

# Applications of Data Assimilation Methods on a Coupled Dual Porosity Stokes Model\*

Xiukun Hu<sup>1</sup> and Craig C. Douglas<sup>2</sup>

<sup>1</sup> University of Wyoming Department of Mathematics and Statistics, Laramie, WY 82071-3036, USA

<sup>2</sup> University of Wyoming School of Energy Resources and Department of Mathematics and Statistics, Laramie, WY 82071-3036, USA  
craig.c.douglas@gmail.com

**Abstract.** Porous media and conduit coupled systems are heavily used in a variety of areas such as groundwater system, petroleum extraction, and biochemical transport. A coupled dual porosity Stokes model has been proposed to simulate the fluid flow in a dual-porosity media and conduits coupled system. Data assimilation is the discipline that studies the combination of mathematical models and observations. It can improve the accuracy of mathematical models by incorporating data, but also brings challenges by increasing complexity and computational cost. In this paper, we study the application of data assimilation methods to the coupled dual porosity Stokes model. We give a brief introduction to the coupled model and examine the performance of different data assimilation methods on a finite element implementation of the coupled dual porosity Stokes system. We also study how observations on different variables of the system affect the data assimilation process.

**Keywords:** Data assimilation · Dual porosity · Stokes equation · Multiphysics.

## 1 Introduction

Hou et al. [6] has proposed the Coupling of dual porosity flow with free flow as a replacement of the widely used Stokes-Darcy family. The proposed model has a better representation than the traditional Stokes Darcy model in modeling fractured porous media with large conduits. Potential applications of this model include petroleum extraction, hydrology, geothermal systems, and carbon sequestration. A finite element implementation of this model using FEniCS has been developed and studied by the authors [8]. Data assimilation is the discipline that studies the combination of mathematical models and observations. In this paper, we will apply data assimilation methods to the implementation of the coupled model to improve the accuracy of the model predictions [4,9].

In Section 2, we give an introduction to the mathematical model of the coupled dual porosity Stokes model proposed by Hou et al. [6]. In Sections 3 and 4 we illustrate the applications of data assimilation methods on the coupled dual porosity Stokes model.

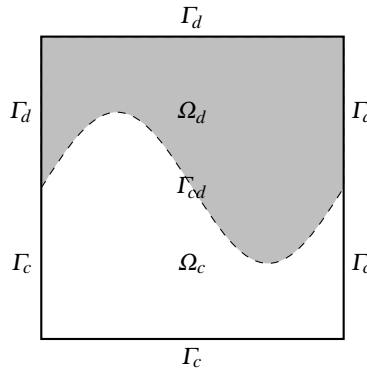
---

\* Supported in part by National Science Foundation grant DMS-1722692.

We set up a data assimilation context from our model in Section 3. We present the numerical results based on synthetic data in Section 4. In Section 5 we draw conclusions and discuss future works.

## 2 A Coupled Dual Porosity Stokes Model

The dual porosity Stokes model proposed by Hou et al. [6] consists of a dual porosity porous subdomain and a conduit subdomain. An example is shown in Fig. 1 where  $\Omega_d$  represents the porous subdomain and  $\Omega_c$  represents the conduit subdomain. Each subdomain has its own set of boundary conditions, represented by  $\Gamma_d$  and  $\Gamma_c$  respectively in the figure. The interface  $\Gamma_{cd}$  is the only place where the two subdomains communicate with each other.



**Fig. 1.** A simplified coupled model in 2D.

Barenblatt et al. [2] first proposed the dual porosity model in 1960. Later in 1963, Warren and Root [16] studied the model thoroughly. In a dual porosity medium, two subsystems are assumed. One is the matrix subsystem, which has high porosity and low permeability, and the other is the microfracture subsystem, which has low porosity and high permeability. The dual porosity equations governing the dual porosity subdomain  $\Omega_d$  in our coupled dual porosity Stokes model are

$$\phi_m C_{mt} \frac{\partial p_m}{\partial t} - \nabla \cdot \frac{k_m}{\mu} \nabla p_m = -Q, \quad (1)$$

$$\phi_f C_{ft} \frac{\partial p_f}{\partial t} - \nabla \cdot \frac{k_f}{\mu} \nabla p_f = Q + q_p. \quad (2)$$

The constant  $\mu$  represents the dynamic viscosity. The constants  $k_m$  and  $k_f$  represent the intrinsic permeability,  $\phi_m$  and  $\phi_f$  the porosities,  $C_{mt}$  and  $C_{ft}$  the total compressibility, of the matrix and the microfracture subsystems respectively. The variables  $p_m$  and  $p_f$  are the flow pressure of the matrix and the microfracture subsystems respectively. The

coefficient function  $q_p$  is the sink/source term. The term  $Q$  denotes the mass transfer rate per unit volume from the matrix subsystem to the microfracture subsystem and is defined as

$$Q = \frac{\sigma k_m}{\mu} (p_m - p_f), \quad (3)$$

where the parameter  $\sigma$  represents the characteristic of the fractured rock and is commonly known as the shape factor. Formulas for calculating  $\sigma$  can be found in Warren and Root [16] and Mora and Wattenbarger [12].

