Maximum Likelihood Bayesian Averaging Of Air Flow Models In Unsaturated Fractured Tuff Using Occam And Variance Windows

H43F-1098

Eric Morales-Casique[1], Shlomo P. Neuman[2], Velimir V. Vesselinov[3]

[1]morales@geo.uu.nl, Department of Earth Sciences, Utrecht University, The Netherlands; [2]neuman@hwr.arizona.edu, Department of Hydrology and Water Resources, The University of Arizona, Tucson, AZ, USA ; [3]vvv@lanl.gov, Los Alamos National Laboratory, Los Alamos, NM, USA

Why prediction with multiple models?

- Complexity of hydrologic environments
- Multiple interpretations / mathematical descriptions

Bayesian model averaging → optimal way of combining the predictions of several competing models and to assess their joint predictive uncertainty.

The maximum likelihood (ML) version of Bayesian model averaging (MLBMA) is compatible with ML methods of model calibration (Neuman, 2003).

Objectives

To test the ability of MLBMA based on both Occam's and variance windows to predict air pressure from pneumatic injection tests conducted in a complex, highly heterogeneous, unsaturated fractured tuff near Superior, Arizona.

MLBMA in practice

The maximum likelihood Bayesian model averaging MLBMA estimate of a variable Δ, conditioned on data set **D**, is given by

$$E[\Delta \mid \mathbf{D}] = \sum_{i=1}^{K} E[\Delta \mid M_i, \mathbf{D}]\, p(M_i \mid \mathbf{D})$$

$$Var[\Delta \mid \mathbf{D}] = \sum_{i=1}^{K} Var[\Delta \mid \mathbf{D}, M_i]\, p(M_i \mid \mathbf{D}) + \sum_{i=1}^{K} \left(E[\Delta \mid \mathbf{D}, M_i] - E[\Delta \mid \mathbf{D}]\right)^2 p(M_i \mid \mathbf{D})$$

Maximum likelihood Bayesian model averaging MLBMA is based on estimating posterior probabilities $P_i = P_i(M_i|\mathbf{D})$ for each alternative model.

$P_i = C \exp(-0.5\, \Delta IC_i)$ $\quad\quad$ $\Delta IC_i = IC_i - IC_{min}$

where C is a normalizing constant, IC_i and IC_{min} are information criteria (either *AIC*, *AICc*, *BIC* or *KIC*) for the *i*-th model and the minimum value among the models, respectively.

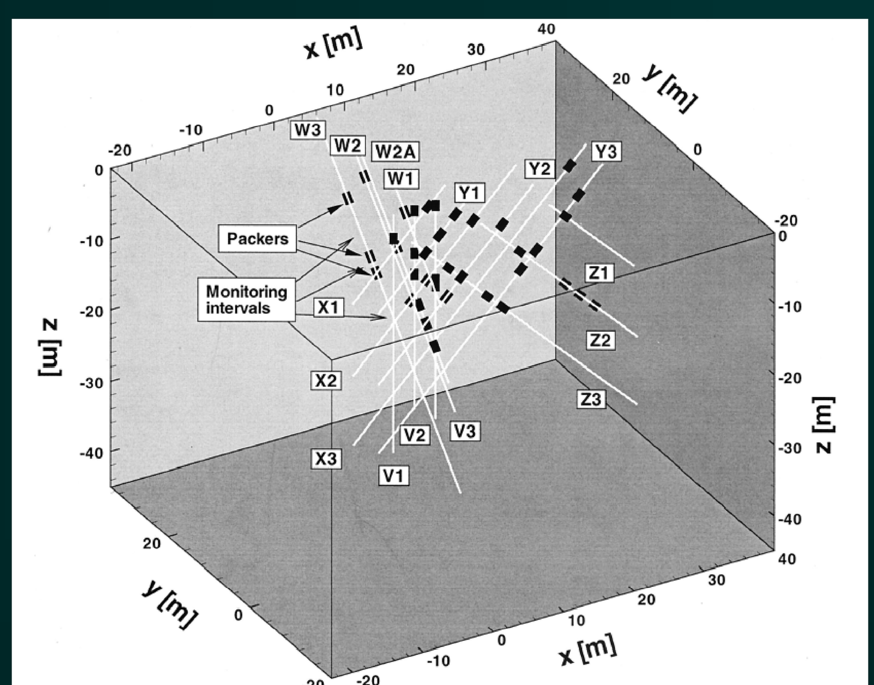
Experience indicates that estimating P_i in this way tends to assign posterior probabilities of nearly 1 to the best model and nearly zero to all other models.

