BAYESIAN PROCESSOR OF OUTPUT: A NEW TECHNIQUE FOR PROBABILISTIC WEATHER FORECASTING

By

Roman Krzysztofowicz and Coire J. Maranzano

University of Virginia

Presented at the

2nd International Symposium on

Quantitative Precipitation Forecasting and Hydrology

Boulder, Colorado

4 June – 8 June 2006

Acknowledgments:

Work supported by the National Science Foundation under Grant No. ATM–0135940.

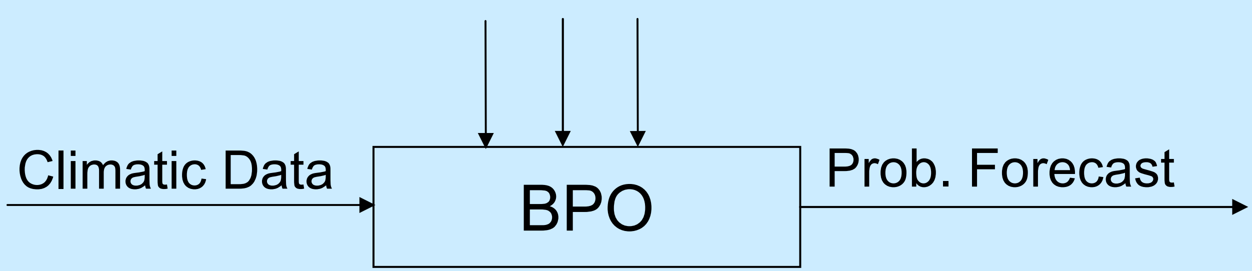
Data provided by the Meteorological Development Laboratory of the National Weather Service.

NEW STATISTICAL TECHNIQUES for PROBABILISTIC WEATHER FORECASTING

Techniques

Bayesian Processor of Output (BPO)
Bayesian Processor of Ensemble (BPE)

NWP Model Output(s)



- processes output
- fuses with climatic data
- quantifies uncertainty

Versions for

binary predictands
multi-category predictands
continuous predictands

BPO for CONTINUOUS PREDICTAND

Variates

W – precipitation amount, conditional on occurrence $W > 0$

$\mathbf{X}$ – vector of predictors

Bayesian Theory

$g(w)$ prior

$f(\mathbf{x}|w)$ conditional (likelihood)

$$\kappa(\mathbf{x}) = \int_{-\infty}^{\infty} f(\mathbf{x}|w)\, g(w)\, dw \qquad \phi(w) = \frac{f(\mathbf{x}|w)}{\kappa(\mathbf{x})}\, g(w)$$

Fusion

- Sample Asymmetry
- Mismatch: Sample Size – Model Complexity

Modeling

- *Marginals*: many forms (non-Gaussian)
- *Dependence*: non-linear, heteroscedastic

=> Meta-Gaussian Distribution (*Kelly, Krzysztofowicz, 1994*)

FORECASTING EQUATIONS

- Posterior Distribution Function

$$\Phi(w) = Q\left(\frac{1}{T}\left[Q^{-1}(G(w)) - \sum_{i=1}^{I} c_i Q^{-1}(\bar{K}_i(x_i)) - c_0\right]\right)$$

- Posterior Density Function

$$\phi(w) = \frac{1}{T}\frac{\exp\left(-\frac{1}{2}[Q^{-1}(\Phi(w))]^2\right)}{\exp\left(-\frac{1}{2}[Q^{-1}(G(w))]^2\right)} g(w)$$

- Posterior Quantile

$$w_p = G^{-1}\left(Q\left(\sum_{i=1}^{I} c_i Q^{-1}(\bar{K}_i(x_i)) + c_0 + TQ^{-1}(p)\right)\right)$$