We assume the flow in the conduit domain is Stokes flow and thus describe it using the Stokes equation in (4) and (5). Note that the model can be extended to other free flow models such as the incompressible Navier-Stokes model, as proposed in [4].

$$\frac{\partial u}{\partial t} - \nabla \cdot \mathbf{T}(u, p) = f, \quad (4)$$

$$\nabla \cdot u = 0. \quad (5)$$

The two variables, the flow velocity vector  $u$  and the flow pressure  $p$ , together describe the state of the flow. The constant  $\nu$  represents the kinematic viscosity. The vector valued function  $f$  is a general body force term. The operator  $\mathbf{T}(u, p) := 2\nu\mathbf{D}(u) - p\mathbf{I}$  is the stress tensor and  $\mathbf{D}(u) := \frac{1}{2}(\nabla u + \nabla^T u)$  is the deformation tensor, where  $\mathbf{I}$  is the identity matrix.

Four interface conditions are imposed:

$$-\frac{k_m}{\mu} \nabla p_m \cdot (-n_{cd}) = 0, \quad (6)$$

$$u \cdot n_{cd} = -\frac{k_f}{\mu} \nabla p_f \cdot n_{cd}, \quad (7)$$

$$-n_{cd}^T \mathbf{T}(u, p) n_{cd} = \frac{p_f}{\rho}, \quad (8)$$

$$-\mathbf{P}_\tau(\mathbf{T}(u, p) n_{cd}) = \frac{\alpha \nu \sqrt{N}}{\sqrt{\text{trace}(\mathbf{\Pi})}} \left( u + \frac{k_f}{\mu} \nabla p_f \right), \quad (9)$$

where  $n_{cd}$  is the unit normal vector of the interface  $\Gamma_{cd}$ , pointing toward  $\Omega_d$ . The function  $\mathbf{P}_\tau$  is the projection operator onto the local tangent plane of  $\Gamma_{cd}$ . The constant  $\alpha$  is dimensionless and depends on the properties of the fluid and the permeable material. The constant  $\rho$  is the fluid density. The constant  $N$  is the space dimension.  $\mathbf{\Pi} := k_f \mathbf{I}$  is the intrinsic permeability of the microfracture subsystem.

Equation (6) represents the no mass exchange condition between the matrix subsystem in  $\Omega_d$  and the conduit. This is an assumption based on of the huge difference in permeabilities between the matrix and the microfracture subsystems. Equation (7) imposes conservation of mass exchange between the conduit and the microfracture subsystem on the interface. Equation (8) balances the two forces on the interface: the kinetic pressure in the microfracture subsystem and the normal component of the normal stress in the free flow. Equation (9) is the empirical Beavers-Joseph interface condition [3], which claims that the tangential component of the normal stress incurred by the free

flow along the interface is proportional to the difference of the tangential component of flow velocities at two sides of the interface.

By introducing test function  $[\psi_m, \psi_f, v^T, q]^T$ , the coupled dual porosity Stokes PDE system defined by (1)–(9) has the variational form,

$$\begin{aligned}
& \int_{\Omega_d} \left( \phi_m C_{mi} \frac{\partial p_m}{\partial t} \psi_m + \frac{k_m}{\mu} \nabla p_m \cdot \nabla \psi_m + \frac{\sigma k_m}{\mu} (p_m - p_f) \psi_m \right) d\Omega \\
& + \int_{\Omega_d} \left( \phi_f C_{fi} \frac{\partial p_f}{\partial t} \psi_f + \frac{k_f}{\mu} \nabla p_f \cdot \nabla \psi_f + \frac{\sigma k_m}{\mu} (p_f - p_m) \psi_f \right) d\Omega \\
& + \eta \int_{\Omega_c} \left( \frac{\partial u}{\partial t} \cdot v + 2v \mathbf{D}(u) : \mathbf{D}(v) - p \nabla \cdot v \right) d\Omega \\
& + \eta \int_{\Gamma_{cd}} \left( \frac{1}{\rho} p_f v \cdot n_{cd} + \frac{\alpha v \sqrt{N}}{\sqrt{\text{trace}(\boldsymbol{\Pi})}} \mathbf{P}_\tau \left( u + \frac{k_f}{\mu} \nabla p_f \right) \cdot v \right) d\Gamma \\
& + \eta \int_{\Omega_c} \nabla \cdot u q d\Omega - \int_{\Gamma_{cd}} u \cdot n_{cd} \psi_f d\Gamma \\
& = \eta \int_{\Omega_c} f \cdot v d\Omega + \int_{\Omega_d} q_p \psi_f.
\end{aligned} \tag{10}$$

A finite element implementation using the automated partial differential equation (PDE) solving platform FEniCS [1,11] has been developed by the authors [8]. The backward Euler time stepping scheme was used for time discretization.

