MIT EARTH RESOURCES LABORATORY
ANNUAL FOUNDING MEMBERS MEETING 2018

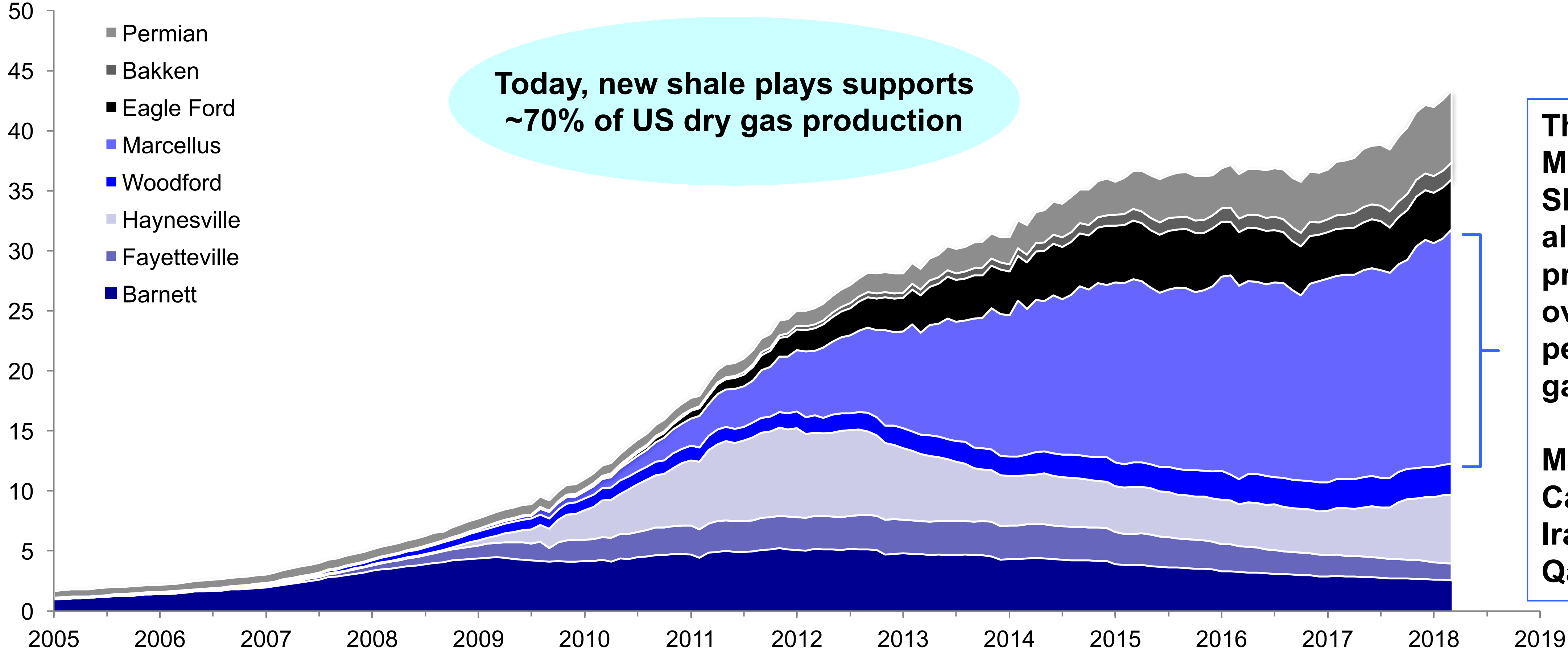
Forecasting shale gas production using a hierarchical Bayesian model

Justin B. Montgomery
PHD CANDIDATE [COMPUTATIONAL SCIENCE AND ENGINEERING]

In collaboration with Francis O'Sullivan and John Williams

Shale gas has become an extremely important resource and continues to alter the energy landscape – Accurately forecasting future levels of production is critical for government and industry

Illustration of gas production growth from the main U.S. shale plays since 2005
Bcf of gas per day



The Marcellus Shale alone now produces over 7 Tcf per year of gas.

More than Canada, Iran or Qatar

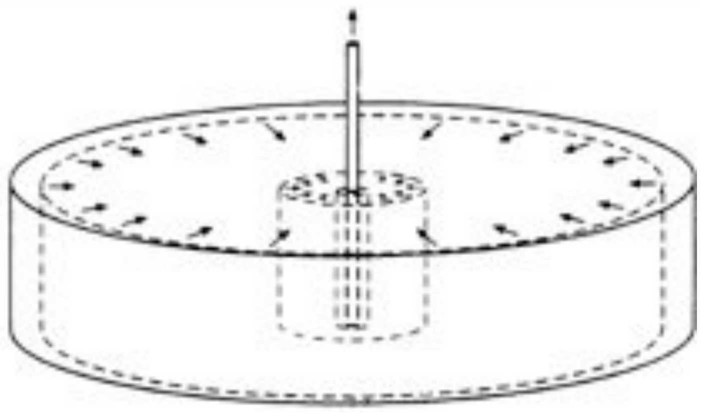
Source: F. O'Sullivan, United States Energy Information Administration, HPDI Production Database

Traditional approaches to forecasting production for conventional reservoirs are not appropriate for shale gas

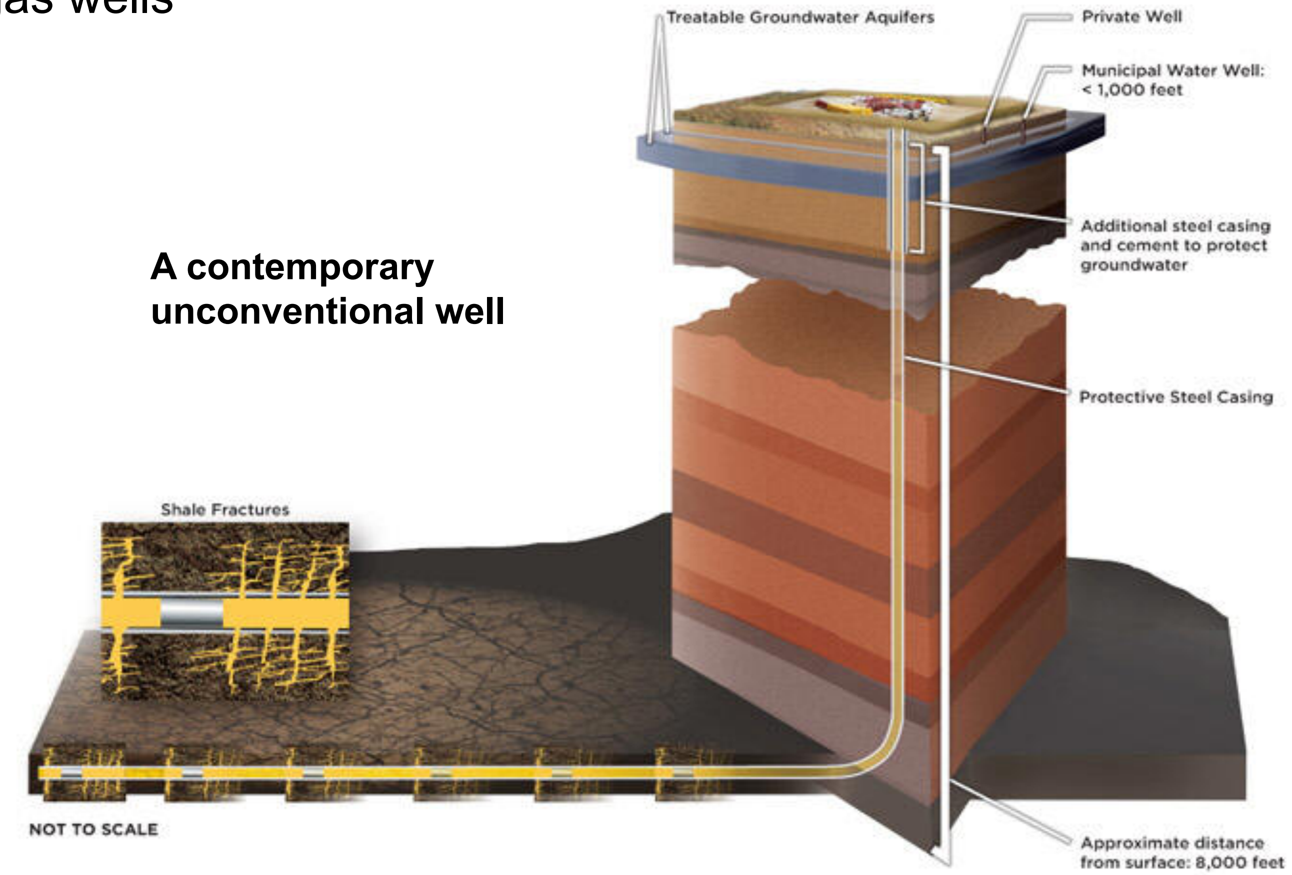
- **Too many unknowns for reservoir simulations**
  - Limited geological data
  - Complex and poorly understood behavior (e.g. fracture propagation and nanoscale flow)

- **Arps' decline curve forecasting has been widely applied to shale wells but is unreliable**
  - Not physically reasonable for horizontal shale gas wells

Radial flow geometry suitable for Arps' decline



A contemporary unconventional well



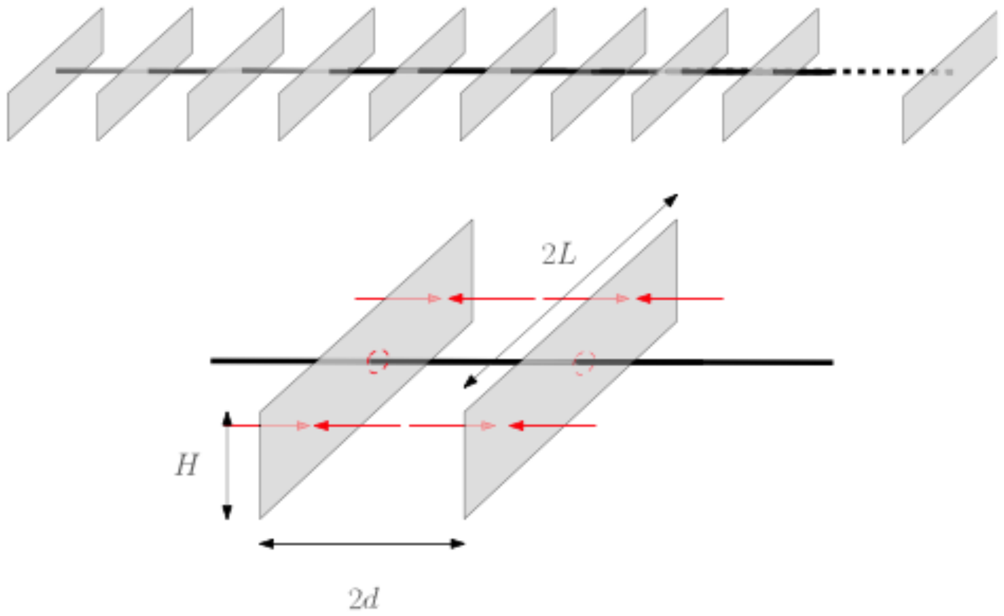
An empirical "scaling curve" with physical basis was introduced by Patzek et al. (2013)

Assumptions:
- Planar fractures with infinite conductivity
- Single phase (gas)
- Darcy flow from zone between fractures