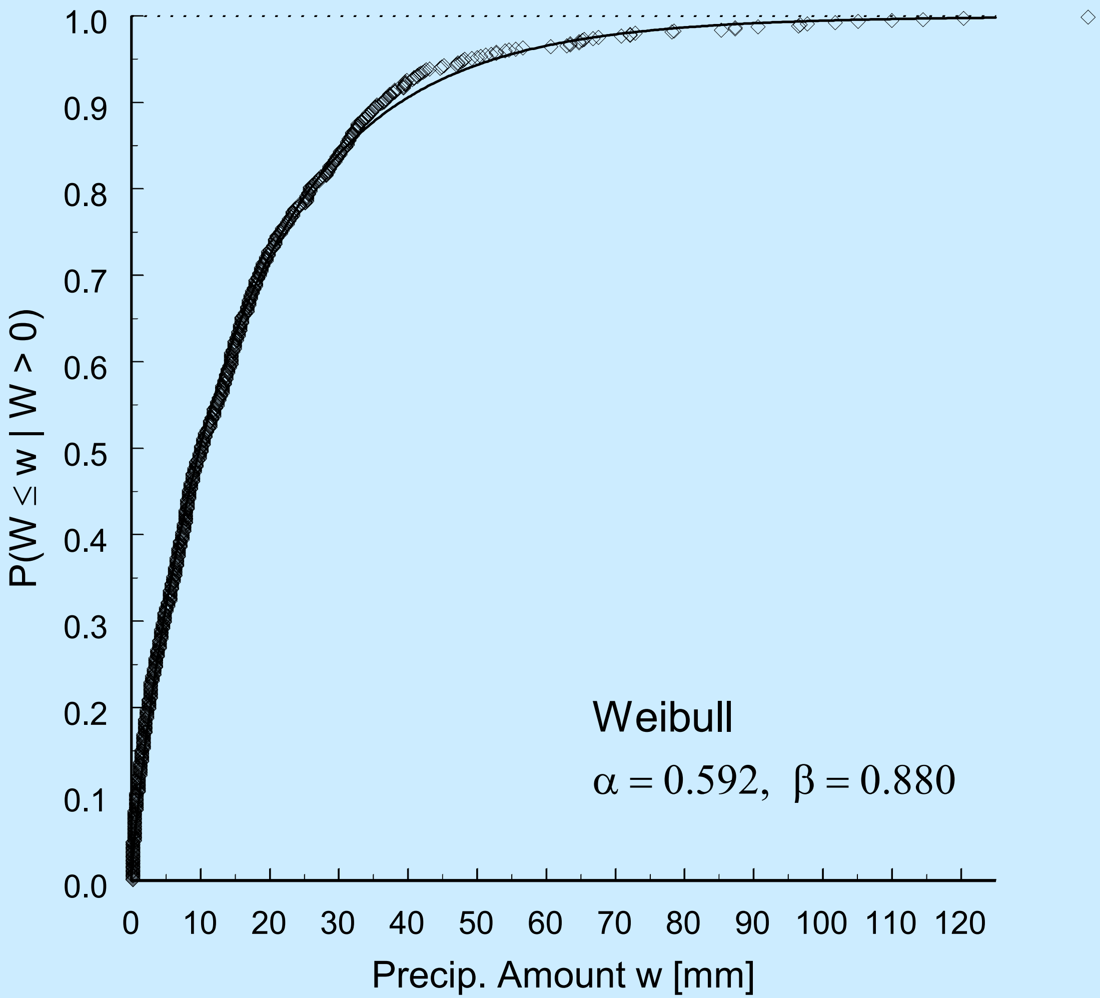
$$0 < p < 1$$

$$p = 0.1,\ 0.25,\ 0.5,\ 0.75,\ 0.9$$

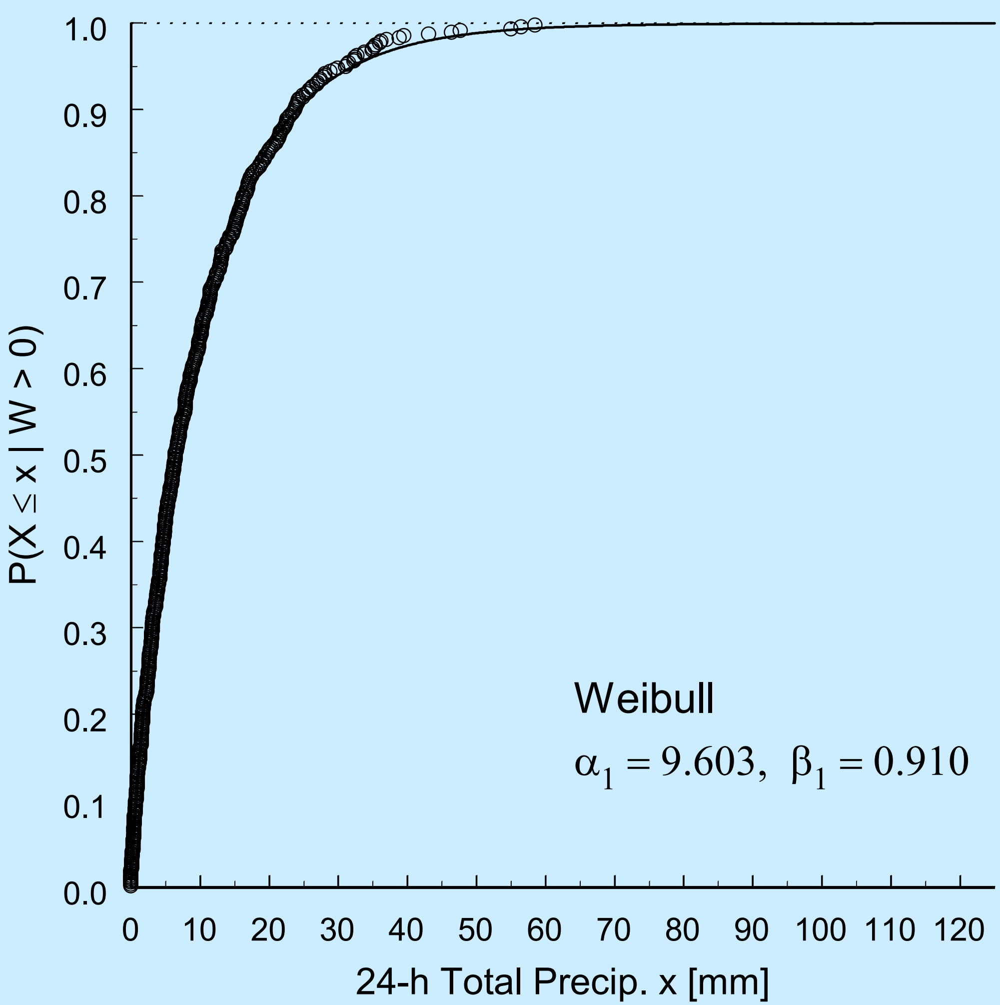