### 3 A Data Assimilation Problem Based on the Coupled Model

In order to apply data assimilation methods to the coupled dual porosity Stokes model, we first convert the dual porosity Stokes model into a discrete dynamical system, and define the observations on it.

Following the finite element analysis with backward Euler scheme, at timestep  $t$  we solve the following equation system for the four variables in four finite functional spaces,

$$\mathbf{A} \begin{bmatrix} p_m^{(t)} \\ p_f^{(t)} \\ \mathbf{u}^{(t)} \\ p^{(t)} \end{bmatrix} = \mathbf{C} \begin{bmatrix} p_m^{(t-\Delta t)} \\ p_f^{(t-\Delta t)} \\ \mathbf{u}^{(t-\Delta t)} \\ p^{(t-\Delta t)} \end{bmatrix} + \mathbf{b}.$$

The matrix  $\mathbf{A}$  is assembled from the bilinear form

$$\begin{aligned}
& a \left( \begin{bmatrix} p_m^{(t)} \\ p_f^{(t)} \\ u^{(t)} \\ p^{(t)} \end{bmatrix}, \begin{bmatrix} \psi_m^{(t)} \\ \psi_f^{(t)} \\ v^{(t)} \\ q^{(t)} \end{bmatrix} \right) \\
&= \int_{\Omega_d} \left( \phi_m C_{mt} \frac{p_m^{(t)}}{\Delta t} \psi_m + \frac{k_m}{\mu} \nabla p_m \cdot \nabla \psi_m + \frac{\sigma k_m}{\mu} (p_m - p_f) \psi_m \right) d\Omega \\
&+ \int_{\Omega_d} \left( \phi_f C_{ft} \frac{p_f^{(t)}}{\Delta t} \psi_f + \frac{k_f}{\mu} \nabla p_f \cdot \nabla \psi_f + \frac{\sigma k_m}{\mu} (p_f - p_m) \psi_f \right) d\Omega \\
&+ \eta \int_{\Omega_c} \left( \frac{u^{(t)}}{\Delta t} \cdot v + 2v \mathbf{D}(u) : \mathbf{D}(v) - p \nabla \cdot v \right) d\Omega \\
&+ \eta \int_{\Gamma_{cd}} \left( \frac{1}{\rho} p_f v \cdot n_{cd} + \frac{\alpha v \sqrt{N}}{\sqrt{\text{trace}(\mathbf{\Pi})}} \mathbf{P}_\tau \left( u + \frac{k_f}{\mu} \nabla p_f \right) \cdot v \right) d\Gamma \\
&+ \eta \int_{\Omega_c} \nabla \cdot u q d\Omega - \int_{\Gamma_{cd}} u \cdot n_{cd} \psi_f d\Gamma.
\end{aligned}$$

The vector  $\mathbf{b}$  is assembled from the linear form

$$L \left( \begin{bmatrix} p_m^{(t-\Delta t)} \\ p_f^{(t-\Delta t)} \\ u^{(t-\Delta t)} \\ p^{(t-\Delta t)} \end{bmatrix} \right) = \eta \int_{\Omega_c} f \cdot v d\Omega + \int_{\Omega_d} q_p \psi_f.$$

The matrix  $\mathbf{C}$  is assembled from the bilinear form

$$\begin{aligned}
c \left( \begin{bmatrix} p_m^{(t-\Delta t)} \\ p_f^{(t-\Delta t)} \\ u^{(t-\Delta t)} \\ p^{(t-\Delta t)} \end{bmatrix}, \begin{bmatrix} \psi_m^{(t)} \\ \psi_f^{(t)} \\ v^{(t)} \\ q^{(t)} \end{bmatrix} \right) &= \int_{\Omega_d} \left( \phi_m C_{mt} \frac{p_m^{(t-\Delta t)}}{\Delta t} \psi_m + \phi_f C_{ft} \frac{p_f^{(t-\Delta t)}}{\Delta t} \psi_f \right) d\Omega \\
&+ \eta \int_{\Omega_c} \frac{u^{(t-\Delta t)}}{\Delta t} \cdot v d\Omega,
\end{aligned}$$

and thus has the form

$$\mathbf{C} = \begin{bmatrix} \frac{\phi_m C_{mt}}{\Delta t} \mathbf{I}_{d_m} & & & \\ & \frac{\phi_f C_{ft}}{\Delta t} \mathbf{I}_{d_f} & & \\ & & \mathbf{I}_{d_u} & \\ & & & \mathbf{0}_{d_p} \end{bmatrix},$$

where  $d_m, d_f, d_u$ , and  $d_p$  are the degrees of freedoms of  $p_m, p_f, u$ , and  $p$ , respectively,

If we let the state variable

$$v_t = \begin{bmatrix} p_m^{(t)} \\ p_f^{(t)} \\ u^{(t)} \\ p^{(t)} \end{bmatrix},$$

the dynamical system can be expressed as

$$v_{t+\Delta t} = \Psi(v_t) + \xi_t, \quad (11a)$$

$$\Psi(v_t) = \mathbf{A}^{-1} \mathbf{C} v_t + \mathbf{A}^{-1} \mathbf{b}, \quad (11b)$$

where  $\xi_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$  represents the model error. This dynamical system is linear. Note that the coefficient matrix  $\mathbf{C}$  is singular as is  $\mathbf{A}^{-1} \mathbf{C}$ , the Jacobian of  $\Psi$  defined in (11). Since some smoothing algorithms involve inverting the Jacobian of the dynamical system, we need to avoid singularities.