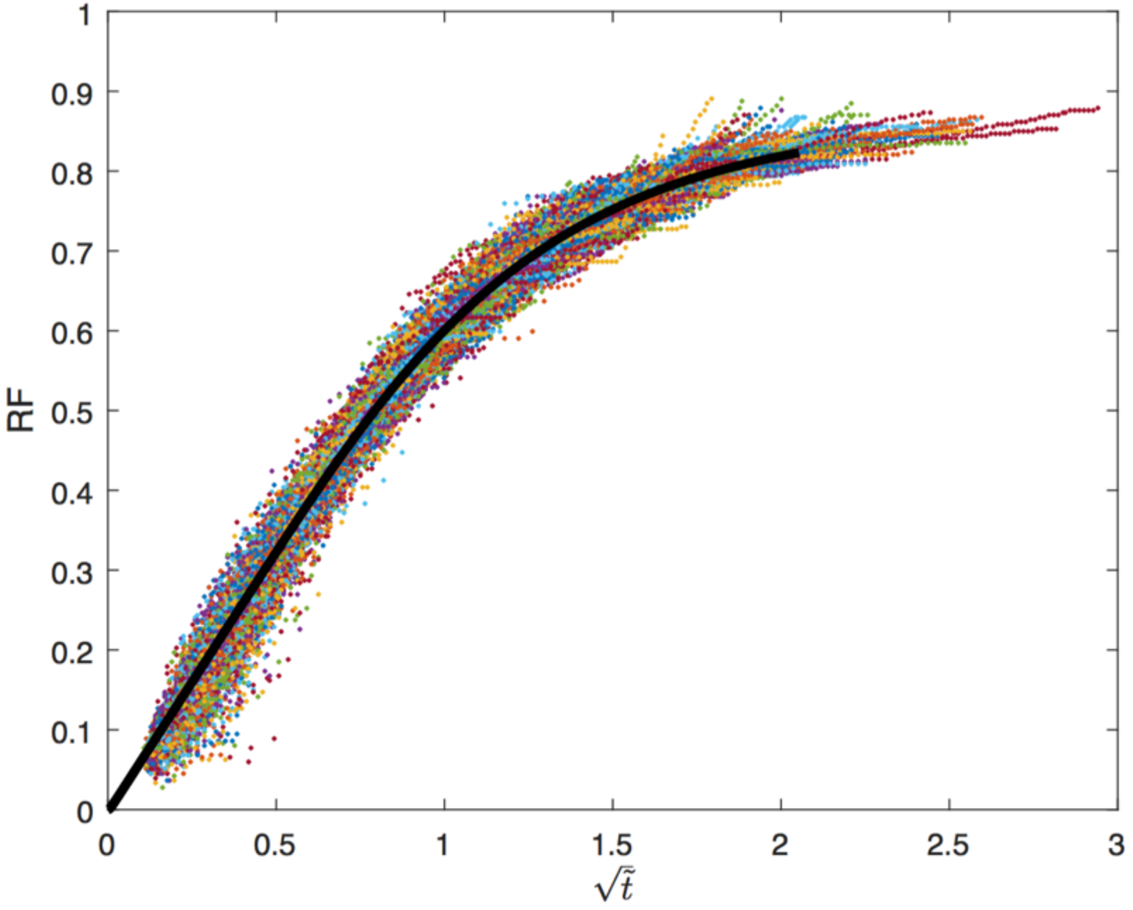
$$m(t) = \mathcal{M}\ RF(t/\tau)$$

Gas in place

Interference time



Correspondence of over 3000 wells in Barnett shale to **scaling curve**



Source: Patzek, T.W., Male, F., & Marder, M. (2013). Gas production in the Barnett Shale obeys a simple scaling theory. *PNAS*, 110 (49), 19731-19736.

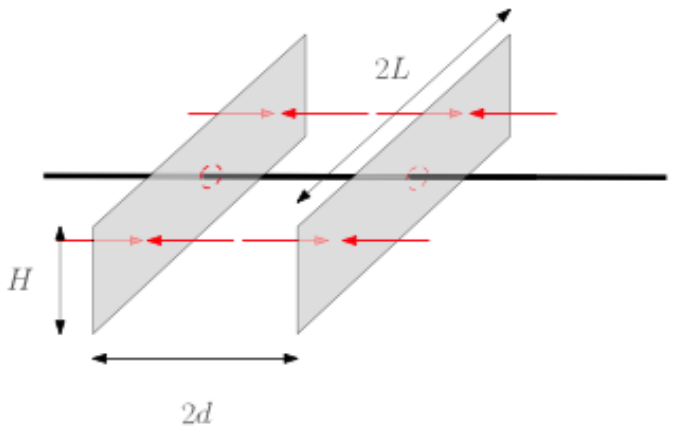
Patzek's scaling curve is unreliable for early life forecasts due to uncertainty about parameters

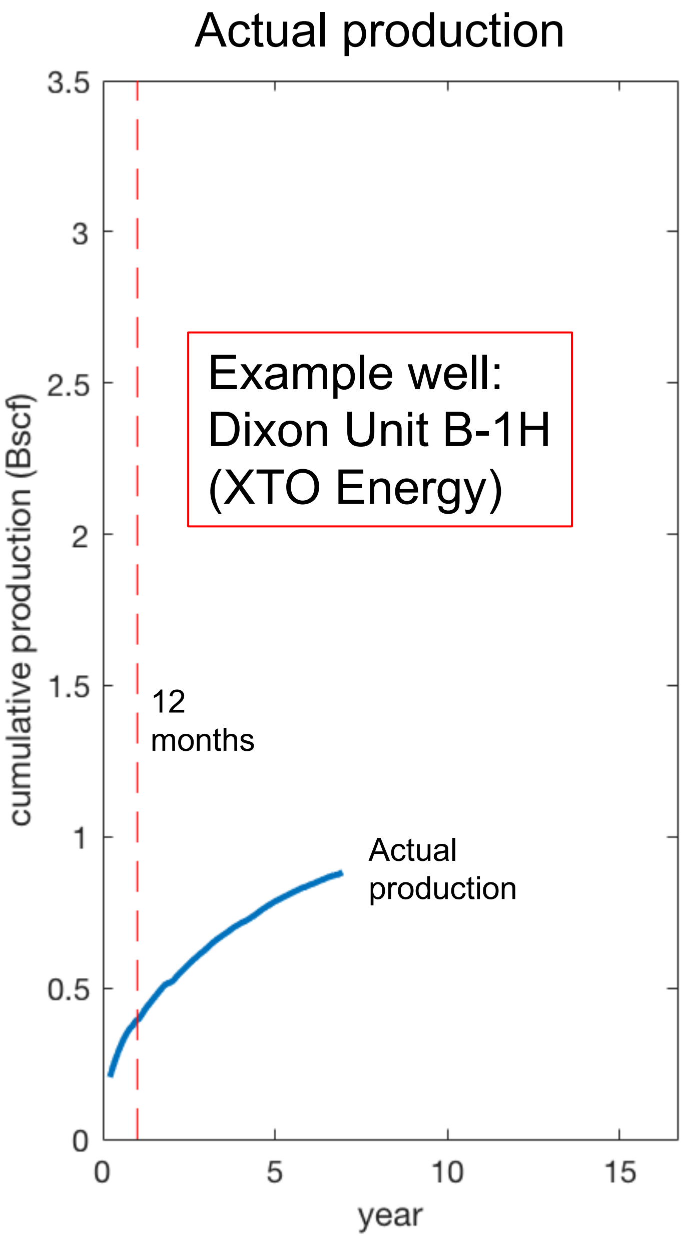
From derivation

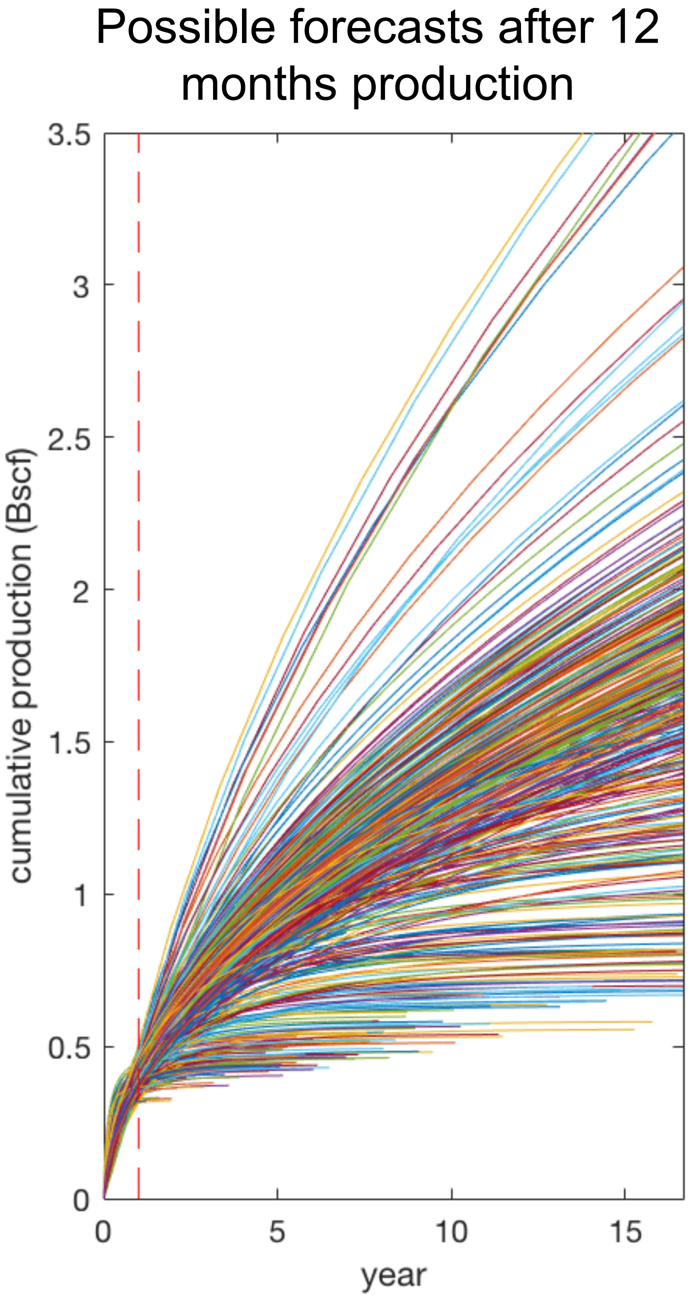
- **Gas in place ($\mathcal{M}$) –** Fracture length
- **Interference time (τ) –** Effective (enhanced) permeability

From curve fitting

- **Ambiguity between relative rate of depletion and total producible amount**

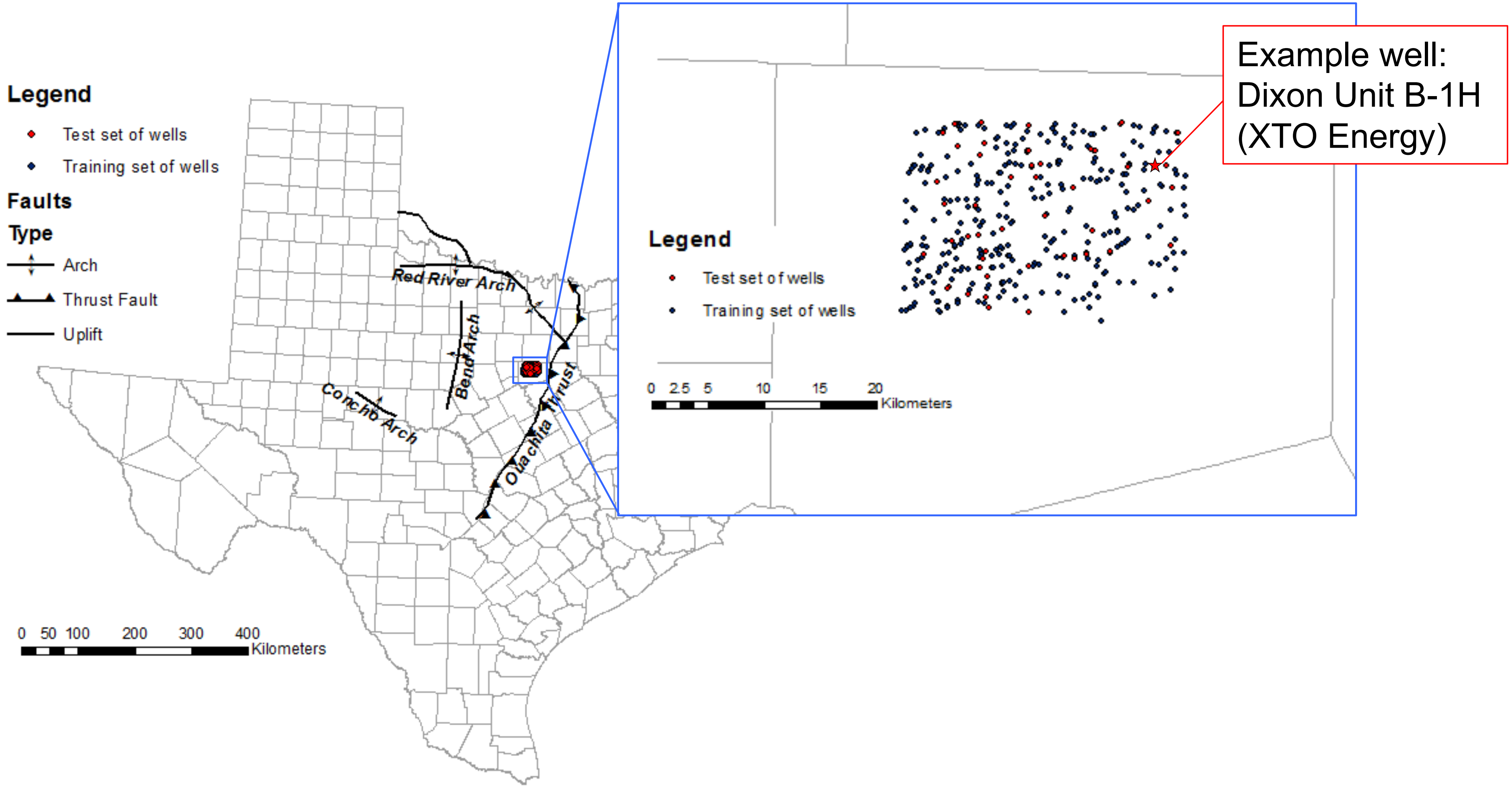






To improve reliability of forecasts we propose using a Bayesian regression framework for Patzek's scaling curve