Tsai and Li (2008) proposed to rely on a broader variance window

$P_i = C \exp(-0.5\, \alpha\, \Delta IC_i)$

α is selected subjectively by the analyst based on a desired level of significance.

Figure. Field setup



Calibration of air flow models

Data

- log permeability ($\log_{10}k$) and log porosity ($\log_{10}\phi$) data (from single-hole pneumatic packer tests)
- Cross-validation + MLBMA of 5 variogram models for $\log_{10}k$ and 4 for $\log_{10}\phi$ → selection of following models

$\log_{10}k$
- Exponential + linear drift – **E1**
- Exponential – **E**
- Power – **P**

$\log_{10}\phi$
- Exponential

- We parameterize $\log_{10}k$ and $\log_{10}\phi$ at selected pilot points and at some single-hole measurement locations and interpolate across the site via kriging.

- Pressure data from cross-hole pneumatic tests (Illman et al. 1998). Pressure response recorded in 32 intervals.

Table 1 Cross-hole tests conditions at ALRS (Illman et al. 1998)

Test	Flow regime	Injection Interval		Injection Rate (kg/s)
		Location	Length (m)	
PP4	Const. Rate	Y2-2	2	10^{-3}
PP5	Step*	X2-2	2.2	10^{-4}
PP6	Step*	Z3-2	2	10^{-4}
PP7	Step*	W3-2	1.2	10^{-4}

* Only data from the first stage was considered

- Values at pilot points are estimated by joint inverse simulation of cross-hole tests **PP4** and **PP5**.

PP4 (calibration) PP5 (calibration) PP6 (validation) PP7 (validation)