In general we can use the singular value decomposition to get around with singularities. In our case, we let our state variable

$$v_t^* = \begin{bmatrix} p_m^{(t)} \\ p_f^{(t)} \\ u^{(t)} \end{bmatrix}.$$

The dynamical system becomes

$$v_{t+\Delta t}^* = \Psi^*(v_t^*) + \xi_t, \quad (12a)$$

$$\Psi^*(v_t^*) = (\mathbf{A}^{-1})^* \mathbf{C}^* v_t^* + (\mathbf{A}^{-1})^* \mathbf{b}^*, \quad (12b)$$

where  $\mathbf{M}^*$  represents the matrix generated by removing the last  $d_p$  rows and columns from a matrix  $\mathbf{M}$ , and  $\mathbf{b}^*$  is the vector from removing the last  $d_p$  components of a vector  $\mathbf{b}$ . In fact (12) can also be formed from applying singular value decomposition to  $\mathbf{A}^{-1} \mathbf{C}$  in (11). Note that  $p^{(t)}$  can still be calculated from  $p_m^{(t-\Delta t)}$ ,  $p_f^{(t-\Delta t)}$  and  $u^{(t-\Delta t)}$ , which in turn can be calculated from  $p_m^{(t)}$ ,  $p_f^{(t)}$  and  $u^{(t)}$ .

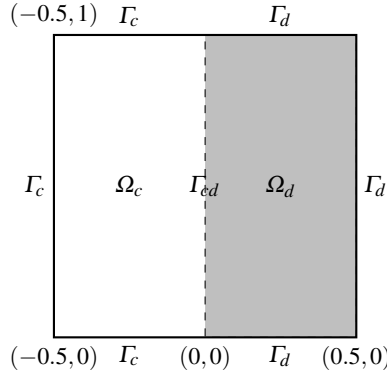
Similarly, the Dirichlet boundary conditions will also cause singularities as they do not depend on previous boundary values. We remove all Dirichlet boundary values from the state variable  $v_t$  using the same technique.

We base the dynamical model on a two dimensional dual porosity Stokes model shown in Figure 2. Let  $\Omega = [-0.5, 0.5] \times [0, 1]$  be a shifted unit square,  $\Omega_c = \{(x, y) \in \Omega \mid x \leq 0\}$ , and  $\Omega_d = \{(x, y) \in \Omega \mid x \geq 0\}$ . The interface is  $\Gamma_{cd} = \{(x, y) \in \Omega \mid x = 0\}$ . The domain is partitioned uniformly into  $\frac{1}{16} \times \frac{1}{16}$  squares.

Dirichlet boundary conditions on  $\Gamma_c$  and  $\Gamma_d$ , initial conditions for all variables, and coefficients  $q_p$  and  $f$  are constructed such that

$$\begin{aligned} p_m &= \cos(\pi t) \cos(x(-y+1)) \\ p_f &= ((x^2 + y^2 - 2y + 2) \cos(\pi t) - 10\pi \sin(\pi t)) \cos(xy - x) \\ u &= \begin{bmatrix} 2x \cos(\pi t) \\ 2x \cos(\pi t) - 2y \cos(\pi t) \end{bmatrix} \\ p &= -10\pi \sin(\pi t) + (x^2 + 2x + y^2 - 2y + 6) \cos(\pi t) \end{aligned}$$

is the solution to our problem.



**Fig. 2.** The 2D example model with a shifted unit square domain  $\Omega = [-0.5, 0.5] \times [0, 1]$ , conduit subdomain  $\Omega_c = \{(x, y) \in \Omega \mid x \leq 0\}$  and dual porosity subdomain  $\Omega_d = \{(x, y) \in \Omega \mid x \geq 0\}$ .

Also, we let  $\Delta t = 0.01$ ,  $\xi_t \sim \mathcal{N}(\mathbf{0}, 5\mathbf{I})$ ,  $v_0^* \sim \mathcal{N}(\mathbf{0}, 100\mathbf{I})$ . The large variance of  $v_0^*$  indicates that we have little knowledge about the initial condition.

For the observations, we assume we have direct observations to every 4 components of  $v_t^*$  at time  $t$ :

$$y_t = h(v_t^*) + \eta_t, \tag{13a}$$

$$h(v_t^*) = \mathbf{H}v_t^*, \tag{13b}$$

$$\eta_t \sim \mathcal{N}(\mathbf{0}, 5\mathbf{I}), \tag{13c}$$

where

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots \\ \dots & & & & & & & & \end{bmatrix}.$$

We observe every 0.01 time unit starting at  $t = 0.01$ . Equations (12) and (13) together defines the data assimilation problem we are solving. Data are generated synthetically.

