MOX-Report No. 44/2022

Forecasting Oil Production Rates in Primary Depletion using the Physics-based Residual Kriging functional approach

Peli, R.; Dovera, L.; Fighera, G.; Menafoglio, A.; Secchi, P.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

Forecasting Oil Production Rates in Primary Depletion using the Physics-based Residual Kriging functional approach

R. Peli*, MOX, Department of Mathematics, Politecnico di Milano; **L. Dovera**, Eni S.p.A.; **G. Fighera**, Eni S.p.A.; **A. Menafoglio** MOX, Department of Mathematics, Politecnico di Milano and **P. Secchi,** MOX, Department of Mathematics, Politecnico di Milano;

*email: riccardo.peli@polimi.it

Keywords: data-driven method; physics-based modeling; geostatistics; functional data analysis

Abstract

In this work, we illustrate a novel functional data analysis approach for the forecast of oil production rates in a mature single-phase reservoir. This model is based on the recently developed Physics-based Residual Kriging predictor, which represents oil rates as functional data and decomposes them as the sum of the predictions of a physical model and the geostatistical modelization of its residuals. In this context, we use the recently introduced FlowNet model to build up the physical term which, through a network-based representation of the reservoir, avoids the burden of three-dimensional full-physics simulations. Furthermore, we propose an extension of the Physics-based Residual Kriging predictor in presence of ensemble of physical models, i.e. when the uncertainty in the model parameters is accounted for by simulating several models corresponding to different parameters samples. The Physics-based Residual Kriging predictor is here applied to the oil rates produced in a realistic reservoir. We analyze three different scenarios in terms of wells drilling schedule, from a simple to a realistic scheme. In each scenario, we compare the predictions given by Physics-based Residual Kriging to the ones obtained with FlowNet and a pure geostatistical approach.

Introduction

Production forecast is a crucial tool for the managing and future development of hydrocarbon reservoirs during all the phases of their life span, from exploration to maturity. Indeed, such a tool is usually required in the decision-making process to optimize production and for the assessment of investment scenarios, in order to maximize asset returns and mitigate the risk (Ahmed and Meehan, 2012; Pathak, 2021). At the stage when some historical production data are already available, the classical approach to production forecast in oil and gas industry is the development of a three-dimensional full-physics reservoir model (Aziz and Settari, 1979; Dake, 1978; Peaceman, 1977), which is simulated to produce predictions of quantities of interest, such as oil production rates at field or well levels. However, the construction of such a model may be very challenging, due to the complexity of fluid flow dynamics and subsurface reservoir structure, implying long development and tuning times and high computational cost. A key step in this process is the history matching (Oliver et al., 2008; Gilman and Chet, 2013), where the model parameters, as the reservoir petrophysical properties and fluids characteristics, are estimated to ensure the match between model predictions and data measured in the reservoir. In this process, iterative ensemble

smoothers can be used (Emerick and Reynolds, 2012, 2013) and, in real world applications, a lot of effort is required to obtain a model which is consistent with the observed data (Alfonzo and Oliver, 2019; Oliver, 2020).

An alternative to three-dimensional full-physics reservoir simulations is represented by surrogate models, also called proxy models. These models are usually simpler, in terms of number of parameters, and have faster simulation times. We distinguish two different approaches for the construction and the use of surrogate models: (i) proxy models which, starting from a full-physics reservoir model, allow one to produce the forecast more efficiently, and (ii) surrogate models which produce predictions without the need of a high-fidelity model. The latter, once conditioned to the past production data, could provide reasonable predictions of the future reservoir behavior.

A notable example of the first type is represented by models that entail the solution of the reservoir governing equations through the use of reduced order models (Ghasemi et al., 2014; Tang et al., 2021; Yoon et al., 2016), which are based on the formulation of the equations in the linear space generated by a set of high-fidelity solutions. Besides their attractiveness, this approach requires the implementation of the full-physics reservoir simulator to generate samples of high-fidelity solutions. Furthermore, their applicability to highly non-linear equations, such as the reservoir equations, is challenging and requires ad hoc formulations.

The second approach for surrogate models can be based on simplified formulations of the reservoir physics. A notable example is the Capacitance Resistance Model (CRM, Wanderley et al. (2018)), that expresses a simplified version of the mass conservation equation, modelling the interaction between the wells. The Interwell Numerical SImulation Model (INSIM, Zhao et al. (2015)) is a more complex version, where the pressure equation is solved on the nodes, i.e. the wells, of a network and the saturation equation on the connections, i.e. the streamtubes. The physics formulation, in this case, neglects capillary and compressibility effects in the solution of the saturation equation. Starting from the base formulation proposed in Zhao et al. (2015), the authors present some extensions in Guo et al. (2017); Guo and Reynolds (2019), which generalize the method in three dimensions and with a better numerical scheme for the equations solution. Within the same line of research, the FlowNet model (Kiærr et al., 2020) has been recently proposed. In this framework, the reservoir is represented in terms of a network, composed of one dimensional connections between wells, and characterized by a set of parameters, such as permeability of the connections and the coefficients of the relative permeability curves. It is based on the solution of the same physics equations as a high-fidelity model, but discretized on the wells network, highly reducing the number of parameters and the grid size. All these models do not need a detailed geological model and a history matching phase is required to ensure the match with the historical data, in particular production data.

In recent years, also pure data-driven surrogate models, in the framework of statistics and machine learning, have been developed for the prediction of production rates. Examples are illustrated in Han and Kwon (2021); Jin et al. (2020), whereas Coutinho et al. (2021) illustrates an example of integration of the physics in the formulation of this type of proxy models. Menafoglio et al. (2016a) propose a data-driven predictor, based on Kriging, that models spatial dependence among production rates in an unconventional reservoir.

Another approach is the Physics-based Residual Kriging, originally introduced in Peli et al. (2022), that combines a simplified physical model with a statistical predictor for the forecast of oil production rates in mature conventional reservoirs. In the Phy-RK framework, oil production rates are represented as functional data, i.e. modelling their variation over continuum time according to the theory of Functional Data Analysis (FDA, Ramsay (2005)), and they are decomposed as the sum of a physical model prediction and its residuals. The first term is given by a suitable version of the mass conservation equation for fluid flow in porous media and the second term is modelled in a

geostatistical framework, exploiting its spatial dependence. In Peli et al. (2022), two applications are presented where the physical model is given by the single-phase mass conservation equation, neglecting fluid and rock compressibilities and fluid compressibility, respectively. In both the cases, the equation is discretized on a fine bi-dimensional grid, prescribing the petrophysical parameters in each cell. Furthermore, the Phy-RK predictor is designed to take into account the wells drilling schedule, which implies a sequential opening of the wells in the reservoir.

We here present the application of the Phy-RK predictor to a realistic reservoir. Due to its three-dimensional geometry and realistic petrophysical properties distribution, this reservoir represents a challenging application of the proposed model, much more complex than the synthetic reservoirs analyzed in Peli et al. (2022). Assuming that the past production rates of the operating wells have been measured and that the drilling schedule of future wells is given, our goal is the prediction of the future production rates of existing and newly drilled wells. The wells drilling schedule is divided into activation intervals, identified by the drilling of a group of wells, and predictions are produced one interval at a time, as soon as the observations in the previous interval becomes available. In order to avoid the three-dimensional high fidelity reservoir modelization, we here opt to use FlowNet as physical model. The limited number of petrophysical parameters and the solution of the high-fidelity physics make FlowNet a powerful and flexible surrogate model and its integration in the Phy-RK method allows one to obtain accurate predictions, without the need to develop a three dimensional full-physics reservoir model.