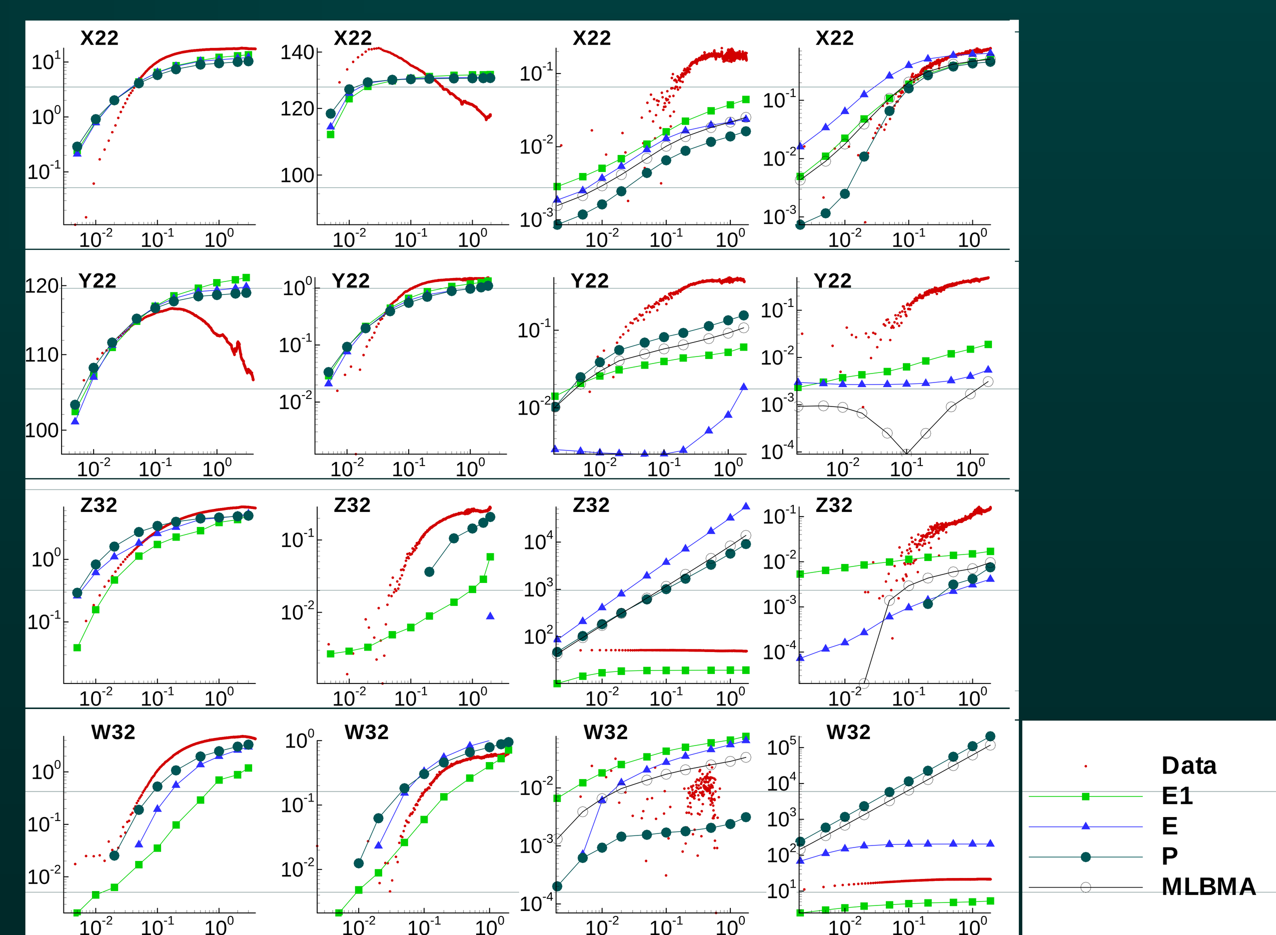


Figure. Pressure buildup (kPa) versus time (days) at four selected monitoring intervals

Calibration results

The traditional Occam's window approach in conjunction with *AIC*, *AICc*, *BIC* and *KIC* assigns a posterior probability of nearly 1 to the power model.

The variance window approach does the same when applied in conjunction with *AIC*, *AICc* or *BIC* but spreads the posterior probability more evenly among the three models when used in conjunction with *KIC*.

Model selection criteria and **posterior probability** for joint calibration of tests PP4 and PP5.

Model	*E1*	*E*	*P*
Parameters at pilot points	64	64	64
Variogram parameters	8	4	4
Pressure data	462	462	462
NLL	2488	2343	2213
ΔAIC, ΔAICc, ΔBIC	275	130	0
ΔKIC	29	52	0
P_{BIC} %, $\alpha = 1$	2E-58	6E-27	~100
P_{KIC} %, $\alpha = 1$	6E-05	5E-10	~100
P_{BIC} %, $\alpha = 0.049$	0.11	3.91	95.99
P_{KIC} %, $\alpha = 0.049$	27.81	15.70	56.50

NLL = Negative log likelihood

AIC = Akaike; AICc = Modified Akaike; BIC = Bayesian; KIC = Kashyap.

P_{IC} = posterior probability based on model information criteria *IC* for a given variance window (α = 1 corresponds to Occam's window).