## 4 Numerical Results

We run the model against the three dimensional variational method (3DVAR), the strong constraint four dimensional variational method (s4DVAR) with a time window with length 0.04, the extended Rauch-Tung-Striebel smoother (ExtRTS) [15], the extended Kalman Filter (ExtKF) [10], the ensemble Kalman Filter (EnKF) [7,5] with 100 particles, and ensemble Rauch-Tung-Striebel smoother (EnRTS) [13] with 100 particles. Note that since we have a linear data assimilation problem, the extended methods ExtRTS and ExtKF are just the Rauch-Tung-Striebel smoother (RTS) and the Kalman

Filter (KF). We also use a baseline filtering method *Forward* that only uses the mathematical model  $\Psi$  and ignores all data. It starts at  $v_0^* = \mathbf{0}$  and then applies  $\Psi^*$  to get an approximation for  $v_t^*$ . Since the model is linear, we expect an optimal solution by ExtKF for filtering and ExtRTS for smoothing.

All numerical experiments were run with the data assimilation package DAPPER [14] on the Teton computer cluster at the Advanced Research Computing Cluster (ARCC) at the University of Wyoming.

The results of filtering on the model with such observations are in Table 1 and the results of smoothing are in Table 2. Since we have a linear system with Gaussian errors, the Kalman Filter and the Kalman Smoother are expected to have optimal data assimilation solutions for filtering and smoothing, respectively, which is in accordance with our numerical results. We see that the Kalman Filter, the Kalman Smoother and 3DVAR are efficient in our small linear model while ensemble methods and s4DVAR are relatively slow.

**Table 1.** Average root mean square error for filtering (rmse.f) and elapsed time

	Forward	3DVAR	ExtKF	EnKF
rmse.f	0.4717	0.2824	0.2604	0.2651
elapsed time	5s	1s	5s	56s

**Table 2.** Average root mean square error for smoothing (rmse.s) and elapsed time

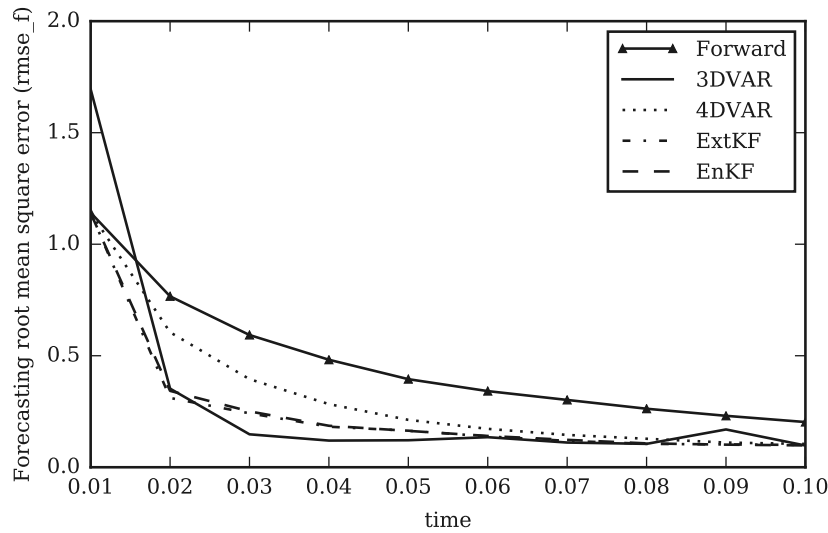
	s4DVAR	ExtRTS	EnRTS
rmse.s	0.3	0.1907	0.2033
elapsed time	121s	32s	62s

The error of different data assimilation methods over time are shown in Figures 3 and 4. Since Forward, 4DVAR, ExtKF, and EnKF all start with an initial guess  $\tilde{v}_0 = \mathbf{0}$ , they all have the same predictions at  $t = 0.01$ . This is why they all have the same error at  $t = 0.01$  for forecasting as shown in Figure 3. The predictions are made every 0.01 time units. ExtKF has a smaller forecasting error than all the other methods except for 3DVAR. Our 3DVAR implementation utilizes all true states to approximate the background covariance  $B_t$ . The exposure to the true states enables the 3DVAR implementation to surpass the theoretical optimal solution from the Kalman Filter. EnKF has a result very similar to that of ExtKF. In EnKF, the calculations of mean and variance of the states are approximated using the Monte Carlo method. Since the states follows a Gaussian process, the approximations converge to the truths as the number of particles increases. We can also see in Figure 4 that by utilizing all observations, the smoothing error at  $t = 0.01$  is reduced by half, comparing to the forecasting error in Figure 3. Note that the Kalman Smoother ExtRTS achieves the best result at all time, and the ensem-



ble Kalman Smoother EnRTS has a very similar result as ExtRTS, but consumes much more computation time as shown in Table 2.