Furthermore, we here extend Phy-RK to deal with ensemble of physical models. This situation appears whenever an ensemble smoother is used in the history matching phase, as in the FlowNet model, ending up with an ensemble of physical models, each corresponding to different realizations of the model parameters. We propose two strategies: the first one, called *Average* Phy-RK, consists in taking the average ensemble prediction as physical term, whereas the second one, called *Multiple* Phy-RK, consists in applying Phy-RK to each physical model obtaining an ensemble of predictions, which allows one to quantify the associated uncertainty.

We consider the application of Phy-RK to three different scenarios, concerning the same reservoir but obtained by varying the wells drilling schedule. In the *reference* case, we consider a regular schedule, where wells are drilled at the starting points of eight time intervals of equal length, in groups with the same number of wells. We then consider two additional cases: in the *variable number* schedule, the wells are drilled in groups of different numbers, while in the *realistic* scenario we further remove the constraint of equally long time intervals, mimicking a realistic perforation campaign. In this last case, we propose an extension of Phy-RK to deal with time intervals of different lengths.

The manuscript is organized as follows. In Section 1, we outline the problem and we describe the reservoir. Section 2 recalls the mathematical formulation of the Phy-RK predictor and its sequential formulation. In Section 3, we describe the FlowNet model and its integration in the Phy-RK predictor. The analyses of the three scenarios are illustrated in Section 4 for the *reference* case and in Section 5 for the *variable number* and *realistic* cases, respectively. Finally, Section 6 outlines conclusions and considerations for future development.

1 Problem description

In this work, our goal is to predict the oil production rates of a mature hydrocarbon reservoir. We here describe the synthetic reservoir that generates the data analyzed in this paper. The three-dimensional reservoir model is simulated using the open source software Open Porous Media (OPM) (Flemisch et al., 2011) and is used only to generate realistic production data and not in the

Phy-RK predictive model. Indeed, the Phy-RK approach has been designed to avoid the burden of the developing of a 3D reservoir model, relying on surrogate physical models, as described in Section 3.

The synthetic reservoir used in this study is a 3D model based on a chalk reservoir. Black-oil modelization is employed and the reservoir is produced in primary depletion. Figure 1 depicts the reservoir geometry and the wells locations. We here describe the wells schedule of the *reference* case, that is analyzed in Subsection 4, whereas other schedules are considered in Section 5. In the *reference* case, over a production schedule of 20 years, 137 oil producers are opened. The wells are incrementally drilled in the reservoir during 8 intervals of equal length of 900 days and, at the beginning of each interval, almost the same number of new wells is drilled. The oil production rates and the wells schedule are shown in Figure 2.

In this framework, our goal is the prediction of the oil production rates of the wells operating in a specified interval, given the observations of the production data in the previous intervals. This analysis is carried out in all the intervals and different approaches are compared: a pure data-driven approach, without modelling the wells interaction and the reservoir perturbation caused by newly drilled wells, a pure physical surrogate model approach, FlowNet, and the Phy-RK predictor, that effectively combines the other two.

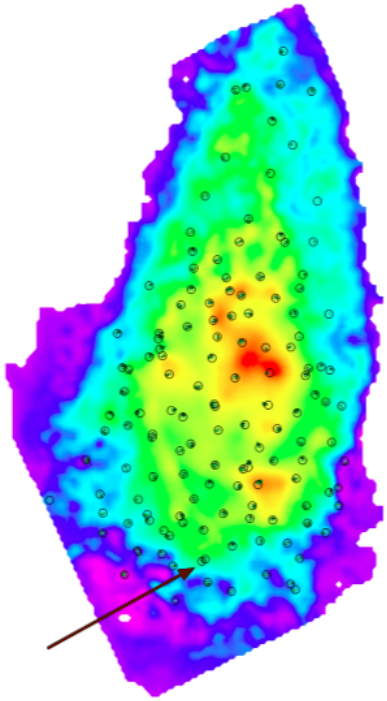


Figure 1: Map view from top of the depth surface of the reservoir model. Colors represent depth, from purple (high) to red (low) and circles represent wells locations. The arrow points a well that is the object of the analysis in Subsection 4.2

2 Physics-based Residual Kriging formulation

In this section we introduce the mathematical formulation of Phy-RK for functional data, referring to Peli et al. (2022) for further details. Consider the functional random variable $\mathcal{Y}_{\boldsymbol{s}}$, which represents the oil production rate of the well drilled at location $\boldsymbol{s}$, which belongs to the reservoir domain D,

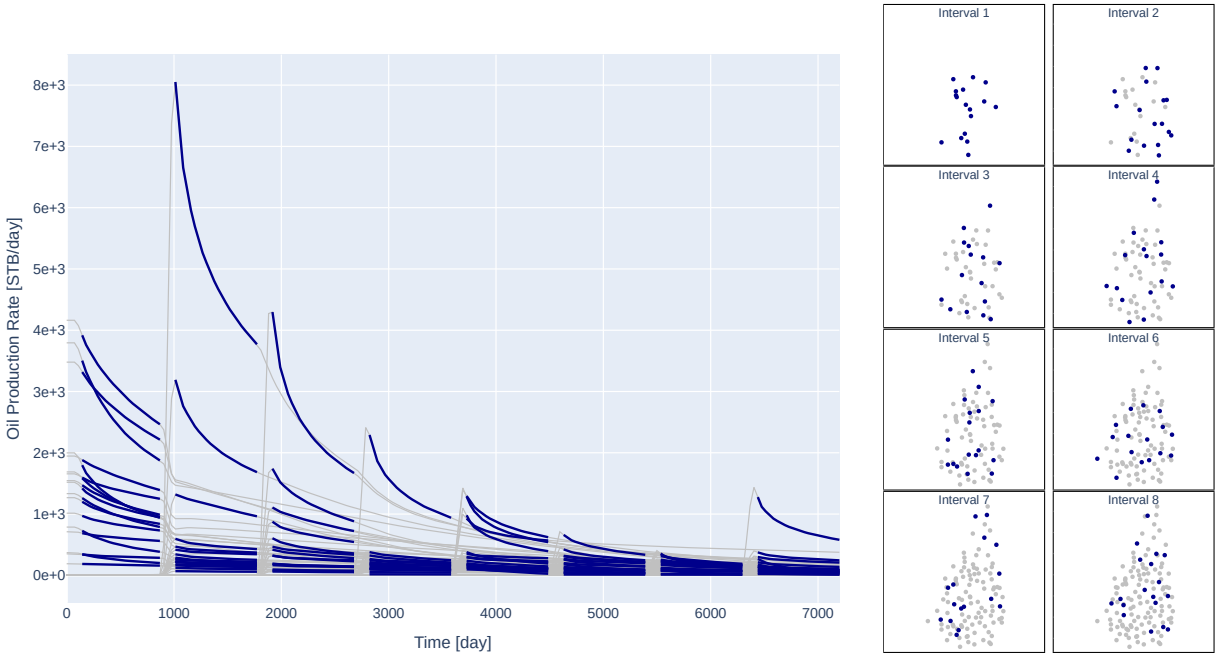


Figure 2: Production rates and schedule for the *reference* case described in Subsection 1. Left panel: blue lines represent the production rates corresponding to newly drilled wells in each interval, gray lines correspond to wells already operating. In the first interval, blue and gray lines coincide, as they correspond to the initial set of drilled wells. Right panel: blue circles represent the newly drilled producers and gray circles the already operating wells

over the continuum time t. We assume that the functional random variable takes values in a generic Hilbert space $\mathcal{H}$. The Phy-RK predictor aims to predict the oil production rate $\mathcal{Y}_{\boldsymbol{s}_0}$ at target site $\boldsymbol{s}_0$, having observed the production rates of the active wells located at $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$. The Phy-RK approach relies on the decomposition of the (random) production curves into the contribution $\mathcal{F}_{\boldsymbol{s}}(\boldsymbol{\theta})$, which is the prediction of a deterministic (surrogate) model with parameters $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^m$, and a stochastic residual $\mathcal{X}_{\boldsymbol{s}}$, as

$$\mathcal{Y}_{\boldsymbol{s}} = \mathcal{F}_{\boldsymbol{s}}(\boldsymbol{\theta}) + \mathcal{X}_{\boldsymbol{s}}.$$

