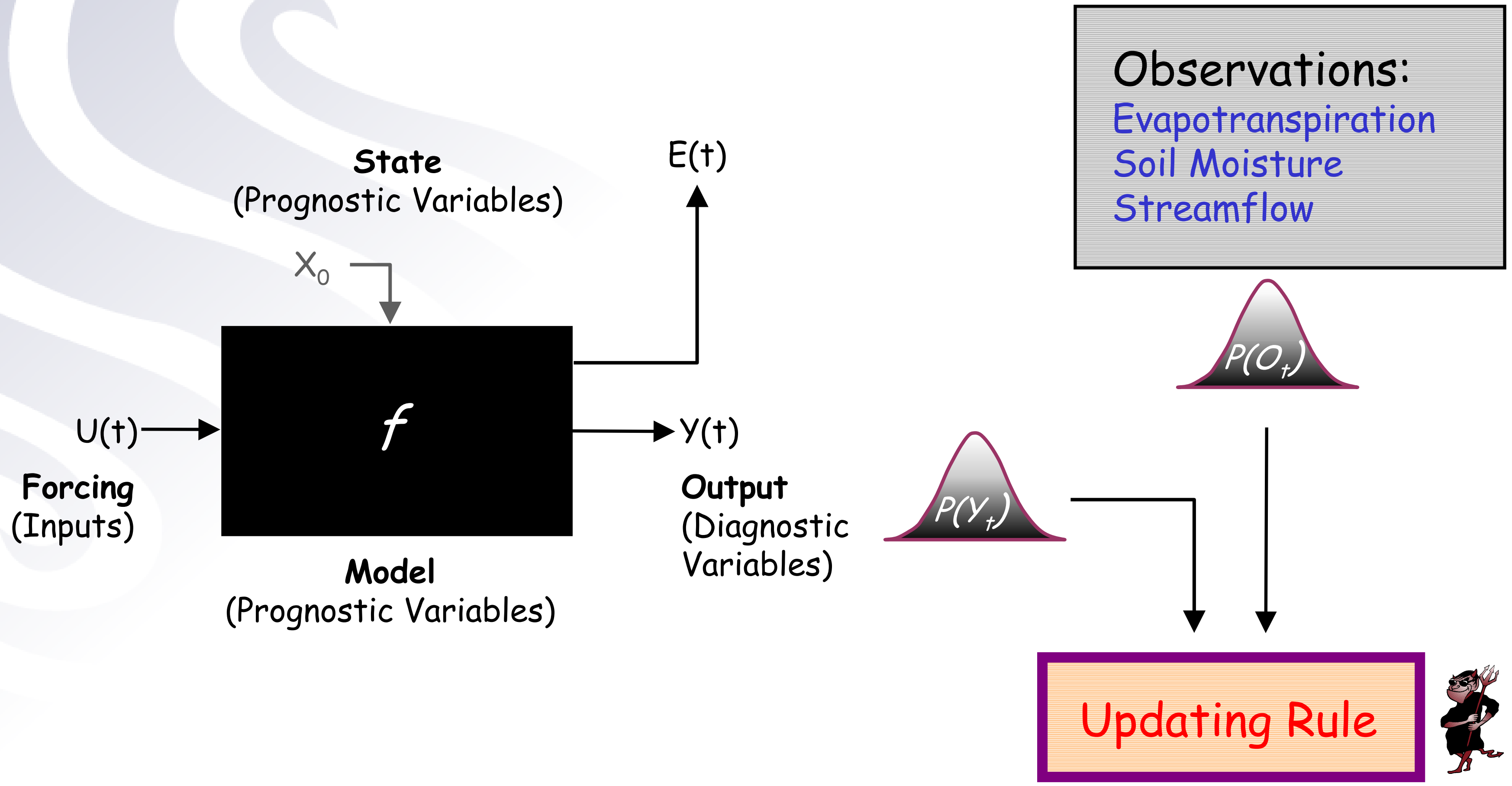
Use Ensembles that sample the space of the uncertainty …