In many shale plays, there is now abundant historical production data from existing wells
→ We want to incorporate this into a prior for the scaling curve parameters in a new (early-life) well



To carry out Bayesian regression we use Markov chain Monte Carlo (MCMC) to draw samples that approximate the posterior distribution

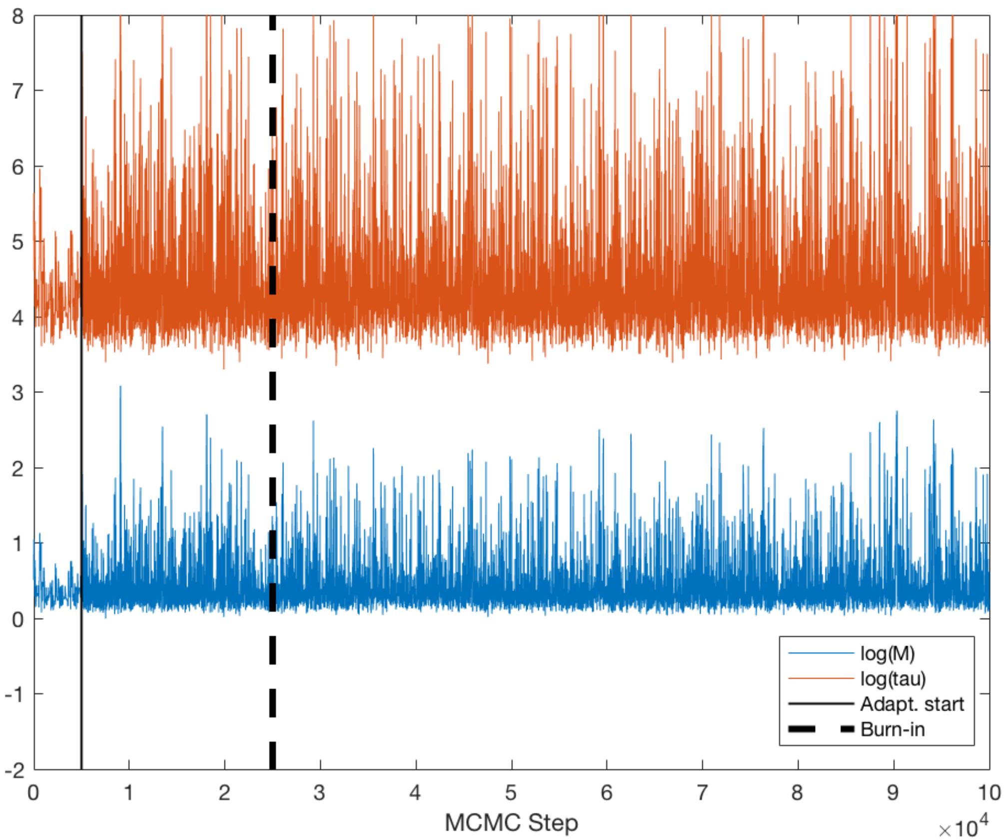
Can include useful information in prior

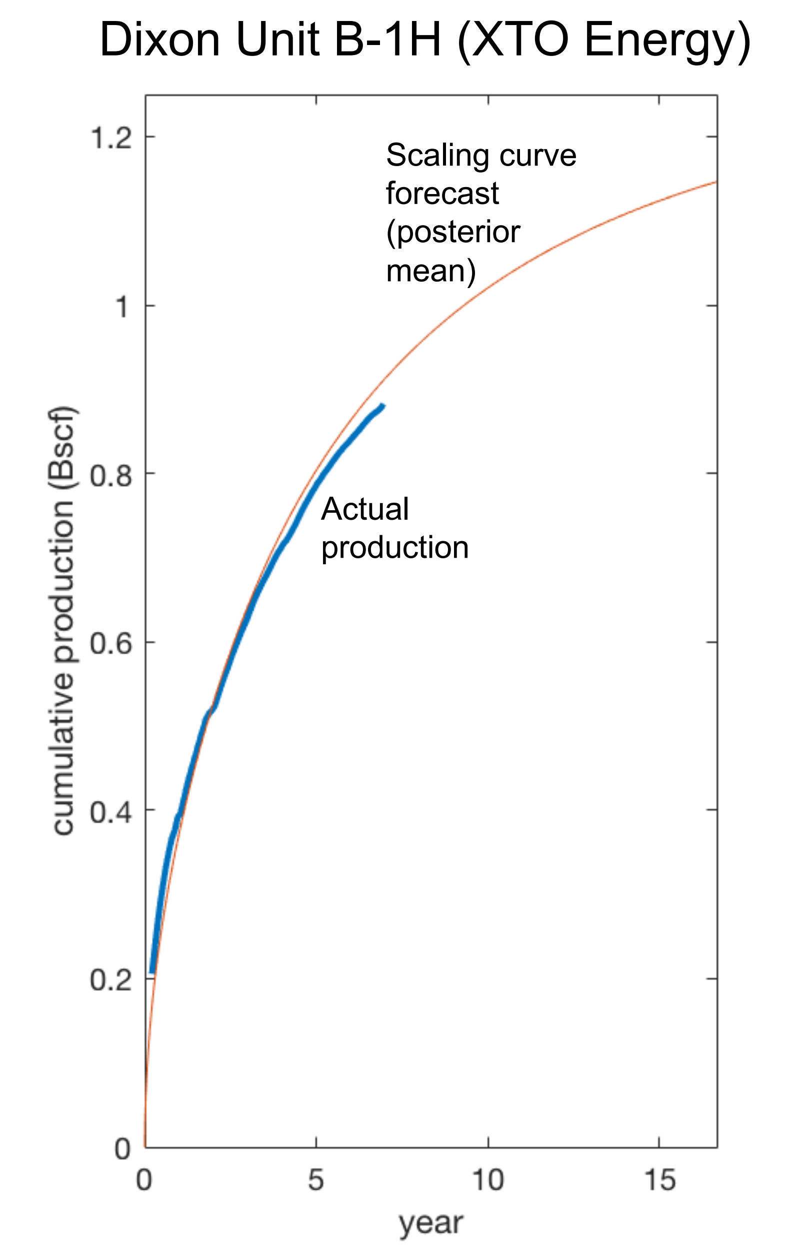
Bayes rule: $\mathbb{P}(\mathcal{M},\tau|\boldsymbol{m}) \propto \mathbb{P}(\boldsymbol{m}|\mathcal{M},\tau)\mathbb{P}(\mathcal{M},\tau)$

Gaussian noise for likelihood: $\mathbb{P}(\boldsymbol{m}|\mathcal{M},\tau) \propto \prod_{t=1}^{T} \mathcal{N}(m_t; \mathcal{M}\ RF(t/\tau), \sigma_\varepsilon{}^2)$

MCMC algorithm

1. Initialize chain (randomly)
2. Propose new state with Gaussian ‘step’ (Metropolis)
3. Accept new state with probability $\min(1, \frac{\Pi_{prop}}{\Pi_{curr}})$
4. After 5000 steps, use covariance of accepted samples



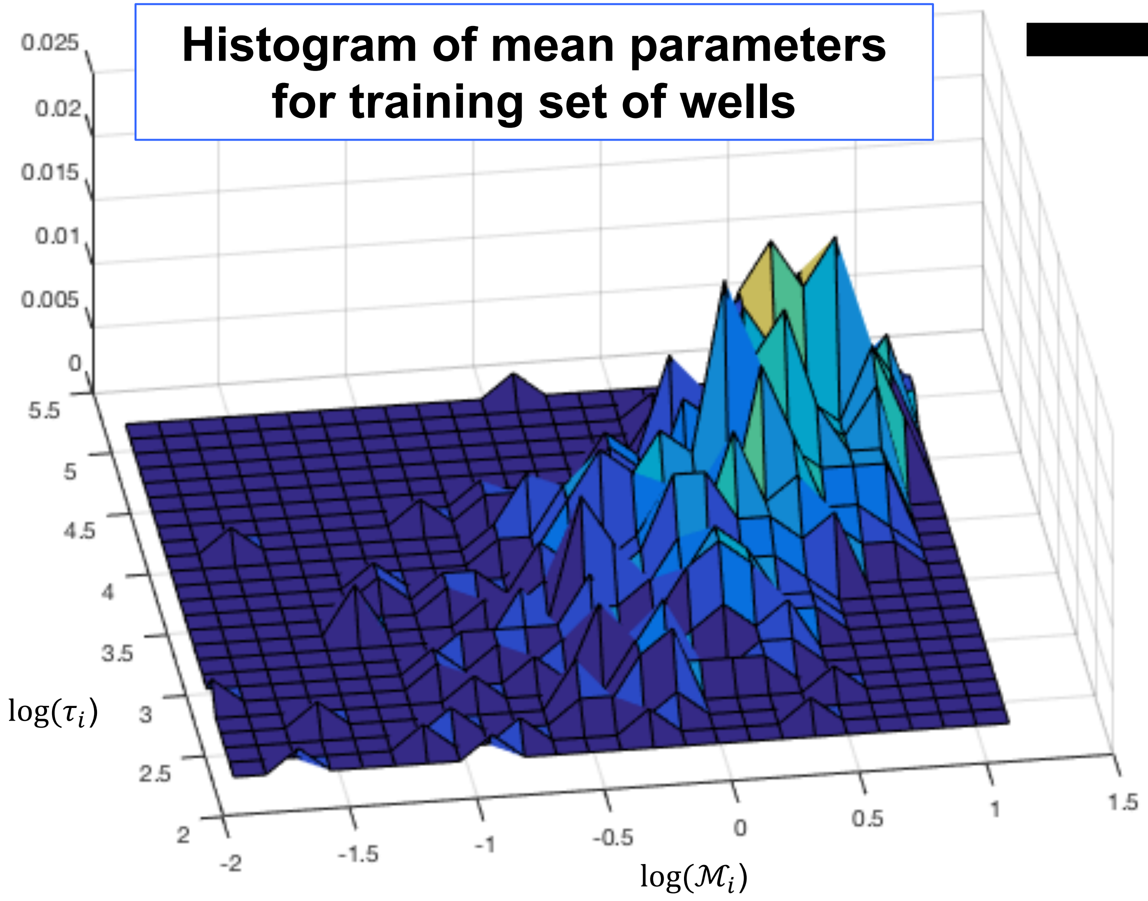


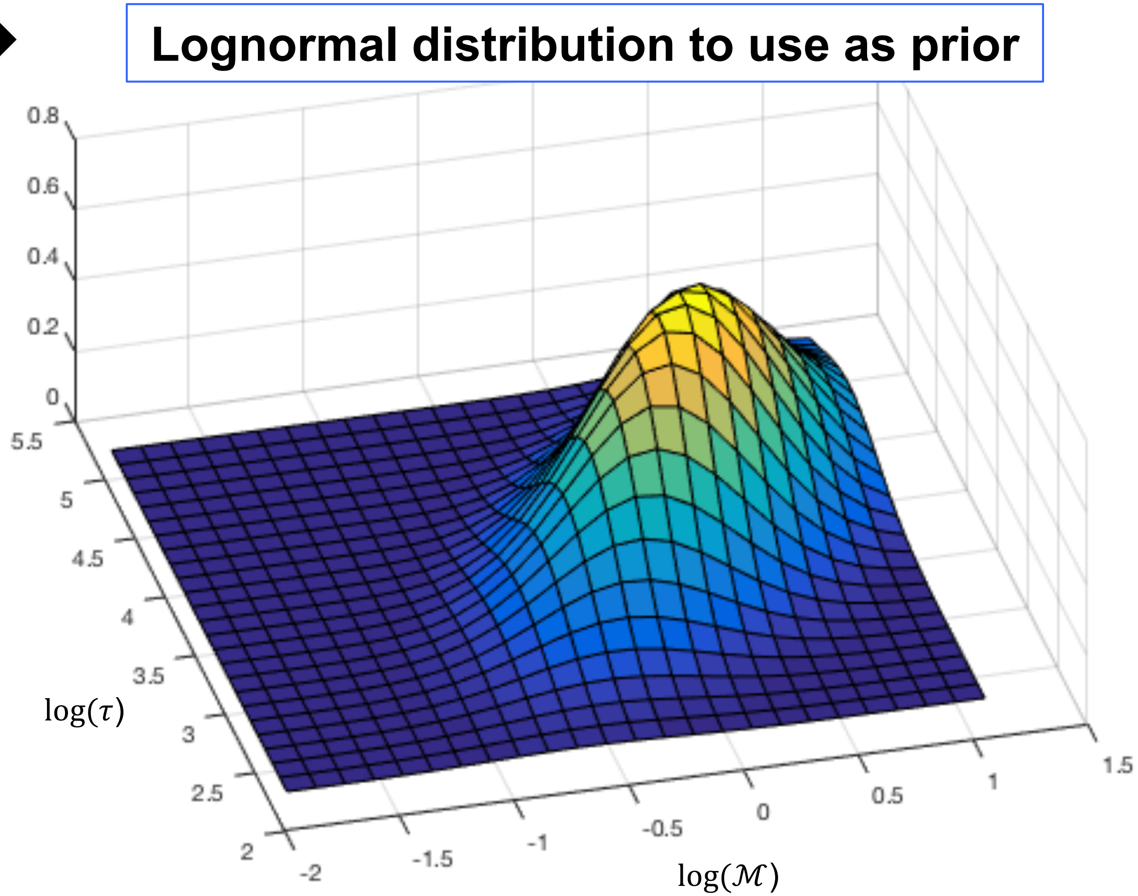
To improve the reliability of early-life forecasts we develop a prior approximating the distribution of posterior mean parameters for all training set wells (entire production history)

$$\mu_{\mathcal{M}} = \frac{1}{N}\sum_{i}^{N} \log(\mathcal{M}_i) \quad \text{(likewise for } \tau\text{)}$$