Thus, it combines a deterministic model, which, in our applications, represents the reservoir physics, with a geostatistical model for the residual random variable. The Phy-RK prediction for the production curve $\mathcal{Y}_{\boldsymbol{s}_0}$, omitting from now on the dependence on $\boldsymbol{\theta}$, is then given by

$$\hat{\mathcal{Y}}_{\boldsymbol{s}_0} = \mathcal{F}_{\boldsymbol{s}_0} + \hat{\mathcal{X}}_{\boldsymbol{s}_0}, \tag{1}$$

where the residual prediction $\hat{\mathcal{X}}_{\boldsymbol{s}_0}$ is obtained through Universal Kriging (UK) predictor for functional data (Caballero et al., 2013; Menafoglio et al., 2013, 2016b), as the optimal linear combination of the observed residuals at locations $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$. Here, optimal refers to the fact the UK is the best linear unbiased predictor in the mean square sense. In this framework, the mean of the residual field (a.k.a. drift) is expressed in terms of a linear model

$$m_{\boldsymbol{s}} = \sum_{i=0}^{L} \beta_l f_l(\boldsymbol{s}), \tag{2}$$

where f_0 is the intercept (i.e, $f_0(\boldsymbol{s}) = 1$ for all $\boldsymbol{s} \in D$), $f_1(\boldsymbol{s}), \ldots, f_L(\boldsymbol{s}) \in \mathbb{R}$ are known spatially dependent covariates, such as the petrophysical properties, the spatial coordinates or, as in our

applications, the pressure at the well location computed by the surrogate physical model. The terms $\beta_0(\cdot), \ldots, \beta_L(\cdot) \in \mathcal{H}$ are functional parameters, which do not depend on the spatial location. Whenever only the intercept f_0 is employed, the Universal Kriging predictor coincides with the Ordinary Kriging (OK) predictor.

Phy-RK can also be used for sequential problems, such as the prediction of oil production rates. Indeed, as depicted in Figure 2, production rates observations become available along consecutive time intervals, in a context where the physical and stochastic processes generating them evolve as time intervals succeed one another. The perturbation is due to the different wells configurations among different time intervals. Indeed, we assume that the wells are opened or closed in N_a time instants, $(t_0, \ldots, t_{N_a-1})$, and, within any time interval $a_i = [t_{i-1}, t_i)$, for $i = 1, \ldots, N_a$, the wells configuration does not change. Hence, the whole time domain is subdivided into N_a time intervals (in Figure 2, $N_a = 8$). We assume that, in each interval a_i, the production rates of the active wells has been sampled at a set of locations $W_i = \left\{ \boldsymbol{s}_j^i \right\}_{j=1}^{n_i}$, which is the set of active wells in interval a_i. In this section, we assume that all the intervals have equal length, but we will relax this hypothesis in Subsection 5.2. There, the predictor is extended to deal with intervals of irregular length.

The general Phy-RK predictor (1) is thus sequentially reformulated to account for the time interval a_i, as

$$\hat{\mathcal{Y}}_{\boldsymbol{s}_0,a_i} = \mathcal{F}_{\boldsymbol{s}_0,a_i} + \hat{\mathcal{X}}_{\boldsymbol{s}_0,a_i}, \tag{3}$$

i.e. the sum of the physics-based prediction $\mathcal{F}_{\boldsymbol{s}_0,a_i}$ in the interval a_i, and the predicted residual, which is obtained with functional UK, having observed the residuals in the interval a_{i-1} at locations W_{i-1}. In this setting, if residual fields among consecutive intervals exhibit a strong variation, Peli et al. (2022) propose to include higher orders residuals terms, updated in each interval, to further correct the predictive model based on the misfit between observations and predictions in the previous intervals. As depicted in Figure 3, at the third interval one has already observed in the second interval the *first-order* residuals, i.e. the difference between observations and physical model predictions $\mathcal{X}^{(1)}_{\boldsymbol{s}_j,a_2} = \mathcal{Y}_{\boldsymbol{s}_j,a_2} - \mathcal{F}_{\boldsymbol{s}_j,a_2}$, and the *second-order* residuals, i.e. the difference between observations and the *first-order* predictions computed in that interval $\mathcal{X}^{(2)}_{\boldsymbol{s}_j,a_2} = \mathcal{Y}_{\boldsymbol{s}_j,a_2} - (\mathcal{F}_{\boldsymbol{s}_j,a_2} + \hat{\mathcal{X}}^{(1)}_{\boldsymbol{s}_j,a_2})$. Hence, in the third interval one can build the predictor

$$\hat{\mathcal{Y}}_{\boldsymbol{s}_0,a_3} = \mathcal{F}_{\boldsymbol{s}_0,a_3} + \hat{\mathcal{X}}^{(1)}_{\boldsymbol{s}_0,a_3} + \hat{\mathcal{X}}^{(2)}_{\boldsymbol{s}_0,a_3},$$

where the two predicted residuals $\hat{\mathcal{X}}^{(1)}_{\boldsymbol{s}_0,a_3}$ and $\hat{\mathcal{X}}^{(2)}_{\boldsymbol{s}_0,a_3}$ are respectively obtained through functional UK of the observed *first-order* and *second-order* residuals in the interval a_2. In practice, one can decide how many residuals to include: (i) no residuals (Phy-RK-0), i.e. the prediction is determined by the physical model only (the only possible predictive model at step a_1), (ii) one residual (Phy-RK-1), correcting the prediction of the physical model with first-order residuals (the complete predictive model at step a_2) or (iii) two residuals (Phy-RK-2), further correcting the predictive model in a_2 with second-order residuals.

As we move towards following intervals, higher-order residual terms become available and the predictor at the i-th step can be defined as

$$\hat{\mathcal{Y}}_{\boldsymbol{s}_0,a_i}(t) = \mathcal{F}_{\boldsymbol{s}_0,a_i}(t) + \sum_{k=1}^{K_i} \hat{\mathcal{X}}^{(k)}_{\boldsymbol{s}_0,a_i}(t) \quad t \in a_i, \tag{4}$$

where $K_i \leq i-1$ is the number of residual terms employed in the i-th interval. Note that, if no residual terms (hereafter denoted by $K_i = 0$) are taken into account, the prediction coincides with

the one generated by the sole physical model. Each residual is predicted by means of functional UK from the observed residuals of the same order in the previous interval

$$\hat{\mathcal{X}}^{(k)}_{\boldsymbol{s}_0,a_i} = \sum_{\boldsymbol{s}_j \in W_{i-1}} \hat{\lambda}_j \mathcal{X}^{(k)}_{\boldsymbol{s}_j,a_{i-1}},$$

where the optimal weights $\hat{\lambda}_j$ are found by UK, being the residuals defined as

$$\mathcal{X}^{(k)}_{\boldsymbol{s}_j,a_{i-1}}(t) = \mathcal{Y}_{\boldsymbol{s}_j,a_{i-1}}(t) - \mathcal{F}_{\boldsymbol{s}_j,a_{i-1}}(t) - \sum_{l=1}^{k-1} \hat{\mathcal{X}}^{(l)}_{\boldsymbol{s}_j,a_{i-1}}(t) \quad t \in a_{i-1}.$$

The parameter K_i is chosen in each interval by selecting the number of residuals which produces the lowest median error in the previous interval. This procedure allows us to avoid overfitting, as, in this sequential framework, the choice of K_i for a_i is based on the available observations (i.e., the observations in the interval a_{i-1}), but the predictor is then employed for the test set, i.e. the production rates in the following interval a_i.