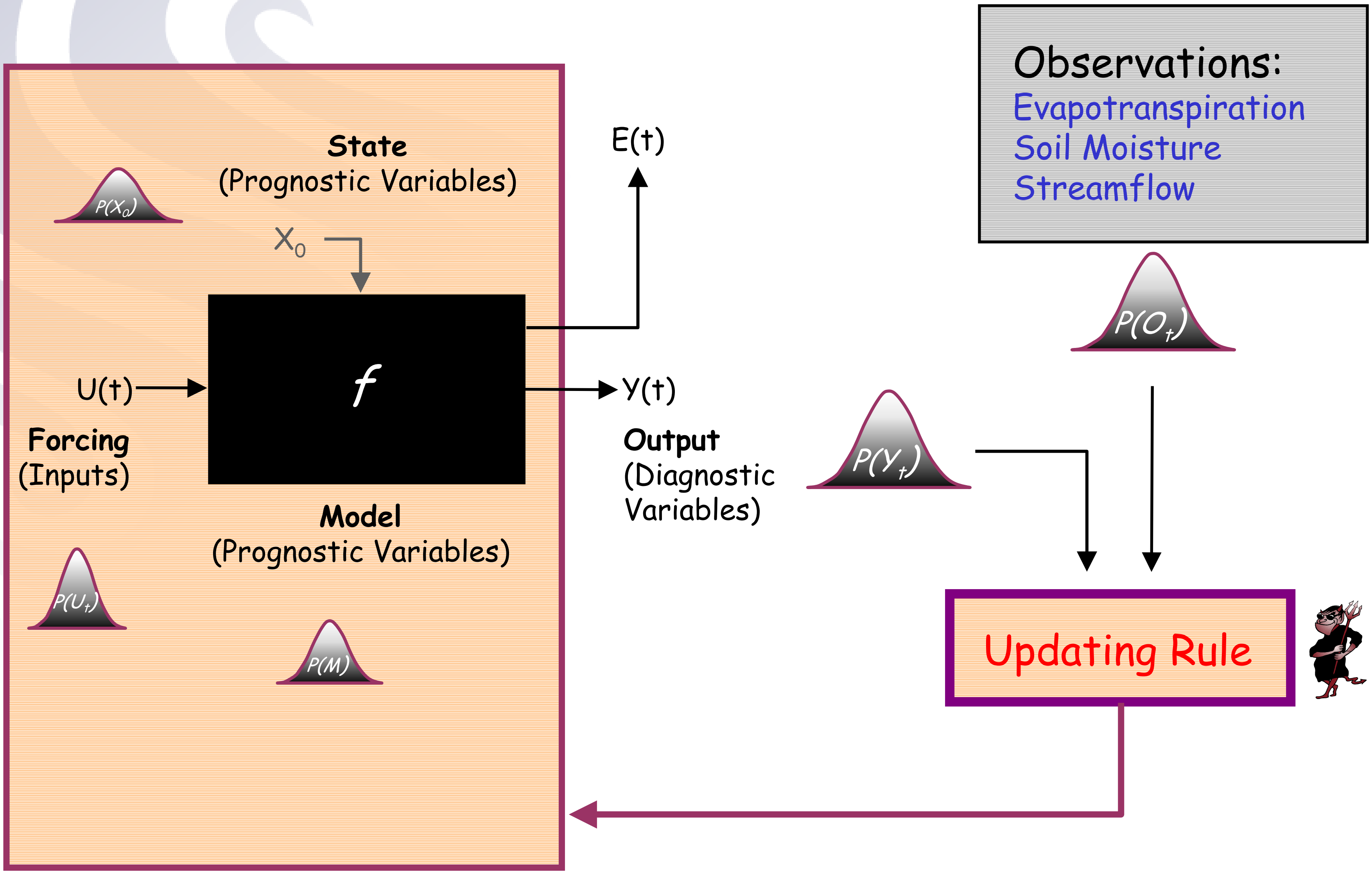
Data Assimilation (Updating the Uncertainty) …