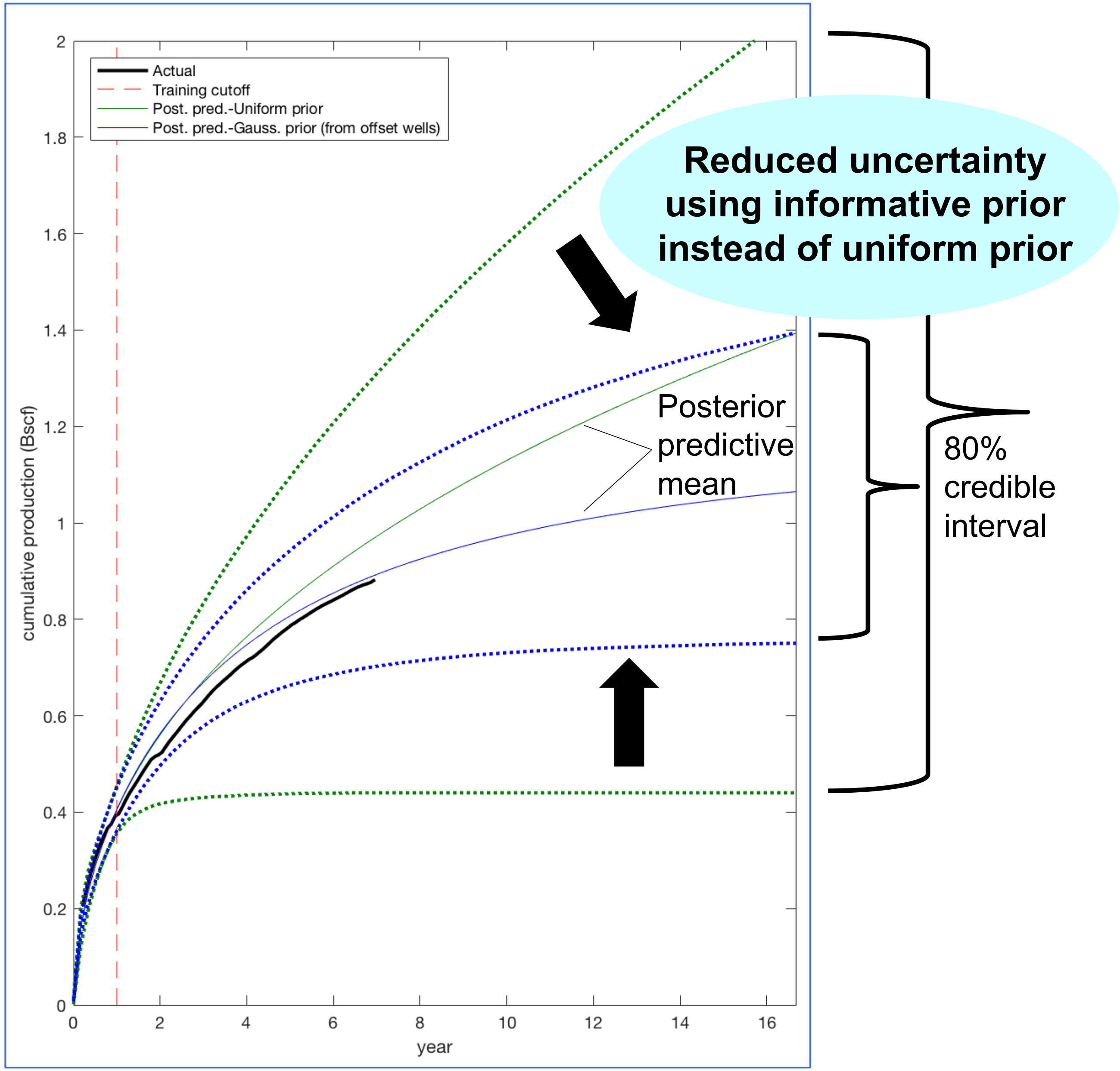
$$\mathbb{P}(\mathcal{M}, \tau) = \mathrm{Lognormal}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\Sigma} = \mathrm{cov}(\log(\boldsymbol{\mathcal{M}}), \log(\boldsymbol{\tau}))$$



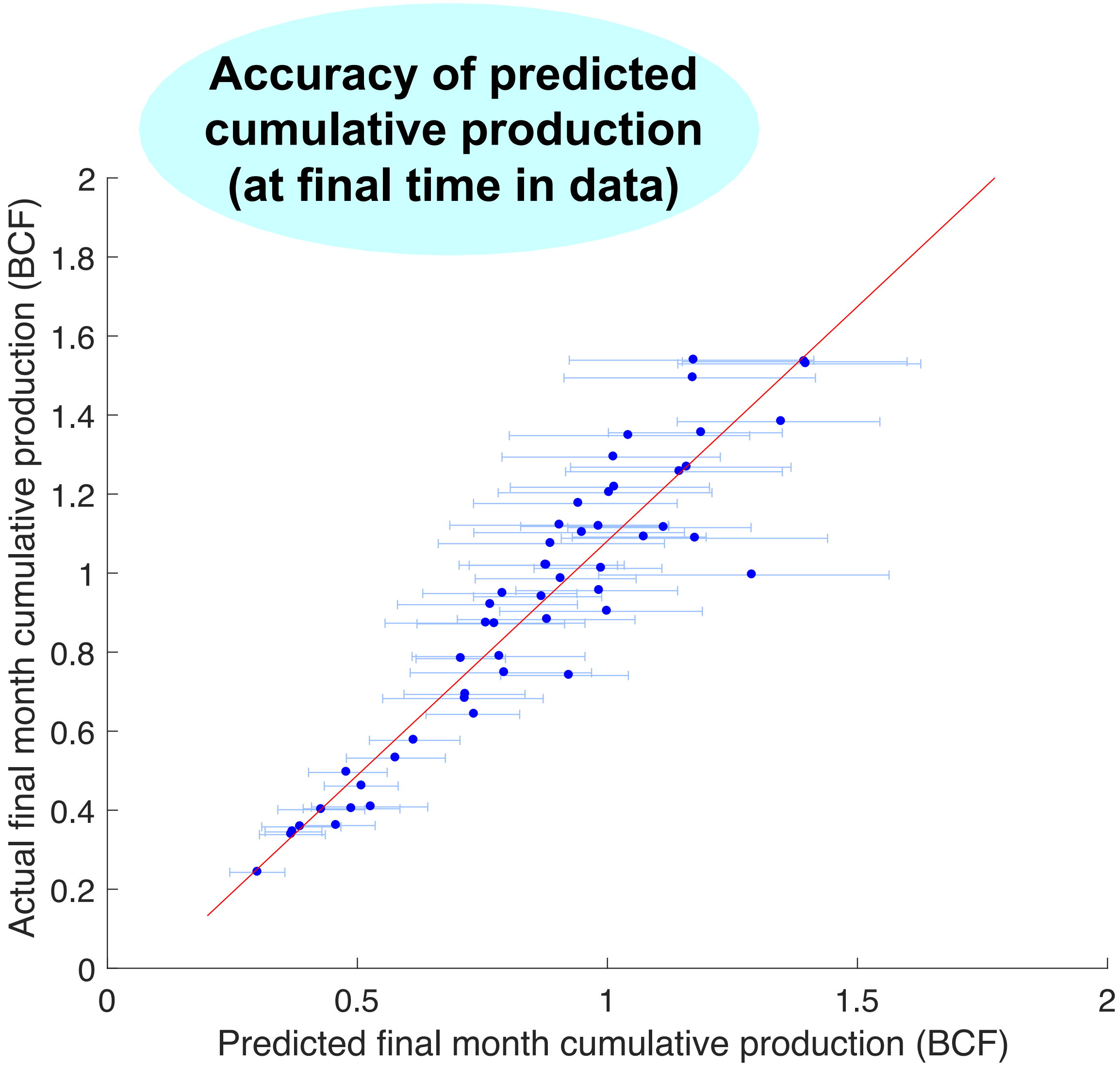


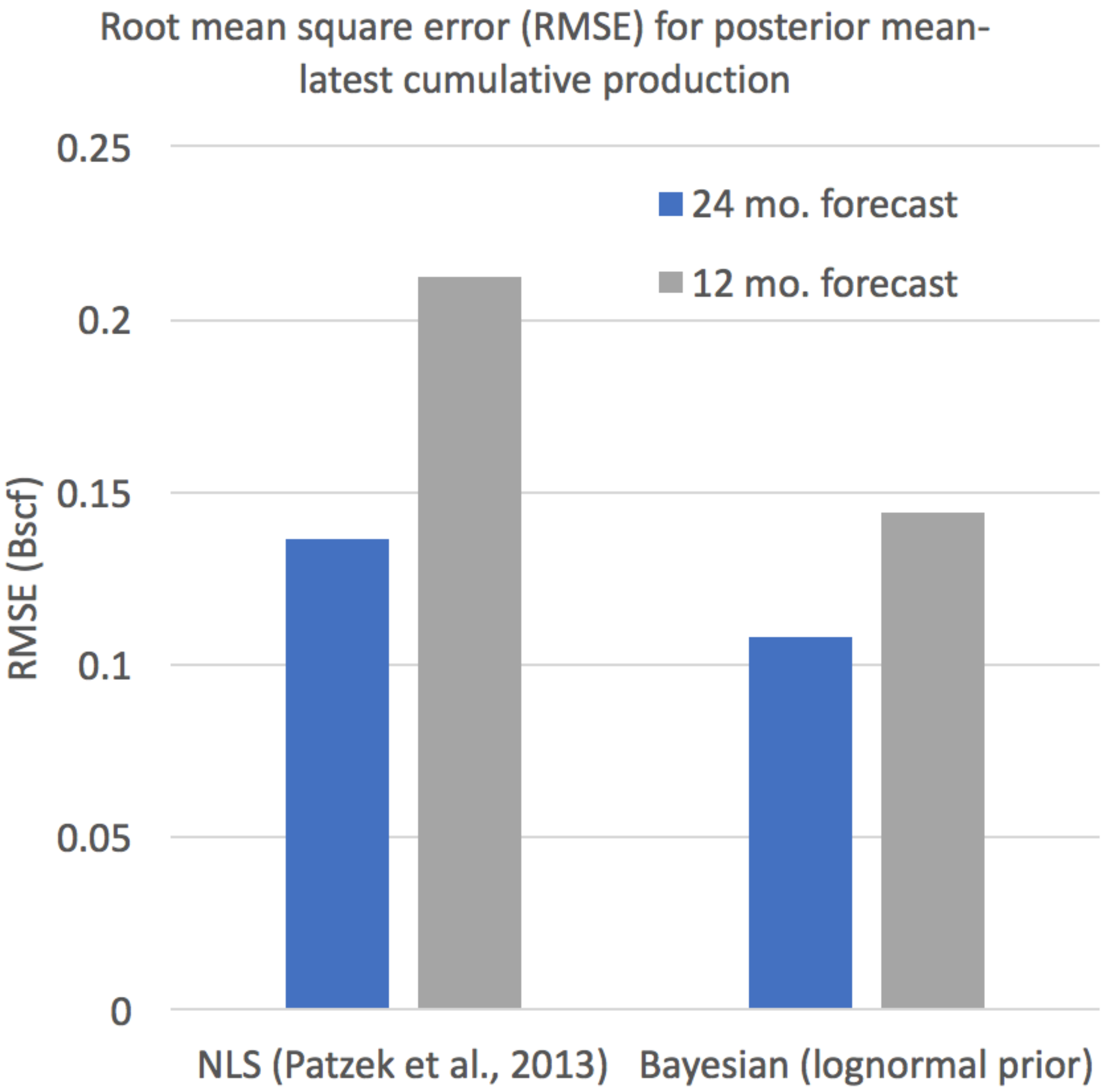
By incorporating offset well information into the prior, we can reduce the uncertainty of early life production forecasts