Note that the baseline method *Forward* also has a decreasing error with respect to time. This is caused by the characteristics of our dynamical system. Because of the essential boundaries in our coupled model, solutions to the PDE system with different initial conditions all converge to each other as  $t \rightarrow \infty$ . This can also be explained by the linear dynamical system. Consider a linear dynamical system with  $\Psi(v_t) = Mv_t$  where  $\|M\| < 1$ . Then  $\Psi^{(n)}(v_t) \rightarrow \mathbf{0}$  as  $t \rightarrow 0$ .

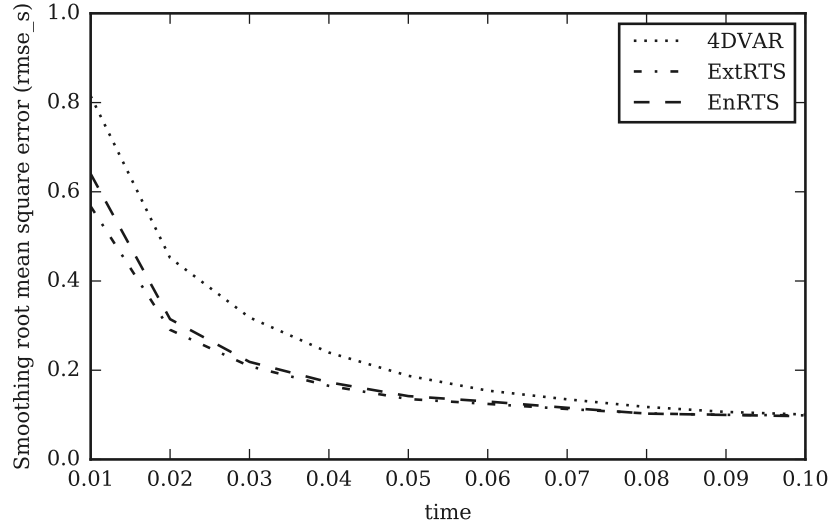


**Fig. 3.** Forecasting error of different filtering algorithms

The results of smoothing at  $t = 0.01$  from s4DVAR and ExtRTS are shown in Figure 5. The results of filtering at  $t = 0.02$  from 3DVAR and ExtKF are shown in Figure 6. We can see from Figure 5 that by using limited observations, 4DVAR and ExtRTS are able to recover the state close to true state. Also in Figure 6, we see that by using only data at  $t = 0.01$ , 3DVAR and the Kalman Filter are able to predict a state at  $t = 0.02$  that is much closer to true state comparing to the Forecast baseline method.

We also explore the importance of observations on different variables. With the same settings on the dynamical system, we apply the Kalman Filter (ExtKF) to observations on  $p_m$ ,  $p_f$ , and  $u$  separately. We still observe from  $t = 0.01$  and observe every 0.01 time unit, but on all grid points. The results are presented in Figure 7

We can see from Figure 7 that the data on the flow pressure  $p_m$  in the matrix subsystem in the dual porosity subdomain provides most of the information while the other two variables provide little improvement over the *Forward* baseline method, which uses no observation at all. This behavior exists in all our test models with different boundary



**Fig. 4.** Smoothing error of different smoothing algorithms

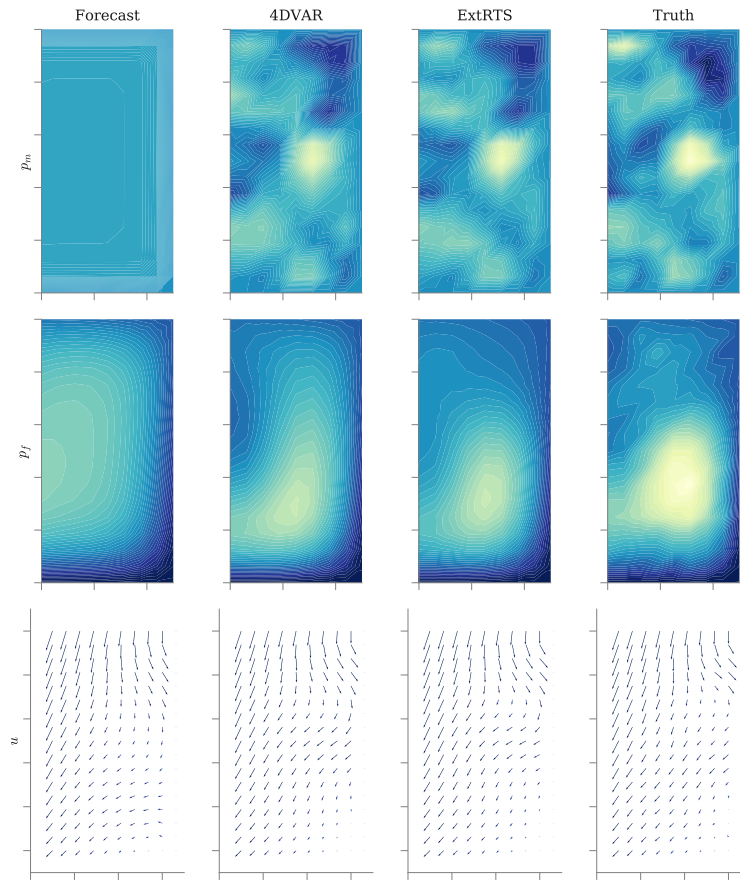
conditions, source terms and geometries. This phenomenon needs further investigation. Here we conclude that in our limited test cases, observations on  $p_m$  provide significant information about the true states while observations on  $p_f$  and  $u$  do not.