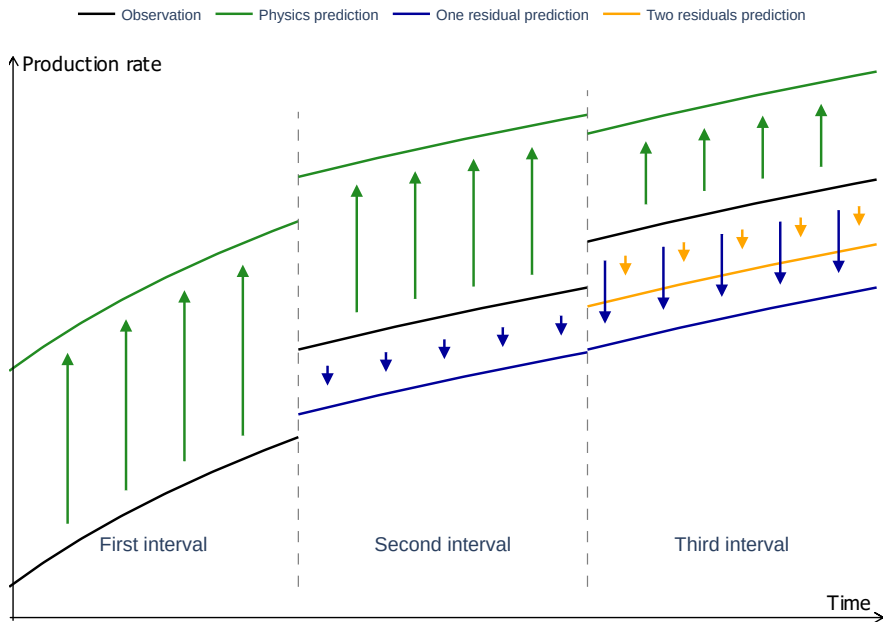


Figure 3: Sequential residuals modelization up to the third interval

3 Surrogate physical model

In the Phy-RK framework considered in this work, we consider as physical model the recently developed FlowNet, proposed by Kiærr et al. (2020). Thus, the term $\mathcal{F}_{\boldsymbol{s}_0,a_i}$ in (4) is given by the FlowNet model predictions. Instead of a full-physics 3D reservoir simulations (Aziz and Settari, 1979), FlowNet relies on a simplified reservoir modelization in terms of a flow network model, where nodes, i.e. wells, are linked by one dimensional connections. Contrary to other network models, such as the Capacitance Resistance Model (CRM, Wanderley et al. (2018)) and the Interwell Numerical SImulation Model (INSIM, Zhao et al. (2015)), in the FlowNet model the connections are discretized in cells and OPM (Flemisch et al., 2011) is used to solve the reservoir governing equation on the

network. This endows FlowNet with the capabilities of an industry-standard reservoir simulator, but with a reduced computational cost compared to high-fidelity 3D simulations.

Starting from the wells locations, FlowNet firstly builds the network, possibly adding additional nodes for refinement and to create more flow paths in the model. Given a prescribed number of additional nodes, their position is determined by a best candidate algorithm, as described in Kiærr et al. (2020). The FlowNet network used in our application is represented in Figure 4. The model is then characterized by connection parameters, such as the permeability, the porosity and the bulk-volume multiplier shared by all its cells, and some global or region-specific parameters, as oil-water contact depth, initial pressure and Corey-type relative permeability curve coefficients. The Ensemble Smoother with *Multiple* Data Assimilation (ES-MDA), originally introduced in Emerick and Reynolds (2013), is used to estimate these parameters in the history matching step. Thus, an ensemble of N_e FlowNet models is generated, independently sampling the parameters from given prior distributions. Then, according to an ES-MDA algorithm, the ensemble parameters are updated, in order to ensure that the ensemble predictions for certain quantities, for instance wells production rates or wells bottom hole pressure, match the observations in the history matching time interval. We refer to Kiærr et al. (2020) for a detailed description of the FlowNet model.

In our case, in the history matching period which corresponds to interval a_1, the bottom hole pressures of the wells active in a_1 are matched, imposing their total reservoir volume production. Therefore, as the reservoir volume rate is imposed for each well and water production is negligible, this results in a perfect match between observed oil rates and FlowNet predictions in a_1. As a consequence, the observed residuals are null in a_1 and cannot be predicted in a_2 (a spatial interpolation would just result in a null prediction over the entire domain). This implies that the interval a_3 is the first one where the Phy-RK predictions are produced with at least *first order* residuals.

The history matching procedure is performed once, considering only the wells operating in a_1, and thus the FlowNet model is not updated in the following intervals. This choice is due to two reasons: firstly, this avoids the burden of the history matching process in each interval and, secondly, it allows us to exploit the residuals with respect to the FlowNet model in the Phy-RK method, which would be null if the history matching were repeated in each interval. After the history matching, each member of the ensemble is characterized by a different set of updated parameters and is simulated to produce predictions according to the given schedule, starting from interval a_2. Note that FlowNet allows to predict production rates only at the locations of the nodes of the network. Indeed, Kiærr et al. (2020) proposed FlowNet to predict production rates at existing wells, without considering the case of infilling wells. Hence, using FlowNet as surrogate model in our framework requires that all the wells of the schedule, along all the time intervals (i.e., all the wells $W = W_1 \cup W_2 \cup ... \cup W_{N_a}$), are included in the network at the initial model calibration. Adding at later stages one or more infilling wells which were not included in the set W would require re-calibration of the model, and repetition of the history matching. The FlowNet implementation is publicly available at (Equinor, 2021).

3.1 Physics-based Residual Kriging for physical models ensembles

In Peli et al. (2022), the Phy-RK predictor is formulated for single physics-based predictions, and, therefore, we here propose an extension to the case of an ensemble of physical model predictions $\{\mathcal{F}^{(m)}_{\boldsymbol{s}_0,a_i}\}_{m=1}^{N_e}$, as for the FlowNet model. We here propose two approaches termed *Average* Phy-RK and *Multiple* Phy-RK. The *Average* Phy-RK consists in using as physical term in (4) the average ensemble prediction, i.e. $\mathcal{F}_{\boldsymbol{s}_0,a_i} = \frac{1}{N_e}\sum_{m=1}^{N_e}\mathcal{F}^{(m)}_{\boldsymbol{s}_0,a_i}$. This leads to a unique Phy-RK predictor given

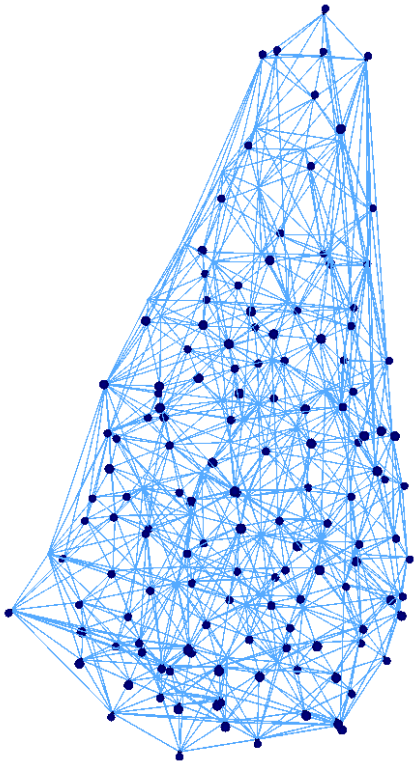


Figure 4: FlowNet network. Dark blue points represent the wells and light blue connections represent the streamtubes

by

$$\hat{\mathcal{Y}}_{\boldsymbol{s}_0,a_i} = \frac{1}{N_e}\sum_{m=1}^{N_e}\mathcal{F}^{(m)}_{\boldsymbol{s}_0,a_i} + \sum_{k=1}^{K_i}\hat{\mathcal{X}}^{(k)}_{\boldsymbol{s}_0,a_i}, \tag{5}$$

where the residuals are computed with respect to the FlowNet ensemble average. Thus, for each well at each time interval, *Average* Phy-RK returns a unique prediction given by (5).