Validation and predictive capabilities

- Independent data set → **PP6** and **PP7**
- Different location of injection → different flow pattern
- Calibration yields a ML estimate of parameters **b*** and covariance of estimation $\mathbf{C}_b$
- We assume a Gaussian distribution with mean = **b*** and covariance $\mathbf{C}_b$

Predictive logscore

The lower the predictive logscore the higher the amount of information recovered by the model in the validation set.

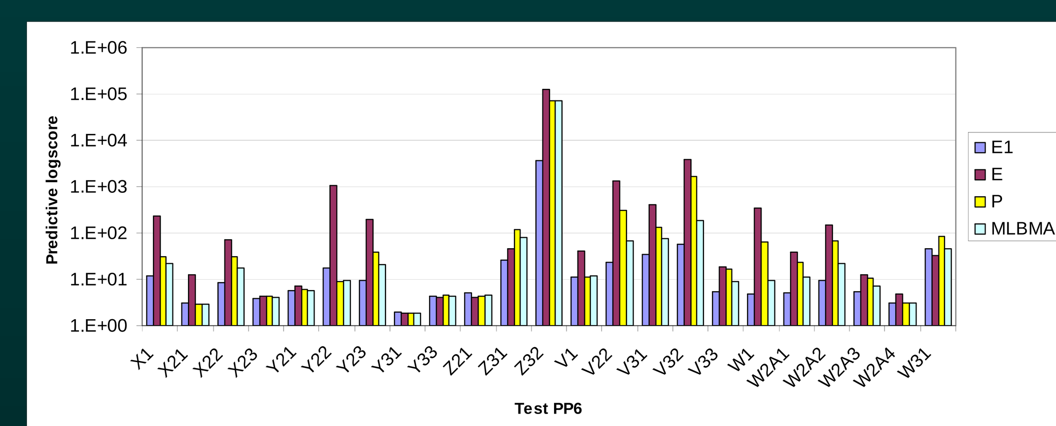
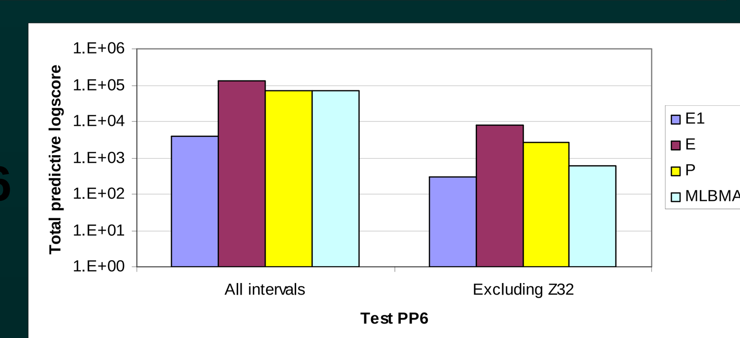