Lastly we show the result of the Kalman Filter and the Kalman Smoother on a 3D coupled dual porosity Stokes model introduced in [8], with the mesh size  $h = 1/8$  and observations on  $p_m$  only. All the other settings are the same as in the 2D model. The results in Figure 8 validate the Kalman Filter and the Kalman Smoother on our 3D models and real world applications.

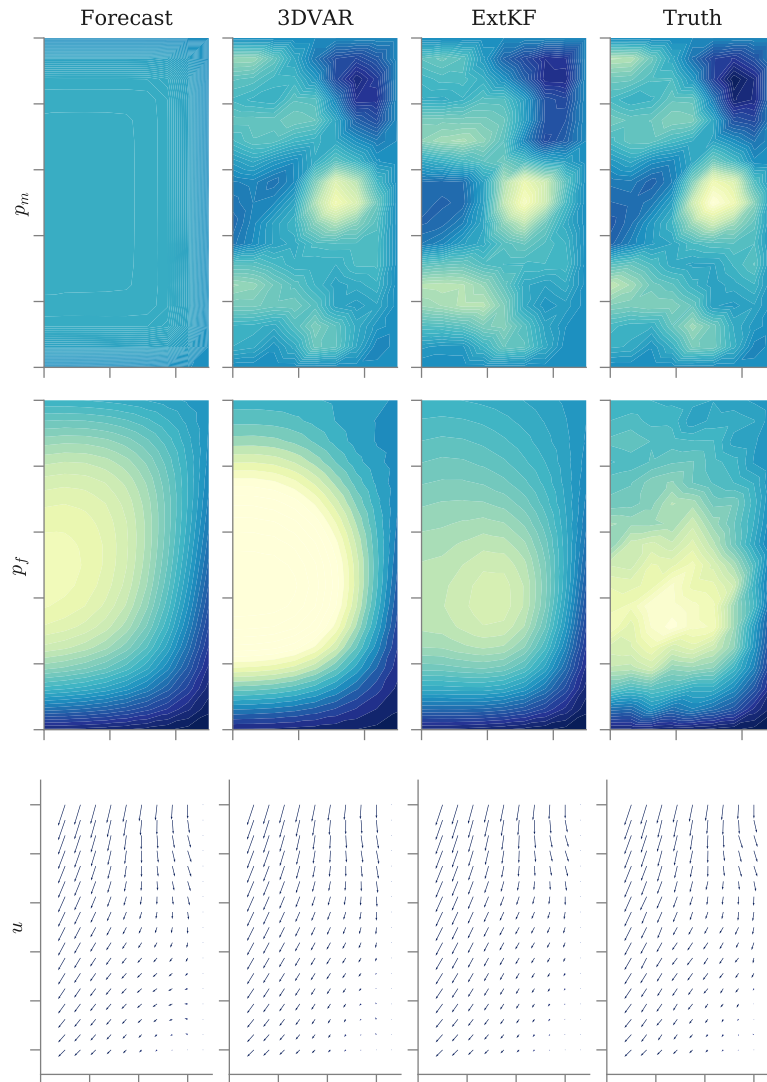
## 5 Conclusions and Future Work

In this paper, we introduced the coupled dual porosity Stokes model. We set up a data assimilation problem based on the coupled model and applied different data assimilations to solve the problem. Due to the linearity of the coupled dual porosity Stokes model, the Kalman Filter and the Kalman Smoother achieve optimal solutions for filtering and smoothing, respectively, as expected. From our numerical experiments we have seen that observations of pressures in the matrix subsystem contain most of the useful information for data assimilation.

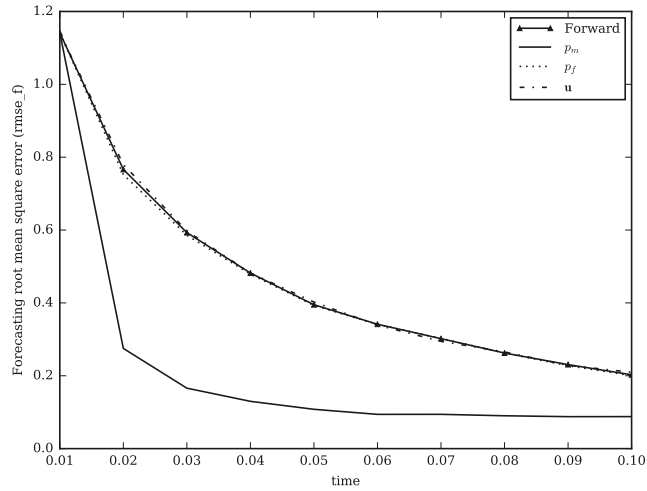
Future work includes exploring different data assimilation methods on the nonlinear coupled dual porosity Navier-Stokes model, applying data assimilation methods with experiment data and investigating the reason behind the uneven distribution of information in different variables.