In *Multiple* Phy-RK, a Phy-RK predictor is built for each ensemble member, leading to N_e predictions, i.e.

$$\hat{\mathcal{Y}}^{(m)}_{\boldsymbol{s}_0,a_i} = \mathcal{F}^{(m)}_{\boldsymbol{s}_0,a_i} + \sum_{k=1}^{K_i}\hat{\mathcal{X}}^{(m,k)}_{\boldsymbol{s}_0,a_i}, \quad m = 1,\cdots,N_e, \tag{6}$$

where the residuals $\mathcal{X}^{(m,k)}$ are defined with respect to the m-th FlowNet prediction. Therefore, for each well, a set of N_e predictions is computed and one can then compute statistics to characterize the prediction, as their average and quantiles, in order to quantify the uncertainty due to the different ensemble models.

4 Analysis of the reference case

Both the approaches described in Subsection 3.1 are here applied to the *reference* case, in order to predict the oil production rates described in Section 1. They are evaluated in terms of relative

error E between the observed data $\mathcal{Y}_{\boldsymbol{s}_0,a_i}$ and the predictions $\hat{\mathcal{Y}}_{\boldsymbol{s}_0,a_i}$ in each interval a_i, as

$$E_i = \frac{\int_{a_i} |\mathcal{Y}_{\boldsymbol{s}_0,a_i}(t) - \hat{\mathcal{Y}}_{\boldsymbol{s}_0,a_i}(t)|\, dt}{\int_{a_i} |\mathcal{Y}_{\boldsymbol{s}_0,a_i}(t)|\, dt}.$$

The relative error E_i is dimensionless and expresses the cumulative absolute difference between the observed and the predicted rates, divided by the total cumulative observed production.

In order to ensure the positivity of the predictions, in the following we always model the logarithm of the oil rates, the prediction for the oil rates being given by

$$\hat{\mathcal{Y}}_{\boldsymbol{s}_0,a_i} = \mathcal{F}_{\boldsymbol{s}_0,a_i} \cdot \exp\left(\sum_{k=1}^{K_i} \hat{\mathcal{X}}^{(k)}_{\boldsymbol{s}_0,a_i}\right). \tag{7}$$

In the analysis, we report the errors computed on the original oil rates, except where explicitly mentioned. We evaluate the prediction of Phy-RK with four competitors, which share with Phy-RK a relative simplicity w.r.t. a full-physics model. The first one is the pure Ordinary Kriging (OK) predictor; here, the physical model predictions in (1) are set to zero, and the model for the stochastic component $\mathcal{X}_{\boldsymbol{s}}$ is formulated by including only the intercept in the drift (2) (i.e., $L = 0$). This model is representative of a purely data-driven prediction, without considering the physics-based knowledge. The second one is the pure physical model (Phy), where the residuals are set to zero and the predictions are exclusively given by FlowNet. This approach is thus the solely surrogate physical model. Additionally, we test two versions of the Physics-based Residual Kriging, that differ only for the linear model of the drift. The Physics-based Residual Ordinary Kriging (Phy-ROK), where only the intercept is included in the drift model, and the Physics-based Residual Universal Kriging (Phy-RUK), where the linear model for the drift includes both the intercept and the pressure, computed by the FlowNet model and averaged over the time interval. Indeed, reservoir pressure at the well site is related to the produced oil rate, as expressed by the well model (Peaceman, 1977). Although already modelled in the FlowNet predictions, we here allow the residuals to depend also on the pressure, as it could be not completely exploited by the surrogate model prediction. On the other hand, if the pressure were not relevant for residuals predictions, because already completely captured by the physics, the estimate of the corresponding coefficient in the linear model for the drift would be close to zero, removing the dependence of the residual prediction on the pressure.

For the FlowNet model, the ensemble size N_e is set to 100, the number of additional nodes besides the wells in W is set to 50, the number of ES-MDA iterations is 6 with corresponding unnormalized weights $\{4, 2, 1, 1, 1, 1\}$, the porosity parameter is fixed to 0.3, the oil-water contact depth is set to 10500 feet, the initial pressure is set to 7103 psia at 10100 feet, whereas the other parameters are sampled according to distributions reported in Table 1. With this configuration, the history matching process takes approximately 4 hours on a laptop equipped with an AMD Ryzen 5 2500U processor, 4 cores at 2 GHz and 8Gb of RAM.

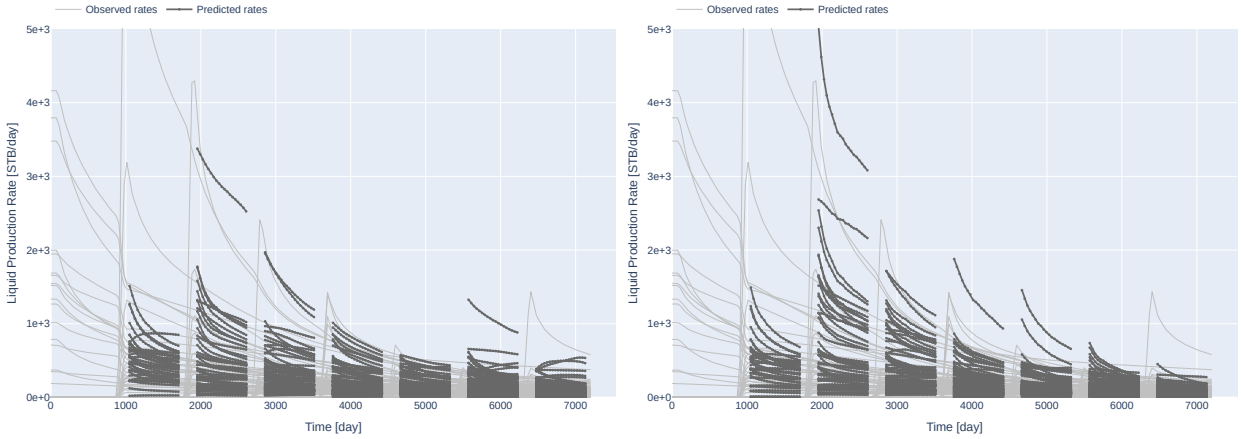
As production rates are observed along the time intervals, we could use Phy-ROK and Phy-RUK with an increasing number of residual terms. The selection of the number K_i of residuals employed in each interval is done by picking the best predictor for the previous interval, in terms of median relative prediction error. Thus, we compare the four approaches, given the selected number of residuals for Phy-ROK and Phy-RUK.

Firstly, in Figure 5, we show, from a qualitative point of view, the predictions for the two approaches *Average* and *Multiple*. In each interval, the optimal number of residuals and the optimal

Parameter	Distribution	Min	Max
Permeability	Log-uniform	0.001	30
Bulk volume multiplier	Uniform	0.171	0.271
Relative permeability			
$S_{w,irr}/S_{w,l}$	Uniform	0.01	0.2
$S_{w,cr}/S_{o,cr}$	Uniform	0.01	0.3
$kr_{w,end}$	Uniform	0.5	0.7
$kr_{o,end}$	Uniform	0.6	0.9
n_w	Uniform	1.5	4.5
n_o	Uniform	2.5	4.5

Table 1: Prior distributions of the FlowNet parameters

drift model are used, as we will describe in Subsections 4.1 and 4.2. For the *Multiple* approach, we report the average predictions, as for each well N_e predictions are obtained. We note a general good agreement between the predictions and the observations, with similar results between the two approaches. In the next subsections, we quantify the prediction errors, quantitatively comparing the two approaches.