KUIL 12-36h Cond. Precip. Amount

Cool

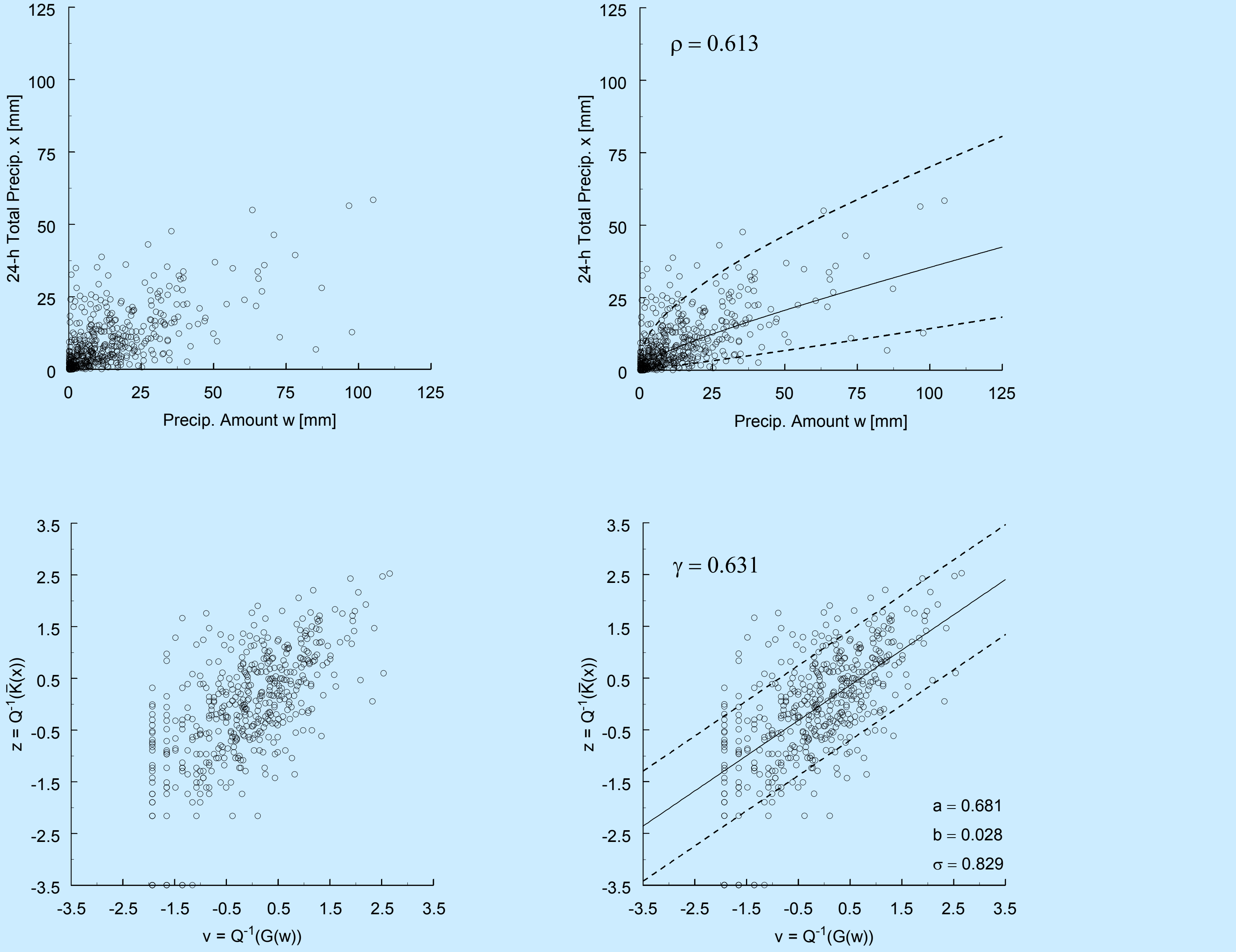
Climatic Prior Dist.