Dixon Unit B-1H
(XTO Energy)

Our Bayesian implementation of Patzek's scaling curve substantially improves the accuracy of early life forecasts – We expect further improvements as we extend hierarchy to include correlations to geology, completion design, and spatial patterns





Other research: Machine learning model that predicts the impact of design choices on resulting well productivity/economics across a resource play

→ Shalestats.com is an interactive web application to explore the "break-even oil price" for different well designs, economic parameters, and locations (currently for Williston Basin)

How to use:

1. Adjust model parameters (or use default settings)
2. Click to choose a location in the Bakken (shaded map area)
3. Click 'Run the model' button and wait to view simulation results

Please send questions and feedback to Justin Montgomery: jbm@mit.edu

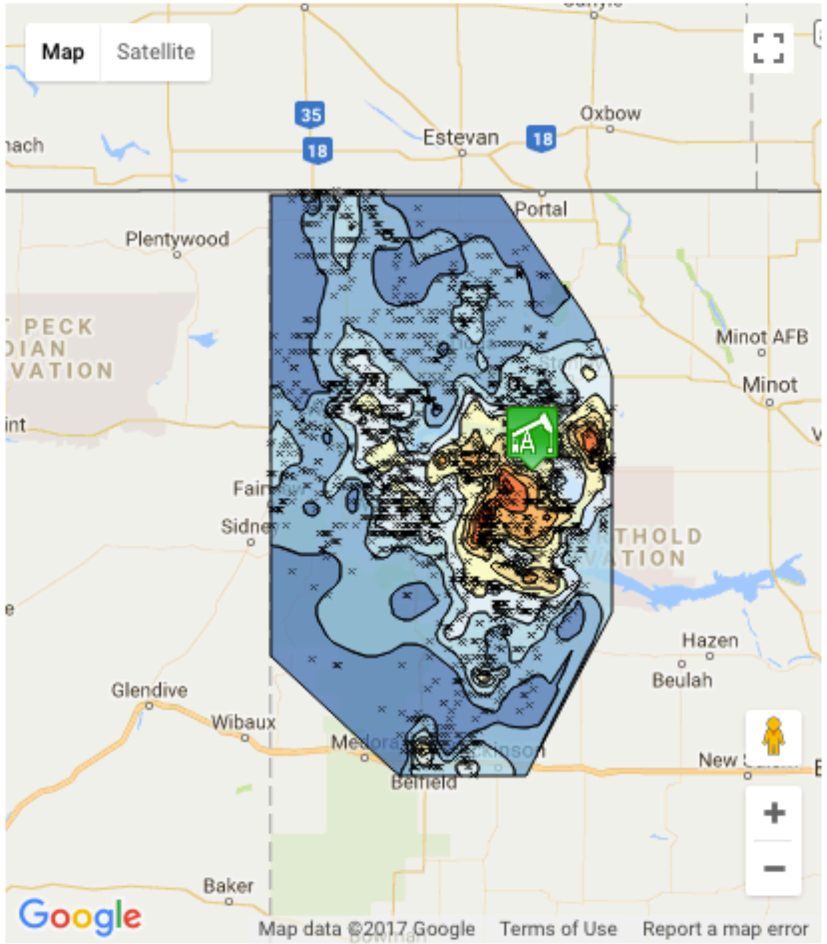
1. Adjust model parameters:

Cost of capital

Royalty/Taxes

Model input variables

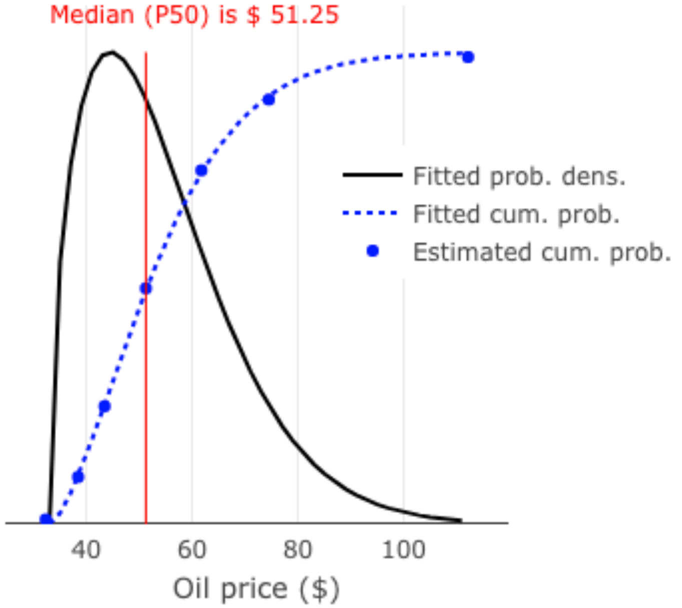
2. Choose a location:



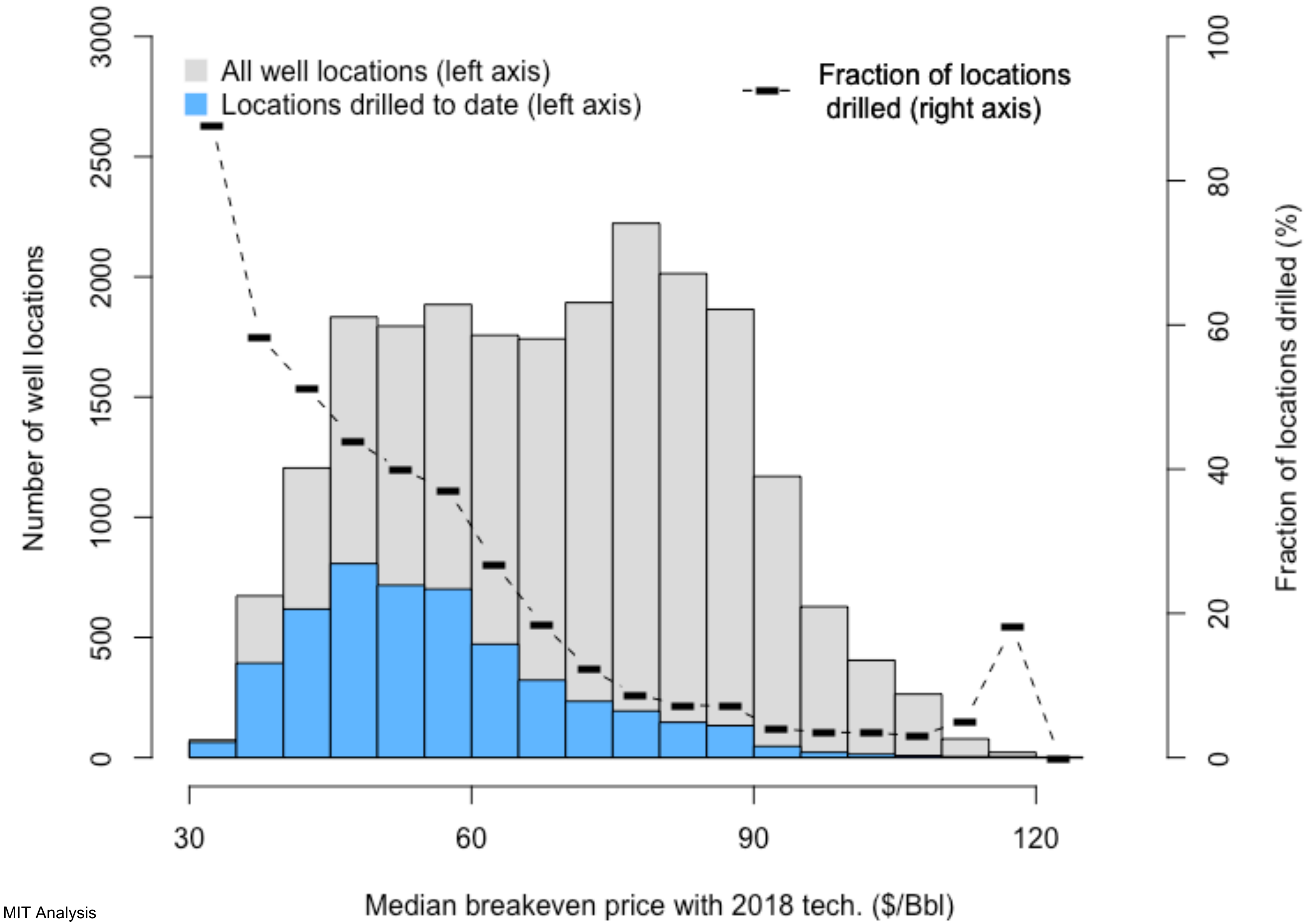
3. Run the model

Results:

Probability of breaking even (well NPV>=0) at different oil prices



This tool is helping us to understand the dynamics behind resource development costs – Drilling activity in the Bakken has been concentrated on the lowest cost (sweet-spot) areas



Source: MIT Analysis

Backup slides

Production forecasting with conventional vertical wells has historically been carried out using Arps' decline curve – derivable for radial transient flow

- Originally introduced as empirical model by Arps, 1944

- Fetkovich (1980) provided physical basis for model

- Led to well testing (inverse problem) for reservoir properties based on fit to analytical models

$$q = q_i(1 + bD_it)^{\left(-\frac{1}{b}\right)}$$

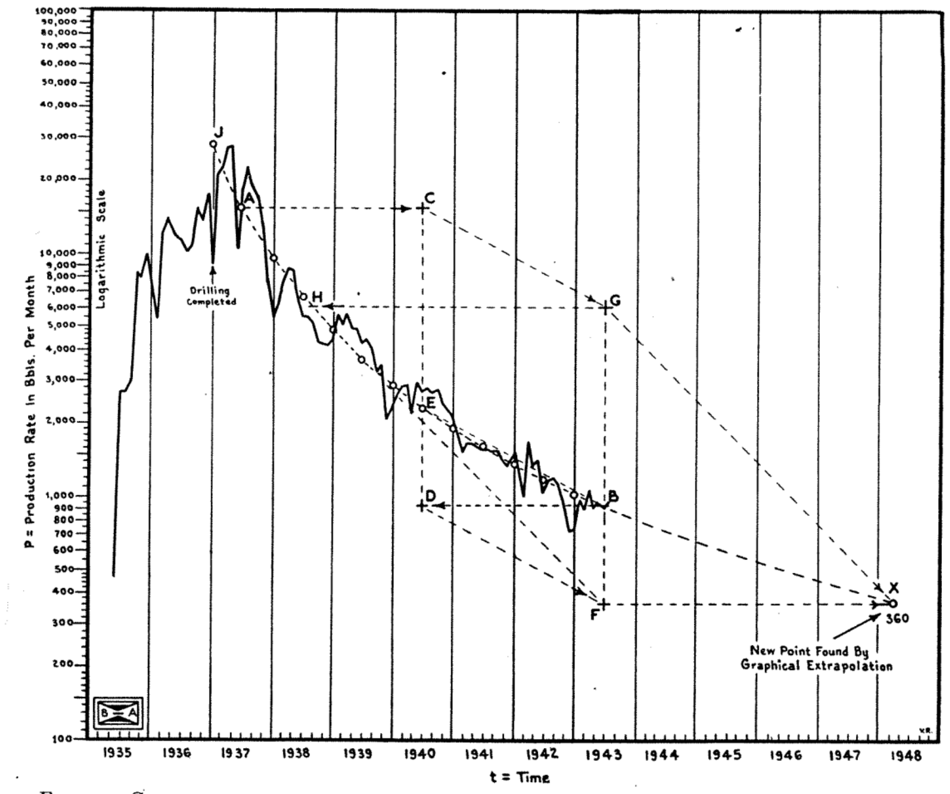
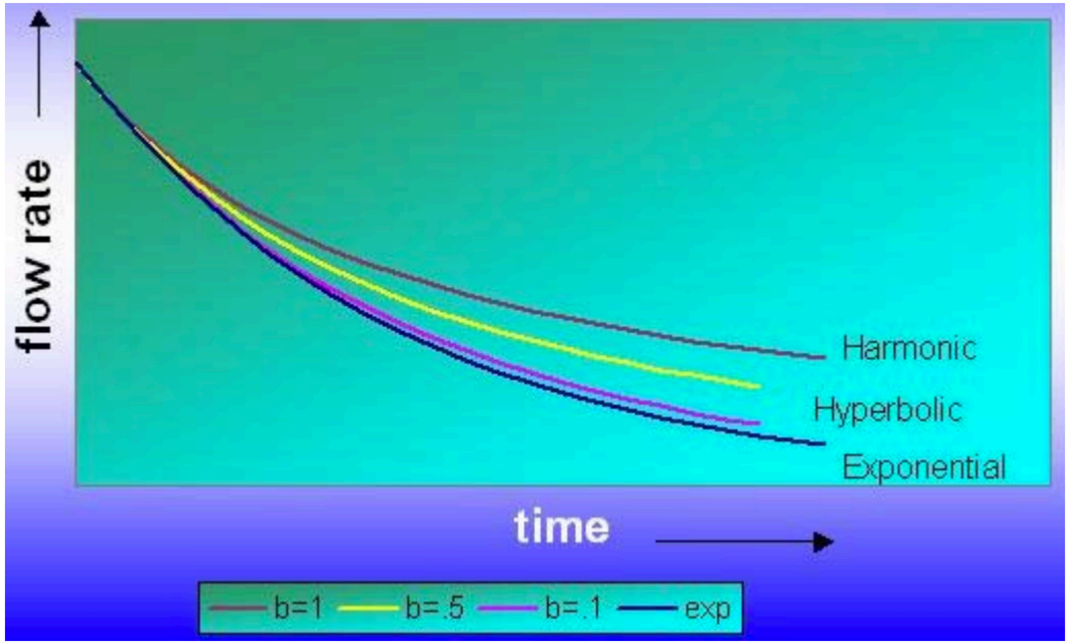