Data Assimilation (Updating the Uncertainty) …



Data Assimilation (Updating the Uncertainty) …



EnKF -- Ensemble Kalman Filter:

Evensen, 1994 // Evensen and van Leeuwen, 1996

$$X^i(+) = X^i(-) + K\,[\,O^i - M\{X^i(-)\}\,]$$

Where the gain "K" depends on the cross covariance of
the forecast Y = M{X} with
the unknown X to be estimated

Note:
The unknown is typically chosen to be the (uncertain) state variable
But could also be the (uncertain) model and/or (uncertain) forcing

Characteristics of EnKF (McLaughlin, Waginingen Workshop, Sept 2001)

- Well-suited for real time applications +
- Provides information on estimation accuracy +
- Flexible, modular, can accommodate wide range of model error descriptions +
- No need for adjoint model, linearizations or other model approximations +
- Robust and easy to use +

- Update assumes states are jointly normal –
- Can be computationally demanding –

Applications of EnKF in Hydrology

State Estimation:

- Soil moisture estimation
  Reichle, McLaughlin, Entekhabi, 2002
- Coastal areas modeling
  Madsen, Canizares, 1999
- Others …

Joint State-Parameter Estimation:

- Simultaneous Optimization and Data Assimilation:
  Vrugt, Gupta, Diks, Bouten, Verstraten, 2004 (in press)
- Dual State-Parameter Estimation
  Moradkhani, Sorooshian, Gupta, Houser, 2004 (to be submitted)

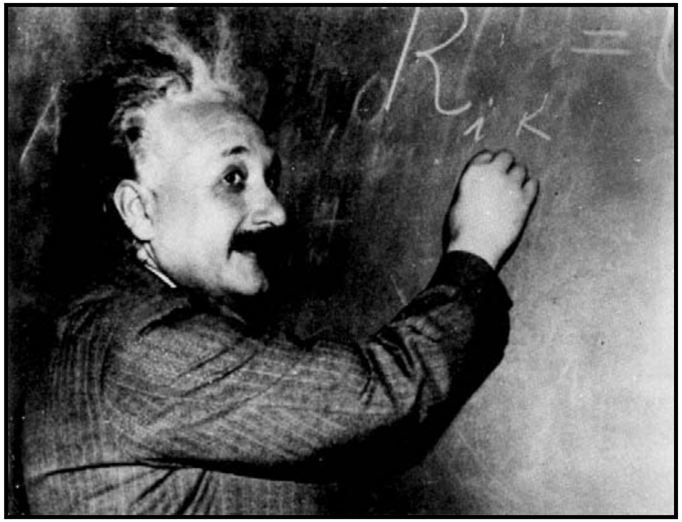
Non-Linear Updating Rules …

Bayesian Non-Linear Updating using Ensembles:

- RSA (Regional Sensitivity Analysis) - behavioral/non-behavioral thresholds
  Hornberger and Spear, 1981
- GLUE (Generalized Uncertainty Analysis)
  Binley & Beven, 1991
- BaRE2 (Bayesian Recursive Estimation)
  Thiemann, Trossett, Gupta, Sorooshian, 2001 // Misirli, Gupta, Thiemann, Sorooshian, 2003
- DyNIA (Dynamic Identifiability Analysis)
  Wagener, McIntyre, Less, Wheater, Gupta, 2003
- SCEM (Shuffled Complex Evolution Metropolis)
  Vrugt, Gupta, Bouten, Sorooshian, 2003
- MOSCEM (Multi-objective SCEM)
  Vrugt, Gupta, Bastidas, Bouten, Sorooshian, 2003
- SODA (Simultaneous Optimization and Data Assimilation: EnKF + SCEM)
  Vrugt, Gupta, Diks, Bouten, Verstraten, 2004
- BMA (Bayesian Model Averaging)
  Neuman, 2003
- Other …

Major areas of research interest today are

Model Structural Uncertainty
&
Forcing Uncertainty



So far as the laws of mathematics refer to reality, they are not certain. And so far as they are certain, they do not refer to reality.

Albert Einstein
Geometry & Experience