KUIL 12-36h, 24-h Total Precip. Cool



Likelihood Dependence Structure



EXAMPLE: Three Predictors

Quillayute, WA; cool season

W — 24-H PRECIP. AMOUNT, 12–36 h after 0000 UTC
X_1 — 24H TOTAL PRECIP. ending 36 h
X_2 — 850 REL. VORTICITY at 24 h
X_3 — 700 VERTICAL VELOCITY at 12 h

- Sample Sizes

Prior: 818 Joint: 470

- Distribution Functions

G	is Weibull:	$\alpha = 0.592,$	$\beta = 0.880$	
$\bar{K}_1$	is Weibull:	$\alpha_1 = 9.603,$	$\beta_1 = 0.910$	
$\bar{K}_2$	is Log-logistic:	$\alpha_2 = 6.212,$	$\beta_2 = 4.863,$	$\eta_2 = -5.0$
$\bar{K}_3$	is Log-logistic (–):	$\alpha_3 = 0.539,$	$\beta_3 = 4.313,$	$\eta_3 = -0.4$

- Posterior Parameters

$c_1 = 0.505$ $c_0 = -0.025$
$c_2 = 0.241$ $T = 0.641$
$c_3 = -0.275$

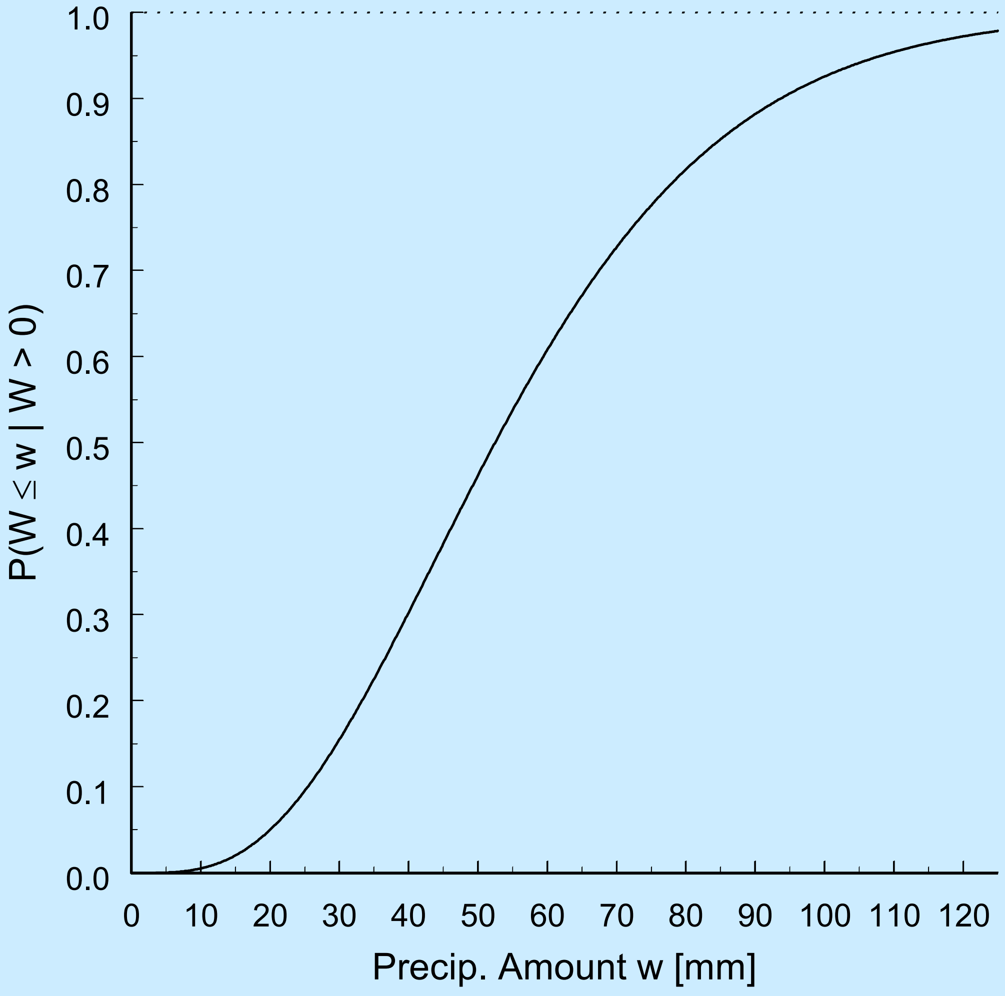
- Informativeness Score, IS

X_1	X_2	X_3	(X_1, X_2)	(X_1, X_3)	(X_1, X_2, X_3)
0.63	0.43	0.48	0.73	0.73	0.77

KUIL 12-36h Cond. Precip. Amount