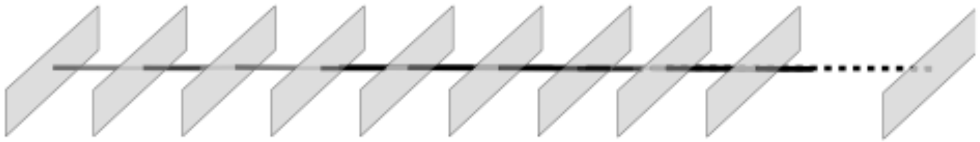
FIG. 3.—GRAPHICAL EXTRAPOLATION OF HYPERBOLIC RATE-TIME CURVE ON SEMILOG PAPER.

Source: Arps, 1945, Analysis of decline curves



Source: Fekete.com

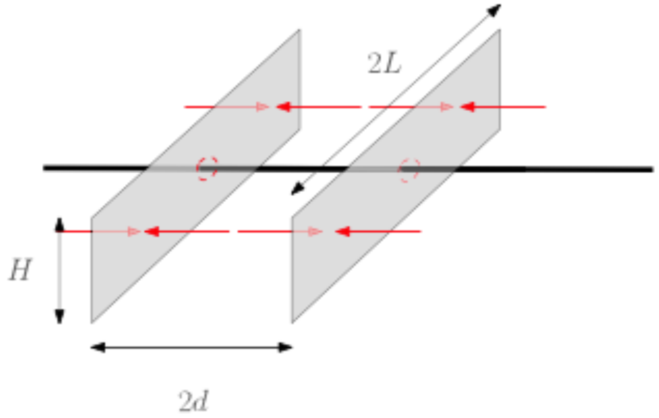
More on derivation of Patzek's model



Interference time: $\tau = d^2/\alpha_i$

Hydraulic diffusivity: $\alpha_i = \left.\frac{k}{\phi S_g \mu_g c_g}\right|_{\text{Initial reservoir } p,T}$

Dimensionless time: $\tilde{t} \equiv t/\tau$



Gas in place: $\mathcal{M} \equiv (N+1)4\rho_i LHd\phi S_g$

Numerically solve PDE: $\frac{\mathbf{m}}{\mathcal{M}} = \mathrm{RF}(\tilde{t}),$ where $\mathrm{RF}(\tilde{t}) \equiv \int_0^{\tilde{t}} d\tilde{t}' \left.\frac{\partial \tilde{m}}{\partial \tilde{x}}\right|_0 (\tilde{t}')$

$$\frac{\partial \tilde{m}}{\partial \tilde{t}} = \frac{\alpha}{\alpha_i}\frac{\partial^2 \tilde{m}}{\partial \tilde{x}^2},$$

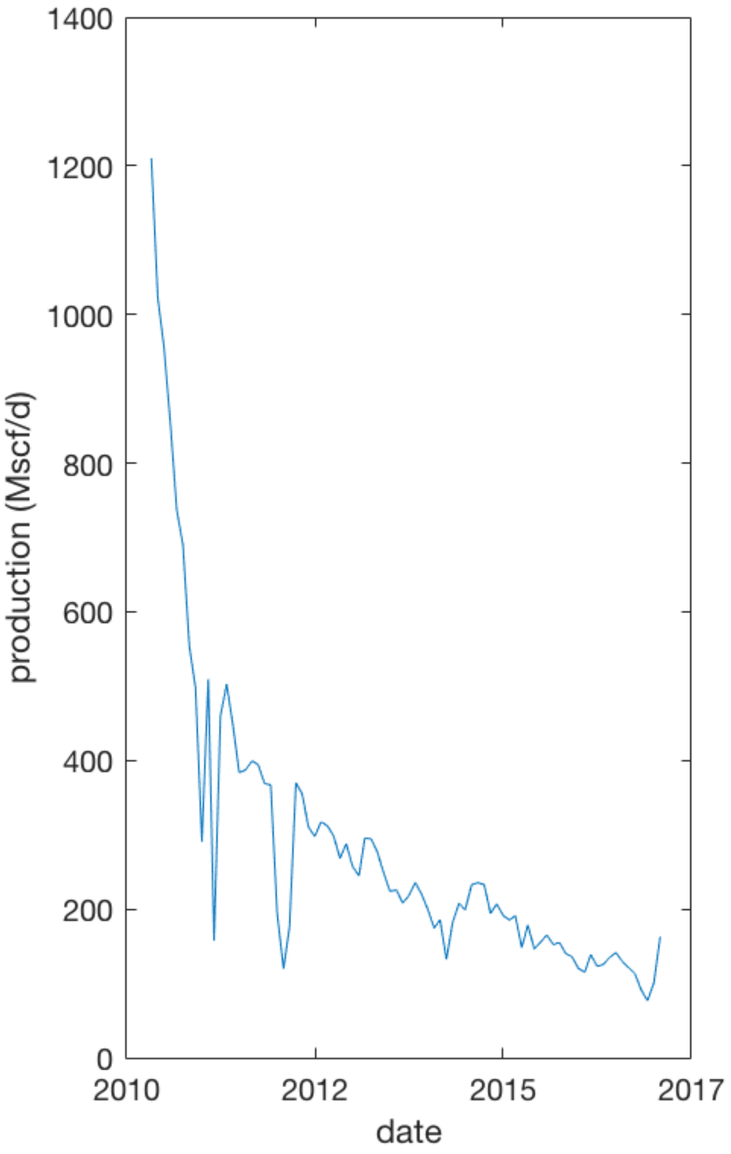
$$\tilde{m}(\tilde{x},\tilde{t}=0) = \tilde{m}_i(\tilde{x}),$$

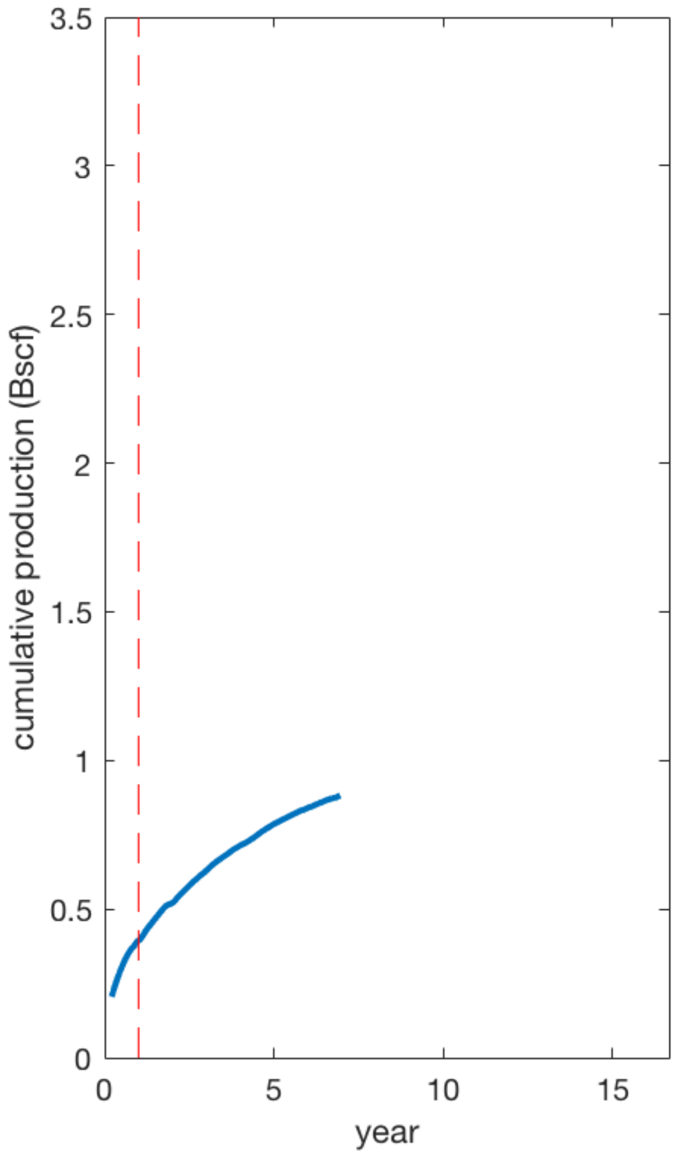
$$\tilde{m}(\tilde{x},\tilde{t}) = 0 \quad \text{for} \quad \tilde{x} = 0$$

$$\partial\tilde{m}/\partial\tilde{x} = 0 \quad \text{for} \quad \tilde{x} = 1.$$

$$\tilde{x} = x/d \qquad \tilde{m} = \frac{1}{2}\left([c_g p]\mu_g Z_g/p^2\right)_i m(x,t)$$

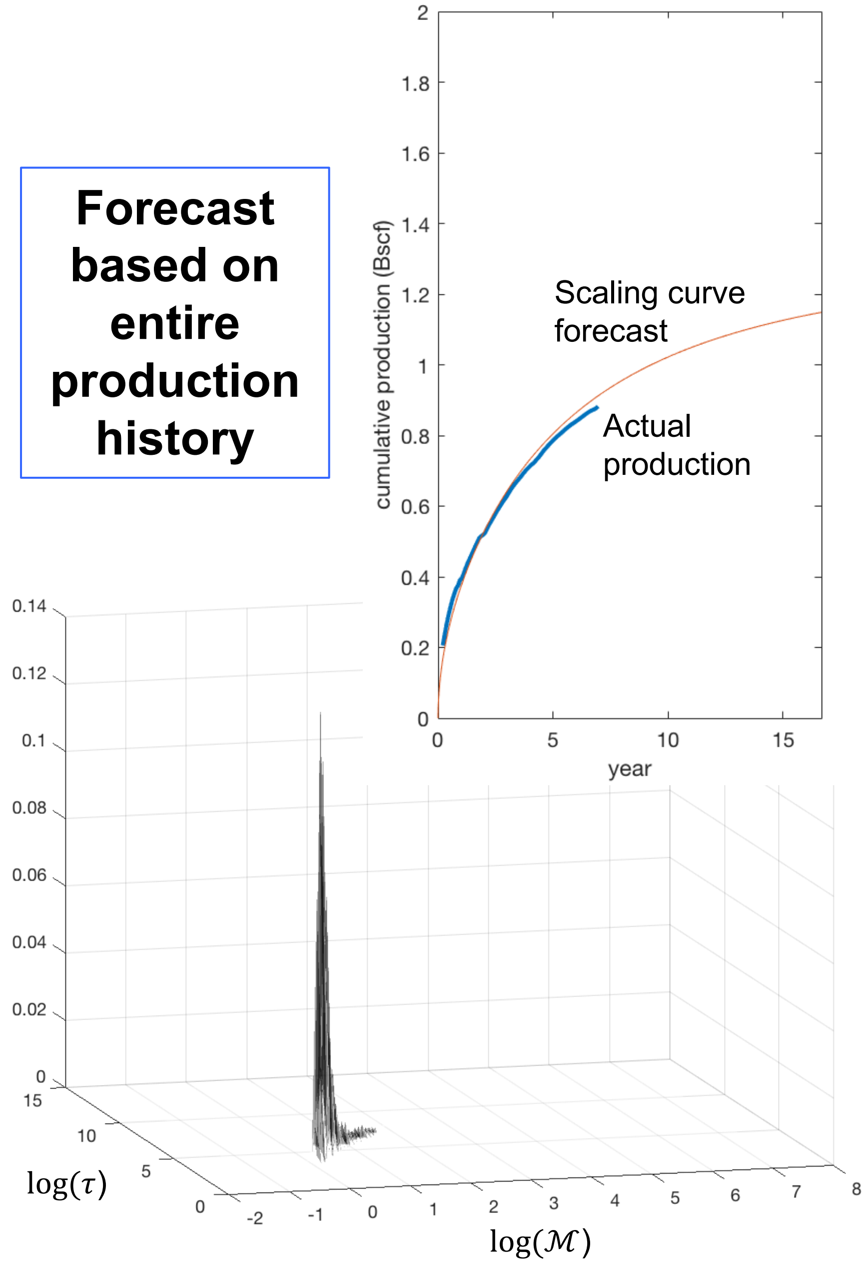
Example well: Dixon Unit B-1H (XTO Energy)