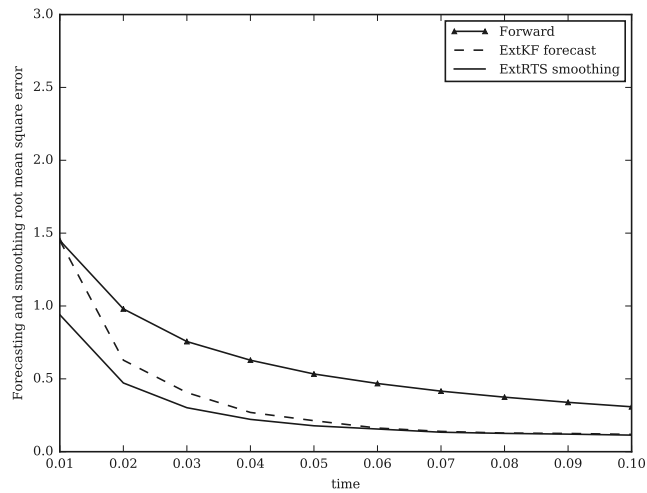
**Fig. 5.** Results of smoothers at time  $t = 0.01$  in the 2D model.



**Fig. 6.** Results of filters at time  $t = 0.02$  in the 2D model.



**Fig. 7.** Forecasting error of Kalman Filter with different observations.



**Fig. 8.** Forecasting and smoothing error of ExtKF and ExtRTS on the 3D model.

## References

1. Alnæs, M.S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., Wells, G.N.: The fenics project version 1.5. *Archive of Numerical Software* **3**, 9–23 (2015)
2. Barenblatt, G.I., Zheltov, I.P., Kochina, I.N.: Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks [strata]. *J. Appl. Math. Mech.* **24**, 852–864 (1960)
3. Beavers, G.S., Joseph, D.D.: Boundary conditions at a naturally permeable wall. *J. Fluid Mech.* **30**, 197–207 (1967)
4. Douglas, C.C., Hu, X., Bai, B., He, X., Wei, M., Hou, J.: A data assimilation enabled model for coupling dual porosity flow with free flow. In: 2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES). pp. 304–307 (2018)
5. Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research* **99**(C5), 10143–10162 (1994)
6. Hou, J., Qiu, M., He, X., Guo, C., Wei, M., Bai, B.: A dual-porosity-stokes model and finite element method for coupling dual-porosity flow and free flow. *SIAM J. Sci. Comput.* **38**, B710–B739 (2016)
7. Houtekamer, P.L., Mitchell, H.L.: Data assimilation using an ensemble kalman filter technique. *Monthly Weather Rev.* **126**, 796–811 (1998)
8. Hu, X., Douglas, C.C.: An implementation of a coupled dual-porosity-stokes model with fenics. In: Rodrigues, J.M.F., Cardoso, P.J.S., Monteiro, J., Lam, R., Krzhizhanovskaya, V.V., Lees, M.H., Dongarra, J.J., Sloot, P.M.A. (eds.) *Computational Science – ICCS 2019*. pp. 60–73. Springer, Cham (2019)
9. Hu, X., Douglas, C.C.: Performance and scalability analysis of a coupled dual porosity stokes model implemented with fenics. *Japan J. Indust. Appl. Math.* **36**, 1039–1054 (2019)
10. Kalman, R.E.: A new approach to linear filtering and prediction problems. *J. Basic Engineering* **82**, 35 (1960)
11. Logg, A., Mardal, K.A., Wells, G.N.: Automated Solution of Differential Equations by the Finite Element Method, *Lecture Notes in Computational Science and Engineering*, vol. 84. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
12. Mora, C.A., Wattenbarger, R.A.: Analysis and verification of dual porosity and cbm shape factors. *J. Canadian Petroleum Tech.* **48**, 17–21 (2009)
13. Raanes, P.N.: On the ensemble rauch-tung-striebel smoother and its equivalence to the ensemble kalman smoother. *Quarterly J. Royal Meteorological Soc.* **142**, 1259–1264 (2016)
14. Raanes, P.N., et al.: nanscenter/dapper: Version 0.8 (Dec 2018). <https://doi.org/10.5281/zenodo.2029296>
15. Rauch, H.E., Striebel, C.T., Tung, F.: Maximum likelihood estimates of linear dynamic systems. *AIAA J.* **3**, 1445–1450 (1965)
16. Warren, J.E., Root, P.J.: The behavior of naturally fractured reservoirs. *SPE J.* **3**, 245–255 (1963)