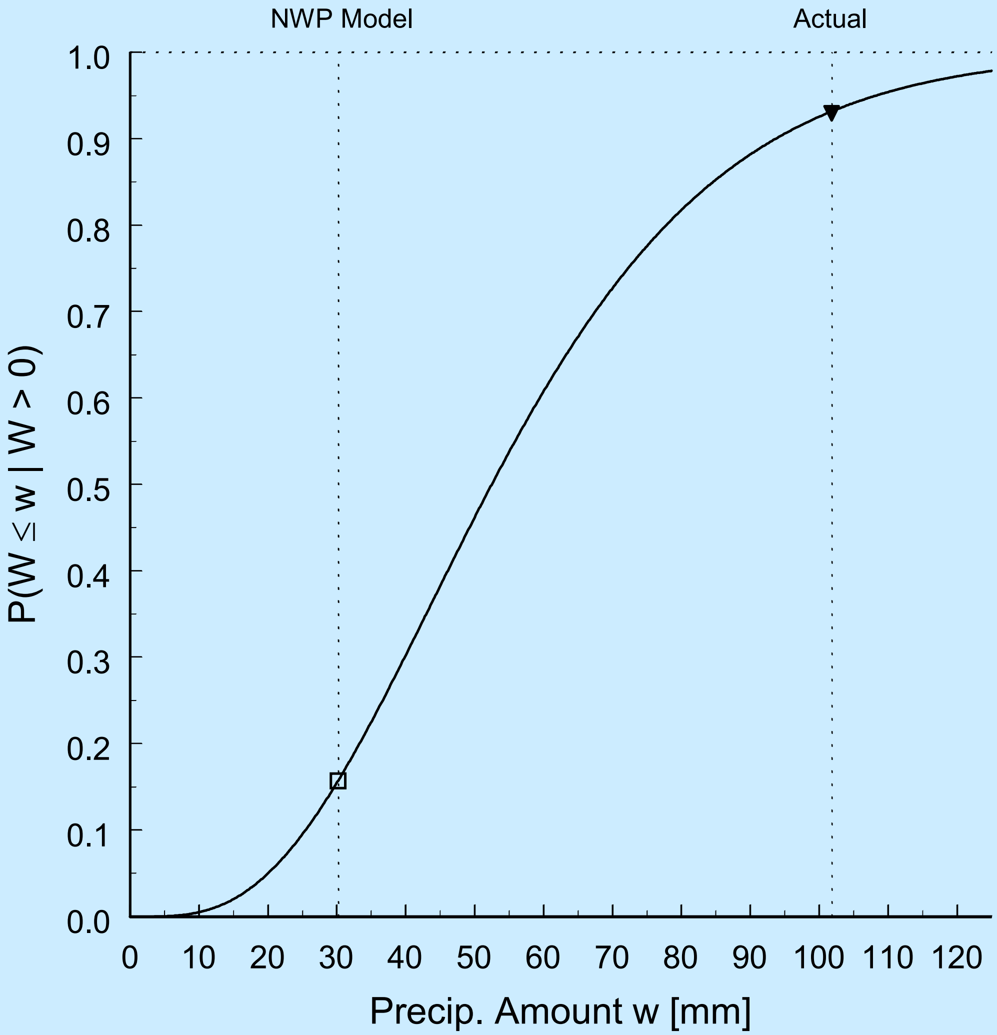
Cool

BPO forecast



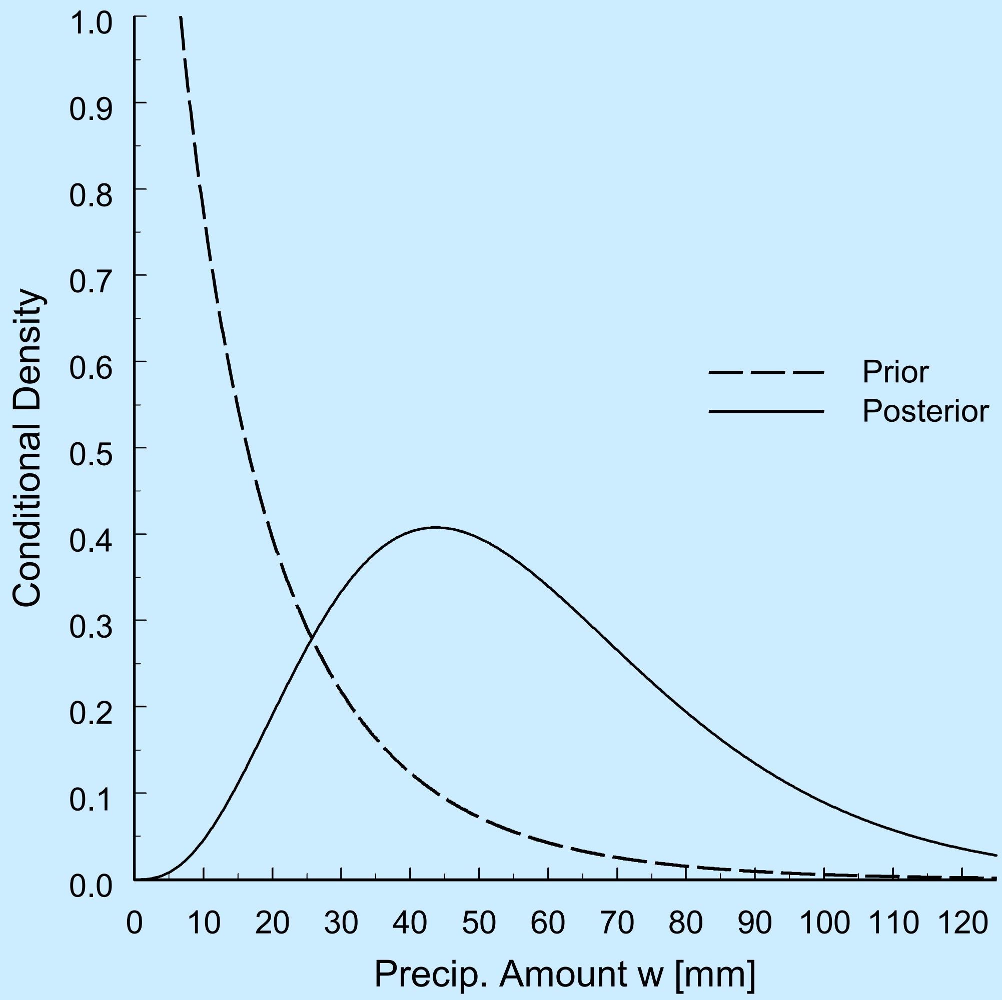
KUIL 12-36h Cond. Precip. Amount Cool

BPO forecast



KUIL 12-36h Cond. Precip. Amount Cool

BPO forecast



EXAMPLE: Conditional PQPF

Quillayute, WA Cool season
24-h, 12–36 after 0000 UTC
21 February 2002

- **BPO: 3 predictors; 15 parameters**

24-H TOTAL PRECIP. ending 36 h	$x_1 = 30.2$
850 REL. VORTICITY at 24 h	$x_2 = 4.8$
700 VERTICAL VELOCITY at 12 h	$x_3 = -0.95$