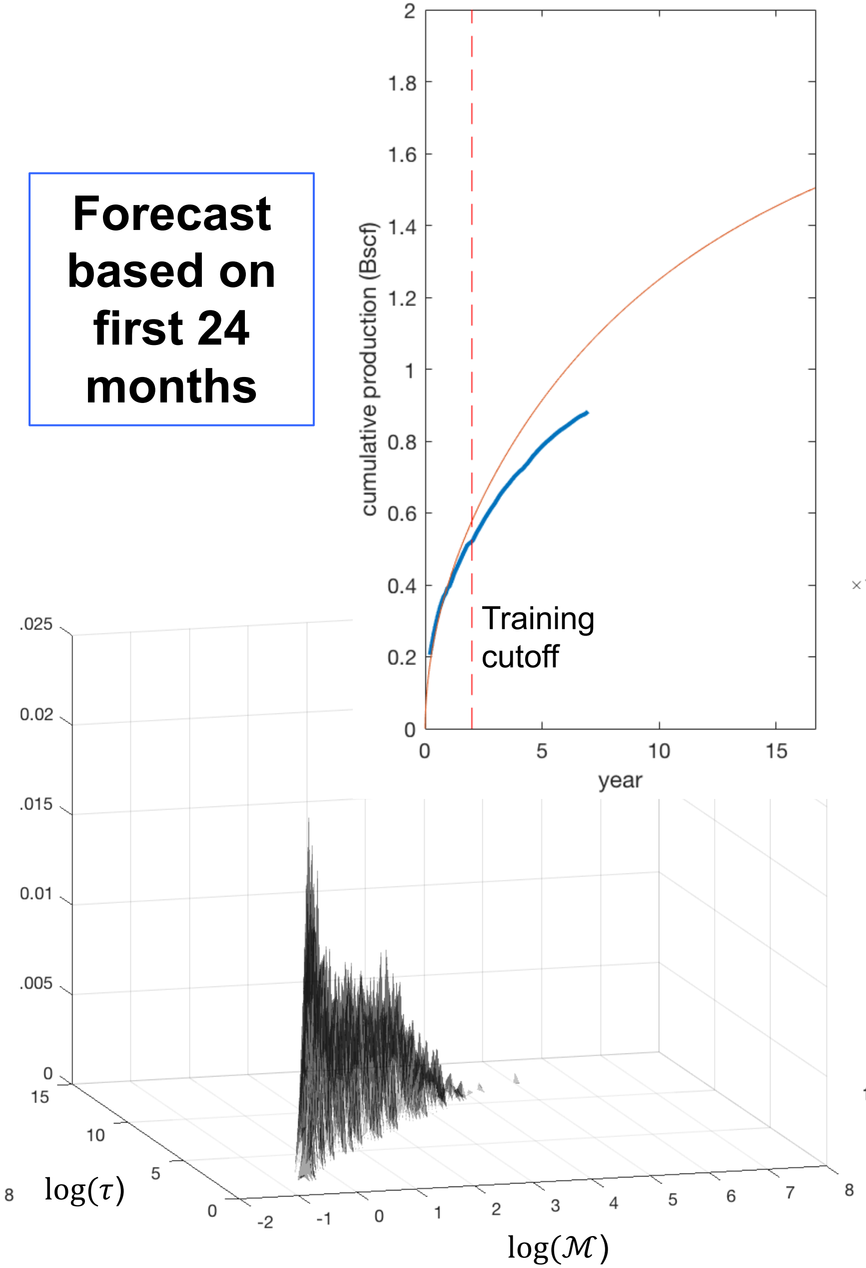


When forecasts are made earlier on, the posterior distribution widens and the posterior predictive mean forecast becomes less reliable

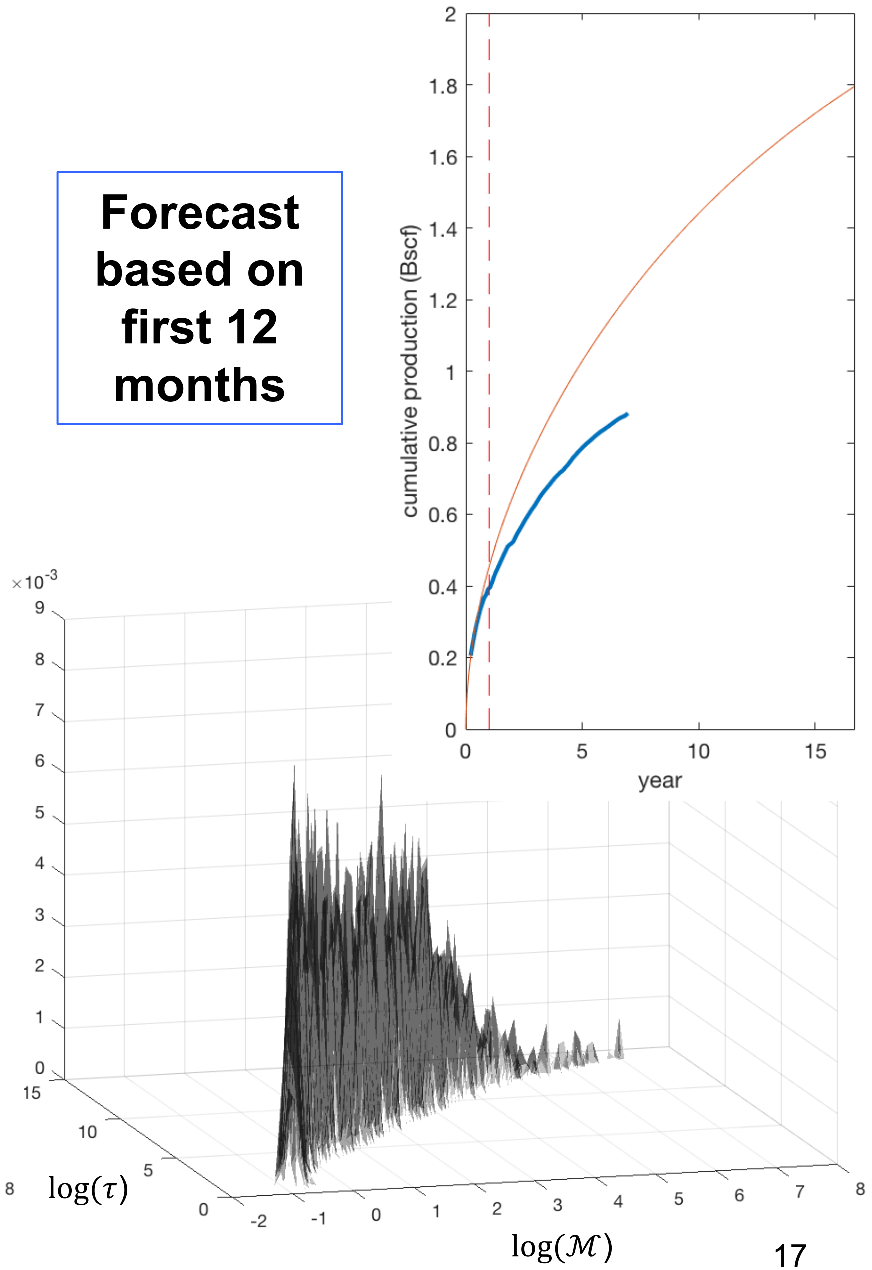
Forecast based on entire production history