Figure. Predictive logscore, test **PP6**



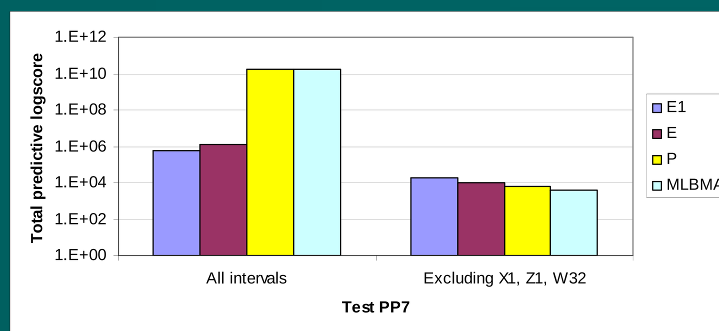
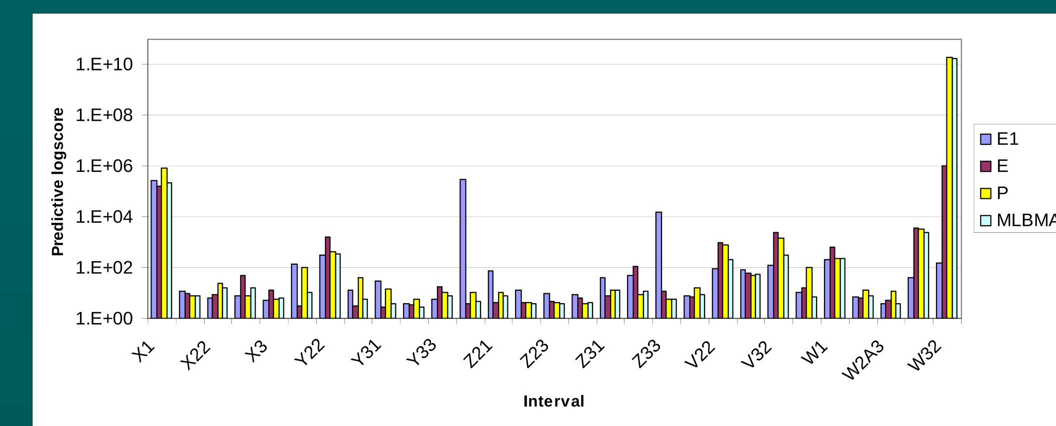
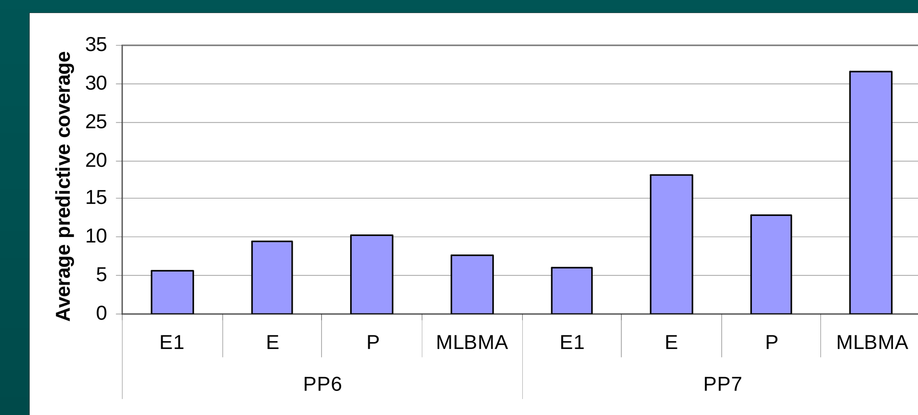


Figure. Predictive logscore, test **PP7**



Predictive coverage

The percentage of observed data that fall within a given prediction interval around average predicted pressure.



Conclusions

- Use of Occam's window led to selecting the model with the lowest fitting error with probability of close to 1 and disregarding all remaining models.
- A variance window gave more evenly distributed posterior probabilities based on KIC. Doing the same based on AIC, AICc or BIC led to one model being assigned a posterior probability of about 1.
- The results of the calibration were validated against an independent data set. Best results were obtained with a model ranked second by KIC but very low by AIC, AICc and BIC.
- Predicted pressures using MLBMA were less accurate than those obtained with some individual models because the individual model with the largest posterior probability was the worst or second worst predictor in both validation cases. In terms of predictive coverage, MLBMA was far superior to any of the individual models in one validation test and second to last in the other validation test.
- We attribute these mixed results to inability of any of our models to capture in a satisfactory manner the complex nature of the ALRS fractured rock system and pressure distribution in it with the available data.

Acknowledgements: This work was supported jointly by U.S. National Science Foundation grant EAR-0407123, the U.S. Department of Energy through a contract with Vanderbilt University under (CRESP) III, and by the International Research Training Group NUPUS.

Reference: E. Morales-Casique, S. P. Neuman, V.V. Vesselinov, 2009. Maximum likelihood Bayesian model averaging of air flow models in unsaturated fractured tuff using Occam and variance windows, *Stochastic Environmental Research and Risk Assessment*, (Accepted)