(a) *Average* Phy-RK (b) *Multiple* Phy-RK

Figure 5: Predictions of the *Average* Phy-RK (left) and *Multiple* Phy-RK (right) and the observed oil production rates for the *reference* case. For the *Multiple* approach, the average prediction is displayed

4.1 Average Phy-RK predictions

The *Average* Phy-RK predictor requires to run the FlowNet models ensemble according to the drilling schedule, taking its average prediction, and then modeling its residuals as explained in Section 2. For the Phy-RUK predictor, the pressure, computed by the FlowNet simulations and averaged over the ensemble members, is employed in the linear model for the drift. The selected numbers of residuals K_i in each interval for the Phy-ROK and Phy-RUK predictors, are reported in Figure 6. For both of them, the numbers of selected residuals is coherent, with two residuals in the last two intervals and one in the others (in the second one, we assume to use the Phy-RK predictor as soon as it becomes available). In Figure 7a, we show the relative errors of prediction along each

time interval for OK, Phy, Phy-ROK and Phy-RUK, with the selected numbers of residuals. In these boxplots, each sample unit is a different well operating in the corresponding interval. Among the four different predictors, one has to decide in advance which of them has to be employed in each interval. Recall that this is done by looking at the method that best performed in the previous interval, in terms of median relative error. The choice for each interval, except the interval a_2 where the aforementioned criterion is not applicable, is shown in Figure 7a, pointed by the *selected model* arrow. We note that the best predictors are the Phy-ROK and Phy-RUK, with no major difference between them. These predictors outperform both the OK and Phy predictors, indicating that Phy-RK successfully combines them. In particular, we note that the pure FlowNet approach results in the worst predictions. Furthermore, we note that, on the contrary of Phy and OK, Phy-ROK and Phy-RUK exhibit a decreasing median predictive error as we move towards final intervals, that is as more data become available. Concerning the computational times, the main burden of the Phy-RK approach is due to the prediction of the FlowNet ensemble, which takes approximately 30 minutes, whereas the prediction of the residuals requires roughly 1 second. Computational times are reported in Table 2.

4.2 Multiple Phy-RK predictions

Using the *Multiple* Phy-RK approach, the Phy-RK predictor is applied to each FlowNet model in the ensemble, ending up with N_e predictions for each well. Thus, in order to compare the results with the *Average* Phy-RK, we show the errors obtained taking their average, whereas their variability is exploited for uncertainty quantification. In this case, for the Phy-RUK predictor, the pressure, averaged over time, computed by each FlowNet model is employed as regressor in the linear model for the drift of the corresponding ensemble member. The selected number of residuals K_i in each interval for the Phy-ROK and Phy-RUK predictors, is reported in Figure 8. In this case, the number of selected residuals coincides for Phy-ROK and Phy-RUK, except for the last interval, where two residual terms are chosen for Phy-ROK and one residual term for Phy-RUK. In Figure 7b, we show the relative errors produced by the *Multiple* Phy-RK, taking its average prediction, along each time interval, comparing OK, Phy, Phy-ROK and Phy-RUK, with selected numbers of residuals. Also in this case, Phy-Rk results to be the best predictor in all the intervals. Furthermore, we note that the *Average* Phy-RK predictor and the average of *Multiple* Phy-RK predictions lead to similar relative errors. However, *Multiple* Phy-RK has an higher computational cost, as residuals are predicted N_e times, whereas for *Average* Phy-RK only one residual is predicted. Its cost is comparable to *Average* Phy-RK, as the main burden comes from the run of the ensemble of physical models. On the other hand, *Multiple* Phy-RK allows one to quantify the uncertainty associated to the prediction. Computational times are reported in Table 2. The Phy computational cost comes from the run of the whole ensemble. For the *Multiple* Phy-RK approach, the residual prediction is repeated N_e times, implying a higher computational time.

FlowNet	**Prediction**			
history matching	OK	Phy	*Average* Phy-RK	*Multiple* Phy-RK
4 hours	1 second	30 minutes	30 minutes + 1 second	30 minutes + 100 seconds

Table 2: Computational times of the FlowNet history matching and the prediction steps. Times are obtained on a laptop equipped with an AMD Ryzen 5 2500U processor, 4 cores at 2 GHz and 8Gb of RAM.

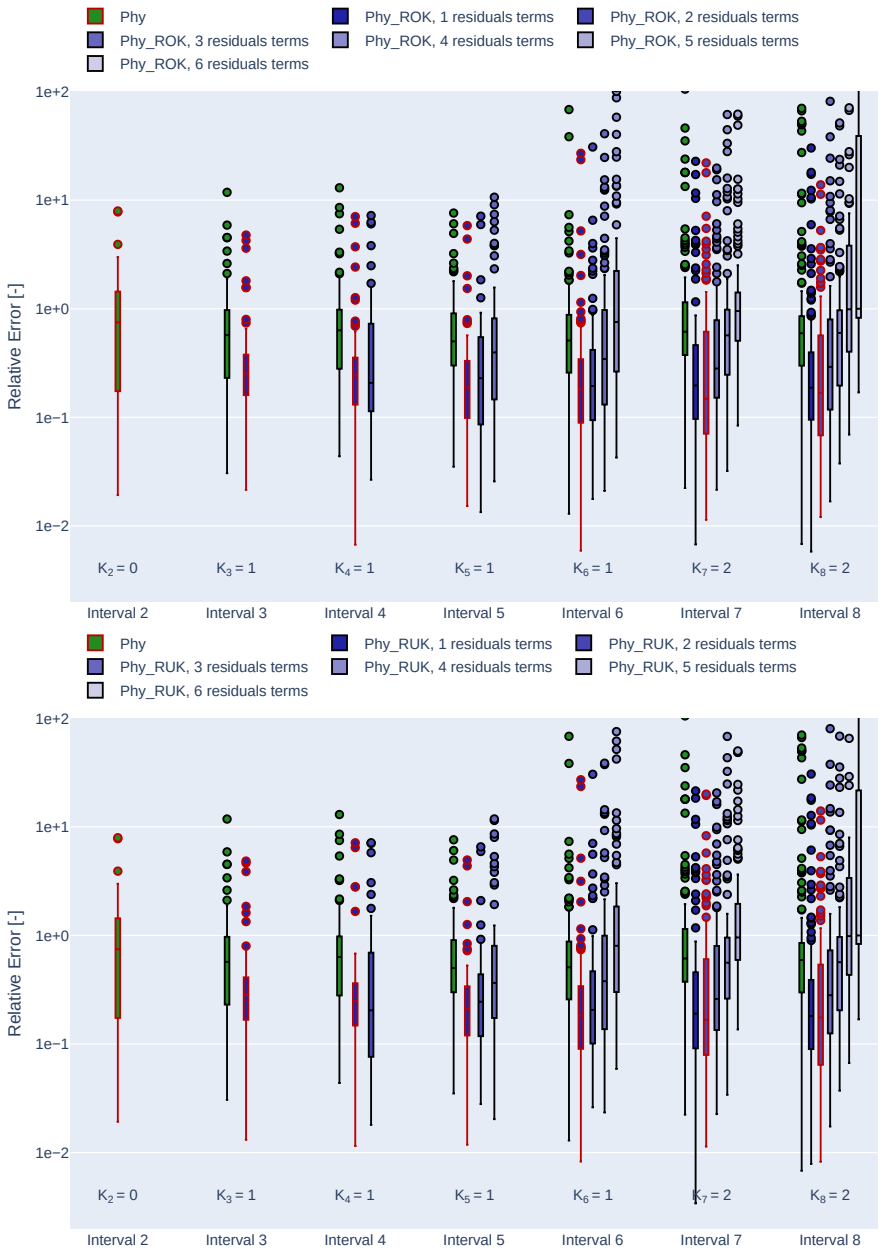
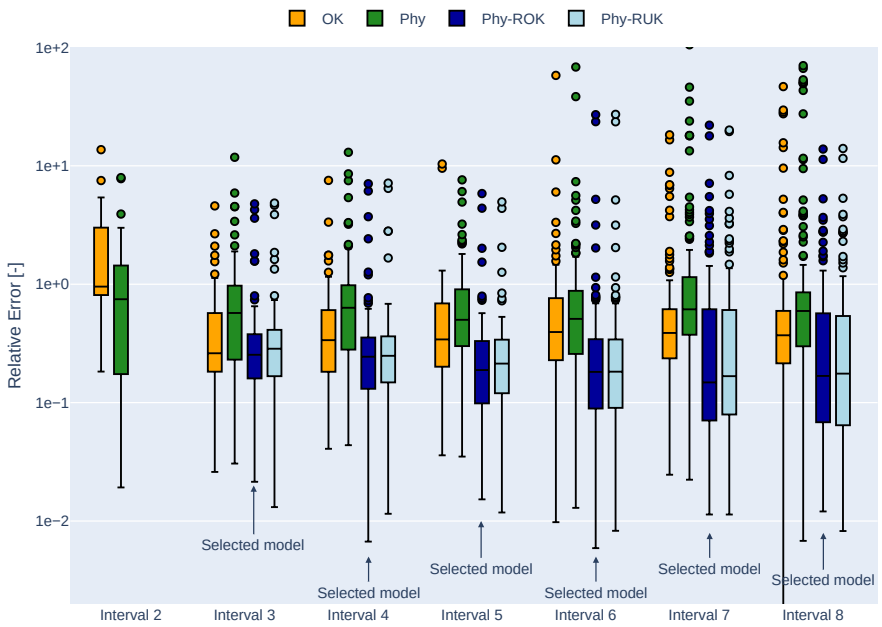