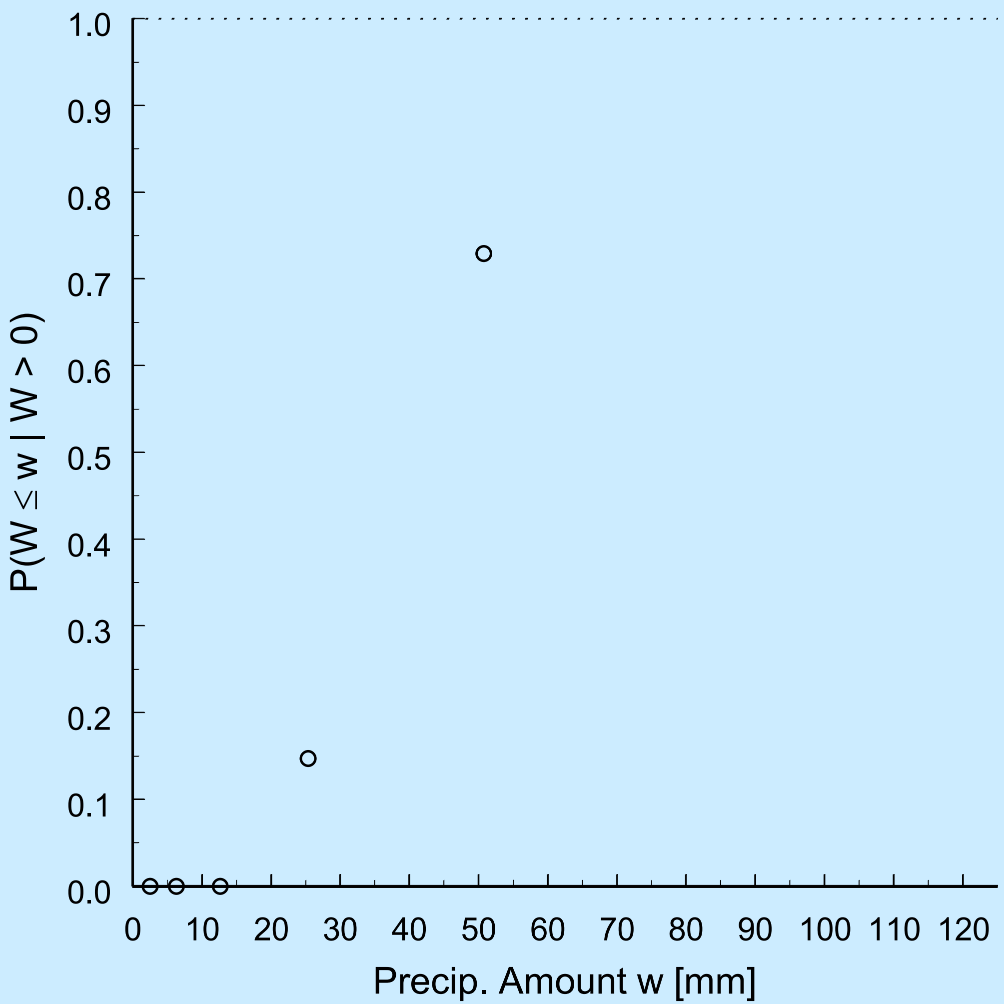
- **MOS: 15 predictors; 80 parameters (5 catego.)**

12-H TOTAL PRECIP. GB (6.35 mm) ending 24 h
12-H TOTAL PRECIP. GB (25.4 mm) ending 24 h
12-H TOTAL PRECIP. GB (0.254 mm) ending 36 h
→ 24-H TOTAL PRECIP. ending 36 h
12-H TOTAL PRECIP. GB (2.54 mm) ending 24 h
24-H TOTAL PRECIP. GB (12.7 mm) ending 36 h
850 REL. VORTICITY at 12 h
• LONGITUDE
12-H TOTAL PRECIP. GB (12.7 mm) ending 24 h
• ELEVATION
• LATITUDE
24-H CONV. PRECIP. GB (0.254 mm) ending 36 h
→ 850 REL. VORTICITY at 24 h
500 VERTICAL VELOCITY GB (-0.9) at 24 h
500 VERTICAL VELOCITY GB (-0.5) at 12 h