Forecast based on first 24 months

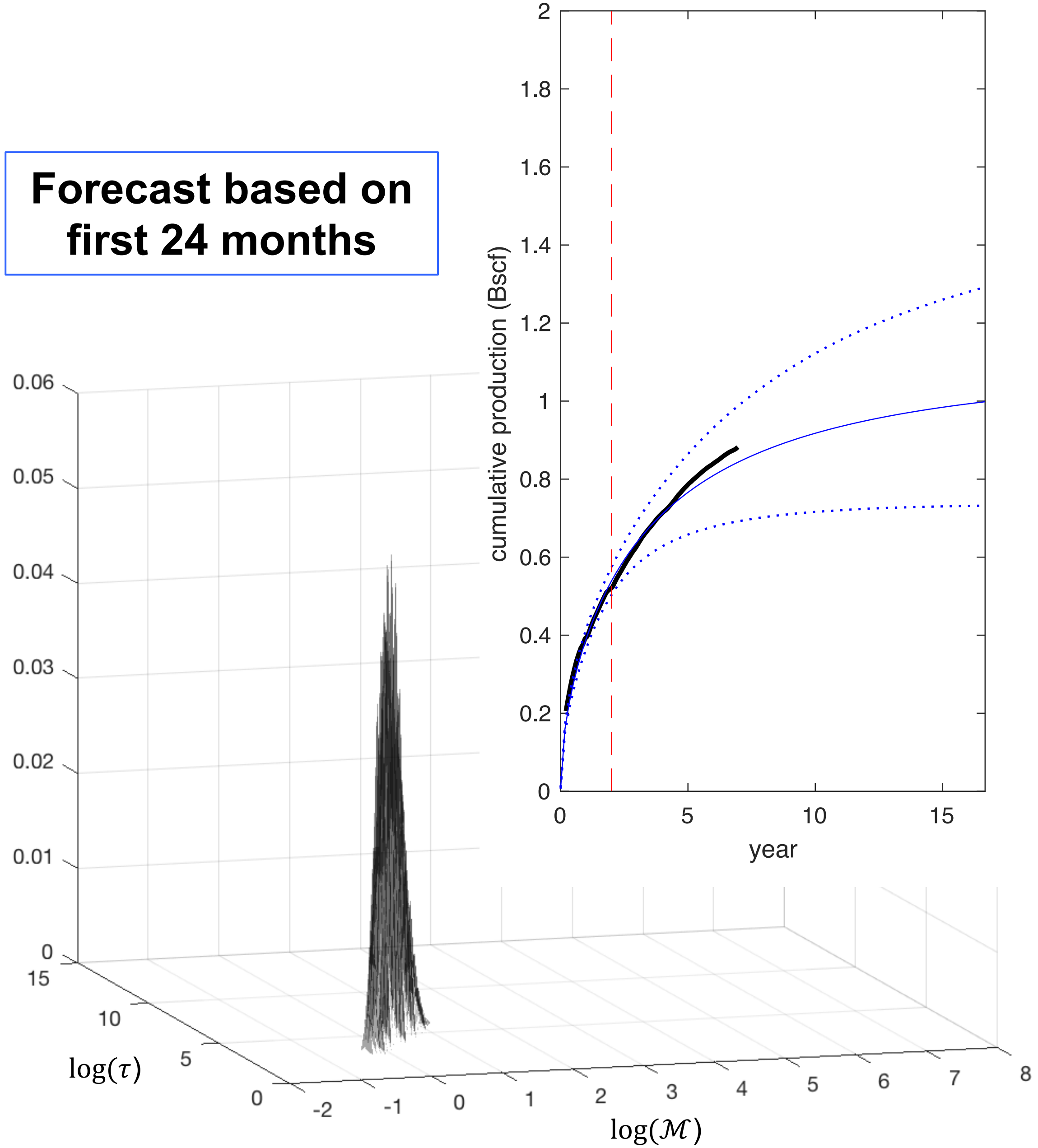


Forecast based on first 12 months

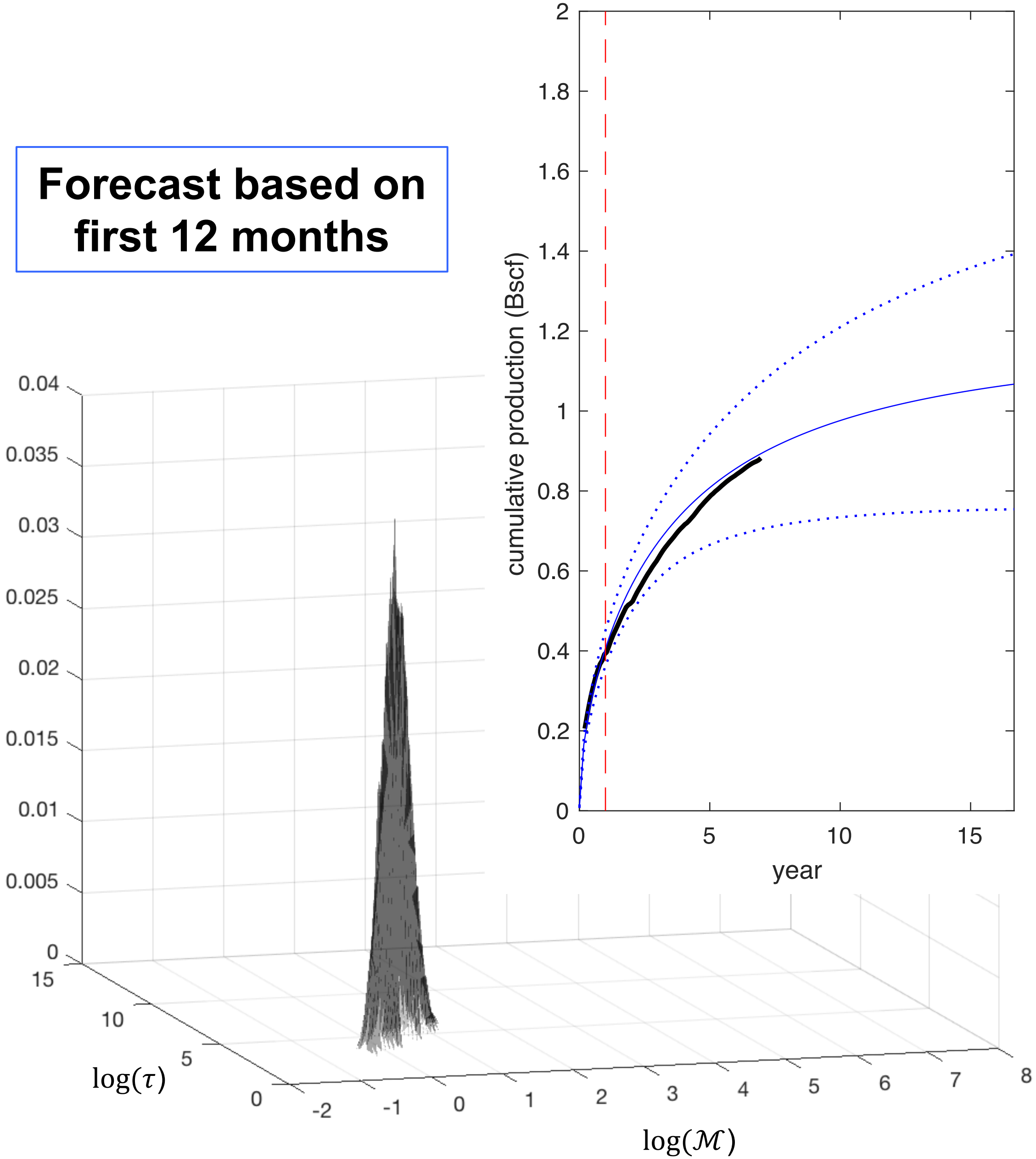


Incorporating this prior into the model helps restrict the parameter space of the posterior according to past observations and improves the forecast

Forecast based on first 24 months

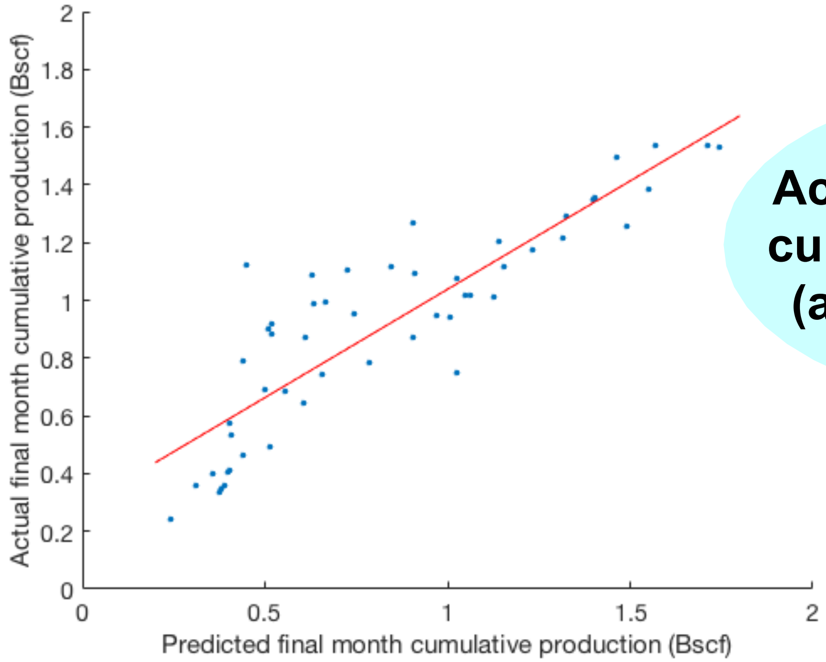


Forecast based on first 12 months



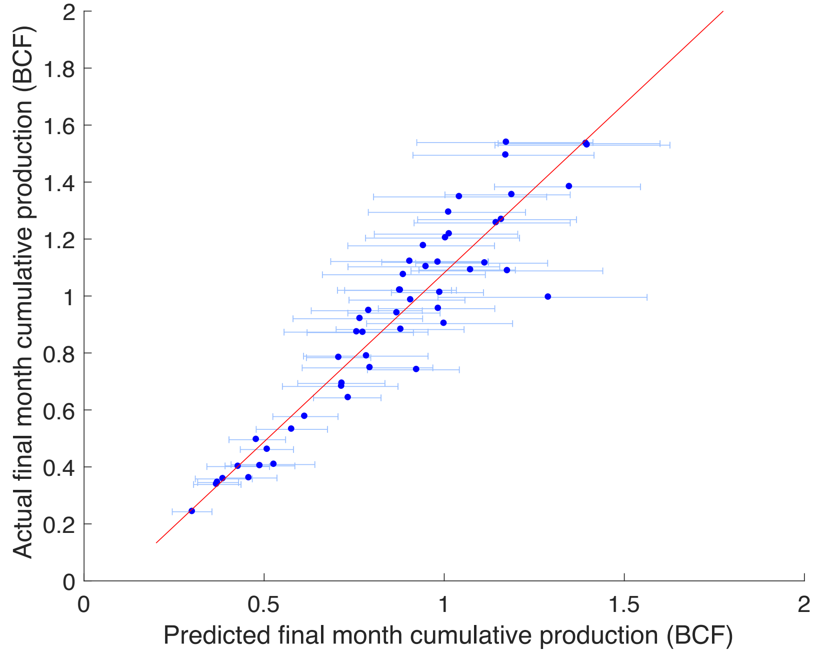
Bayesian approach improves accuracy of early life (12 mo.) production forecasts for test set

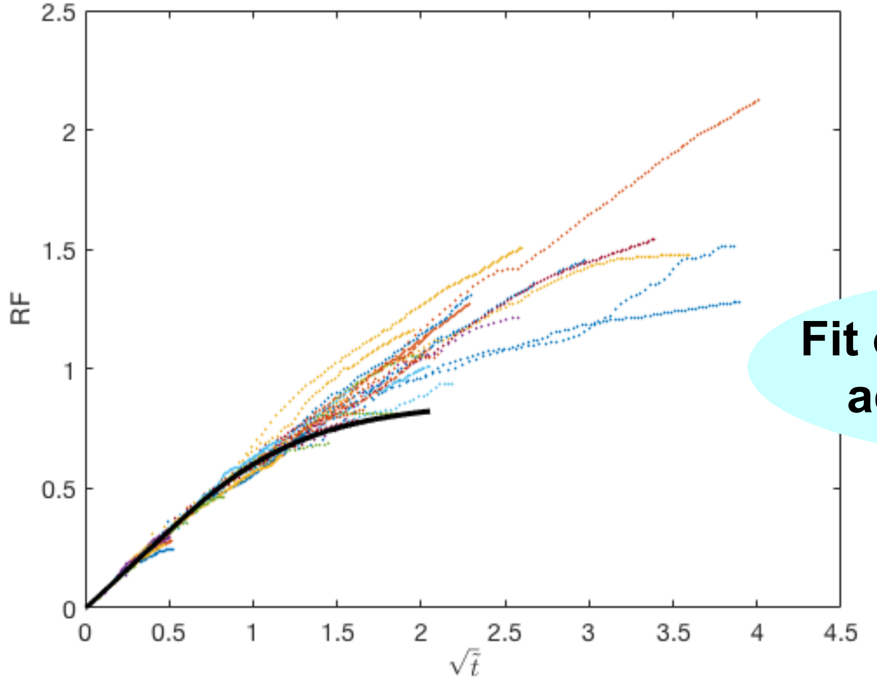
Nonlinear least squares (Patzek et al., 2013)



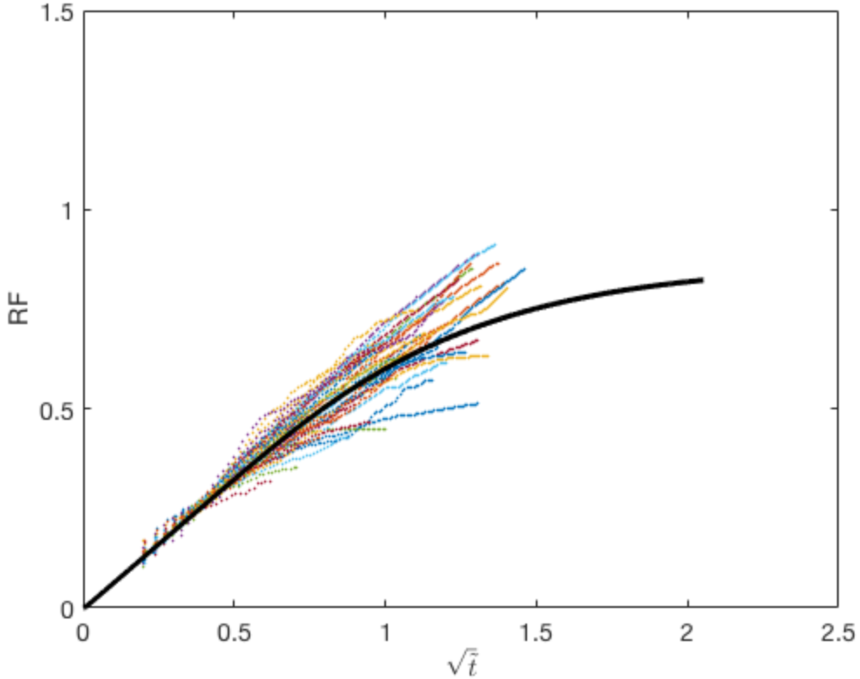
Accuracy of predicted cumulative production (at final time in data)

Hierarchical approach





Fit of scaling curve to actual production



Adaptive Metropolis algorithm (Haario et al., 2001)

Acceptance probability:

$$\alpha(X_{t-1}, Y) = \min\left(1, \frac{\pi(Y)}{\pi(X_{t-1})}\right)$$

Proposal is Gaussian centered at X_{t-1} with covariance:

$$C_t = \begin{cases} C_0, & t \leqslant t_0, \\ s_d \operatorname{cov}(X_0, \ldots, X_{t-1}) + s_d \varepsilon I_d, & t > t_0. \end{cases}$$

Empirical covariance calculated for $t = t_0 + 1$:

$$\operatorname{cov}(x_0, \ldots, x_k) = \frac{1}{k}\left(\sum_{i=0}^{k} x_i x_i^{\mathrm{T}} - (k+1)\bar{x}_k \bar{x}_k^{\mathrm{T}}\right)$$

Recursive formula used thereafter to reduce computation:

$$C_{t+1} = \frac{t-1}{t} C_t + \frac{s_d}{t}\left(t\bar{X}_{t-1}\bar{X}_{t-1}^{\mathrm{T}} - (t+1)\bar{X}_t \bar{X}_t^{\mathrm{T}} + X_t X_t^{\mathrm{T}} + \varepsilon I_d\right).$$

Source: Haario, H. Saksman, E., & Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli*, 7(2), 223-242.

MCMC for well