Figure 6: Relative prediction error boxplots for the *reference* case produced by Phy-ROK (left panel) and Phy-RUK (right panel) with different numbers of residual terms for the *Average* Phy-RK approach. The boxplots with red outlines are the ones corresponding to the selected residuals number

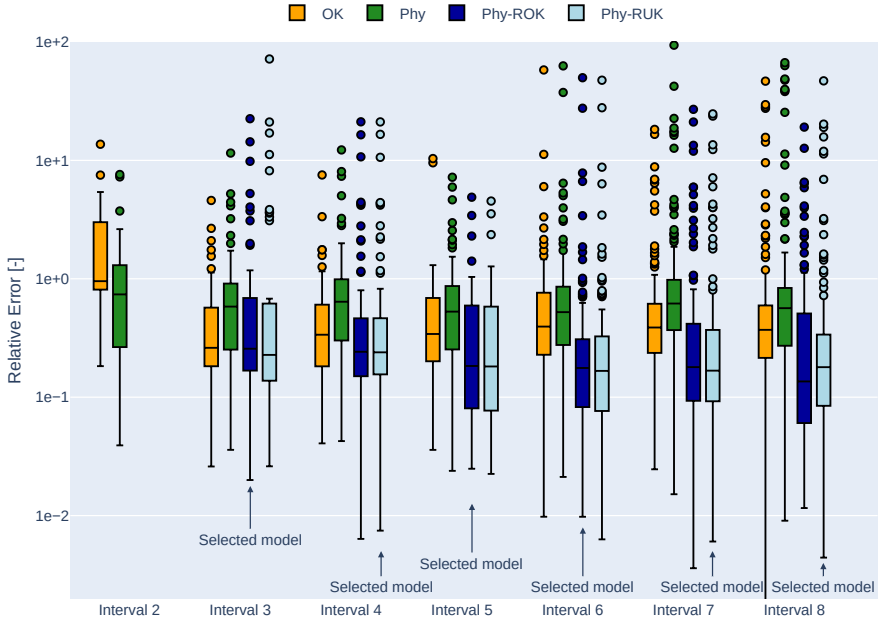
4.3 Uncertainty quantification

The uncertainty associated with the prediction of the *Multiple* Phy-RK approach is due to two components: the variability of the FlowNet ensemble and the variance associated with the kriging prediction of the residuals. Throughout this analysis, we investigate both the variances in order to evaluate the contribute of the different sources of uncertainty.

On one hand, one can evaluate the distribution obtained using kriged residuals, where for each



(a) *Average* Phy-RK



(b) *Multiple* Phy-RK

Figure 7: Relative errors boxplots in the *reference* case for *Average* Phy-RK (left) and *Multiple* Phy-RK (right) produced by OK, Phy, Phy-ROK and Phy-RUK. The selected model arrow points to the model that results to be the one with lowest relative error in the previous interval. Note that the y-axis is represented on a log scale

member of the FlowNet ensemble, the prediction is given by equation (6), for a total of N_e predictions. In this way, the uncertainty associated to residuals prediction is not considered. On the other hand, in order to quantify the total prediction variance, we perform multiple stochastic simulation of the residuals, from their approximate distribution conditioned to the observations, as illustrated