KUIL 12-36h Cond. Precip. Amount

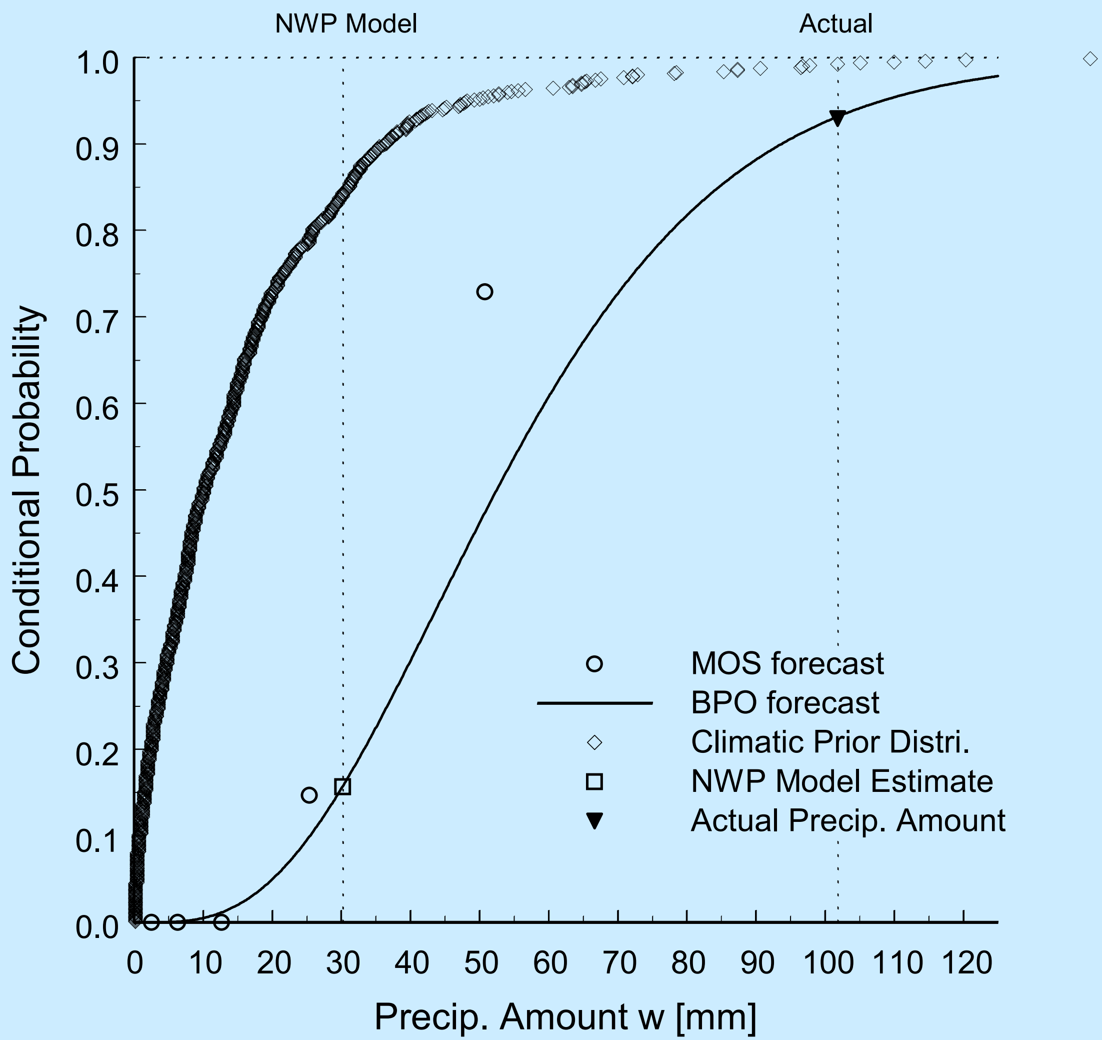
Cool

MOS forecast



KUIL 12-36h Cond. Precip. Amount

Cool



ATTRIBUTES OF BPO
Implied by Statistical Theory

1. Correct theoretic structure:
    - Always valid
    - Framework for different

        modeling assumptions

        estimation procedures

2. Appropriate parametric models:
    - *Marginals*: any form (non-Gaussian)
    - *Dependence*: non-linear, heteroscedastic

3. If models correct and estimation proper, then BPO
    - guarantees calibration
    - maximizes informativeness

IMPLICATIONS

1. Beneficial utilization of climatic data
   - stable calibration
   - user-specific calibration

BPO	MOS
Point-specific	Regional
Month-specific	Seasonal

2. Robustness when joint samples are smaller

3. Extension to ensemble processing less demanding
   - modeling complexity
   - computing requirements