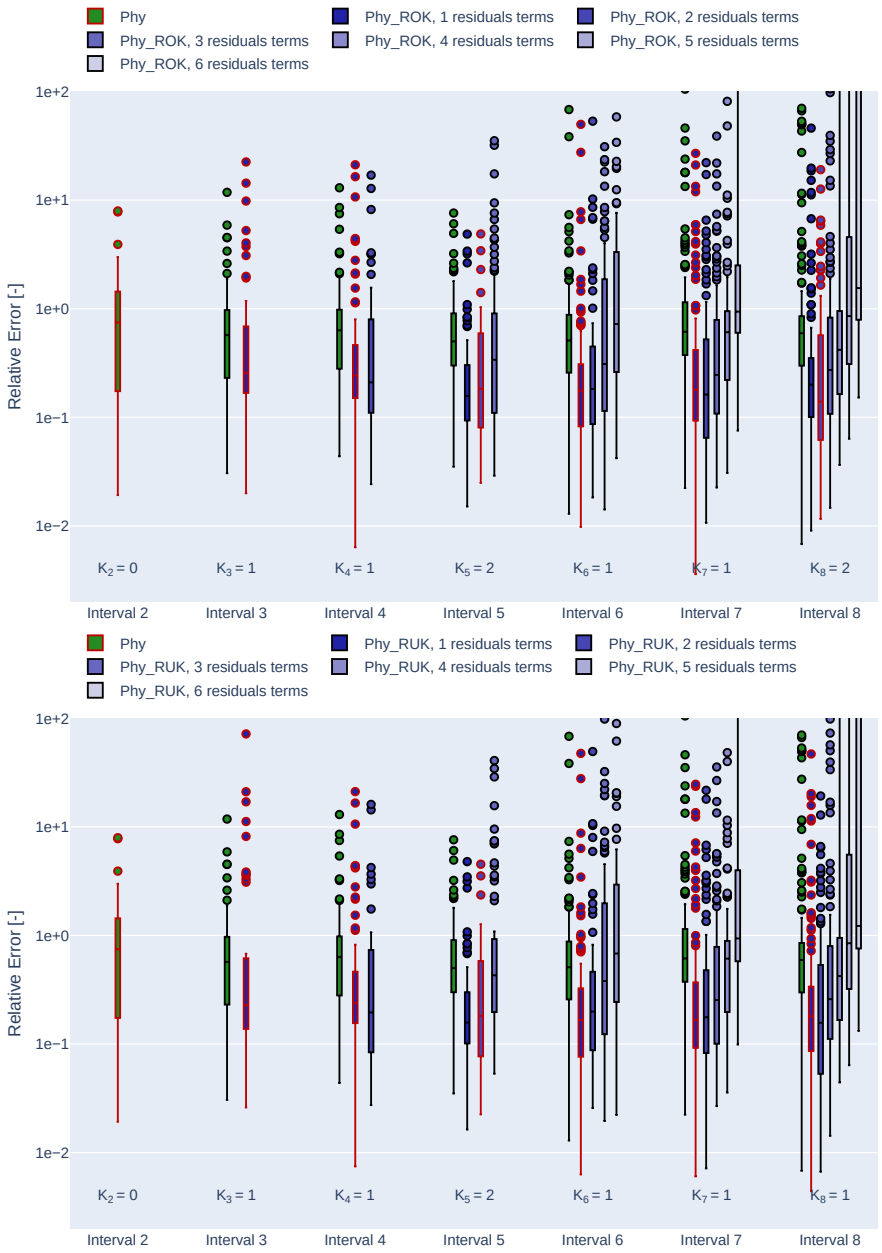


Figure 8: Relative prediction error boxplots for the *reference* case produced by Phy-ROK (left panel) and Phy-RUK (right panel) with different numbers of residual terms for the *Multiple* Phy-RK approach. The boxplots with red outlines are the ones corresponding to the selected residuals number

in Menafoglio et al. (2016c). Hence, for each FlowNet ensemble model, $N_r = 100$ Montecarlo samples of the residual functional field are simulated, conditional to the residuals observed at the operating wells in the previous interval. In this way, $N_e \times N_r$ simulations are produced for each well, allowing to characterize their distribution, and thus the total prediction variance. In order to limit the complexity of the analysis, we carry out the uncertainty quantification on the logarithm of oil production rates predicted with Phy-ROK with one residual, that, as illustrated in Subsection 4.2, is often the best predictor.

We start presenting the results on a single well, the one pointed by the arrow in Figure 1 and drilled at the third interval. In Figure 9, we display the point-wise median and 5% and 95% percentiles of the predictions obtained with kriged (left) and conditionally simulated (right) residuals. Therefore, on the left, only the ensemble variance is displayed, whereas on the right also the kriging variance is considered. We note that there is a difference between the two distributions only in the first interval after well drilling, with a higher variance associated with simulated residuals. This is due to the fact that, after that interval, the residual has been observed (because the well has been opened), and thus its simulations and kriged prediction both coincide with the observed one.

We now iterate the same argument for all the wells. For the sake of visualization, we take the time-average predictions over each interval, and we then compute their (scalar) quantiles. In Figure 10, we show the median prediction (Fig. 10a) and the Inter Quartiles Range (IQR), i.e. the difference between the $75th$ and the $25th$ percentiles, starting from interval a_3, of the FlowNet ensemble (Fig. 10b), and of the *Multiple* Phy-ROK predictions with kriged (Fig. 10c) and conditionally simulated (Fig. 10d) residuals. We note a decrease in the median prediction as we move forward in time, coherently with the observed rates in Figure 2. Concerning the variance, the IQRs are similar between the FlowNet ensemble and the predictions with kriged residuals, with roughly a doubling with conditionally simulated residuals. We point out that, when residuals are conditionally simulated, the variance is higher for the newly drilled wells than for the ones already operating, because, in this case, the predicted residual coincides with the observed one in the previous interval, due to the interpolation property of the kriging predictor. Furthermore, we observe that the overall IQR tends to reduce towards last intervals, where the number of observations is higher.

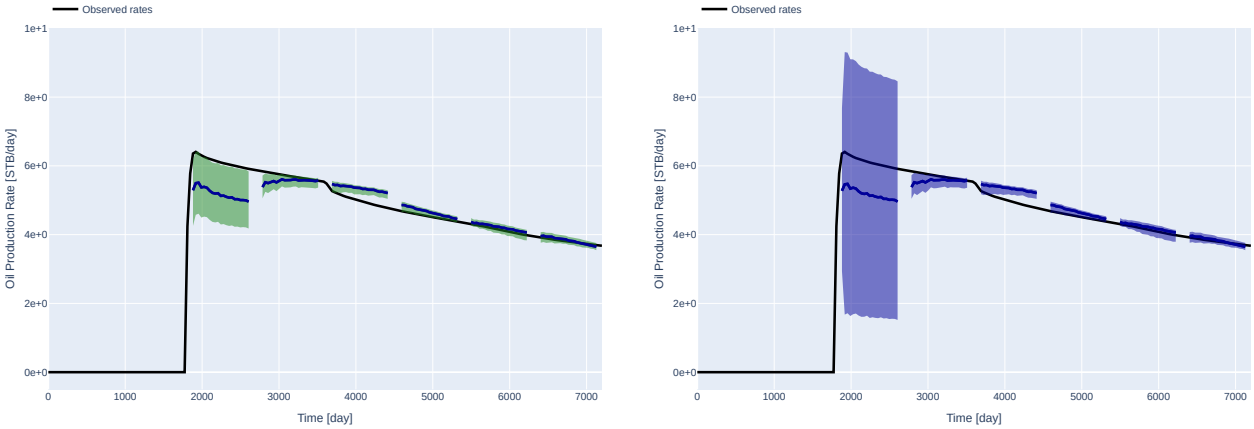
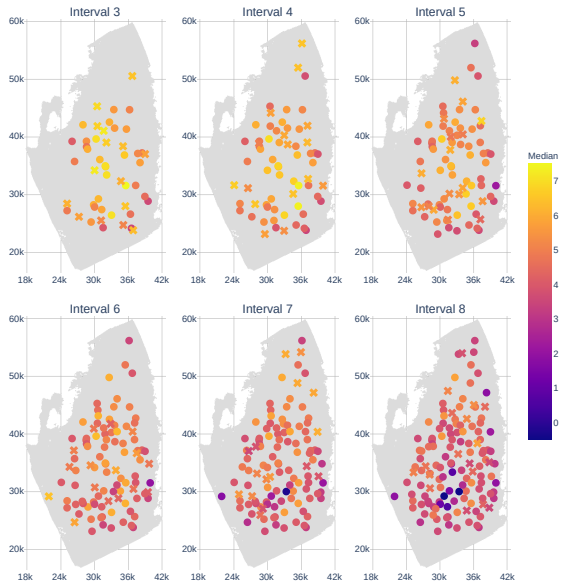


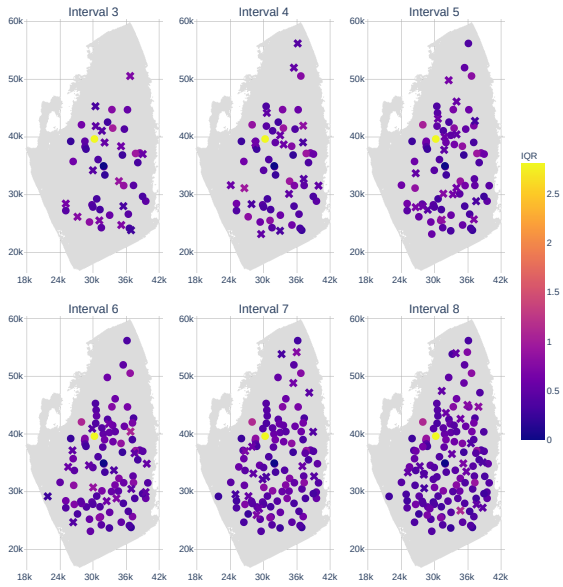
Figure 9: Point-wise median, 5% and 95% quantiles of the ensemble predictions of *Multiple* Phy-RK with kriged (left) and conditionally simulated (right) residuals, for well PRED64

5 Additional analyses

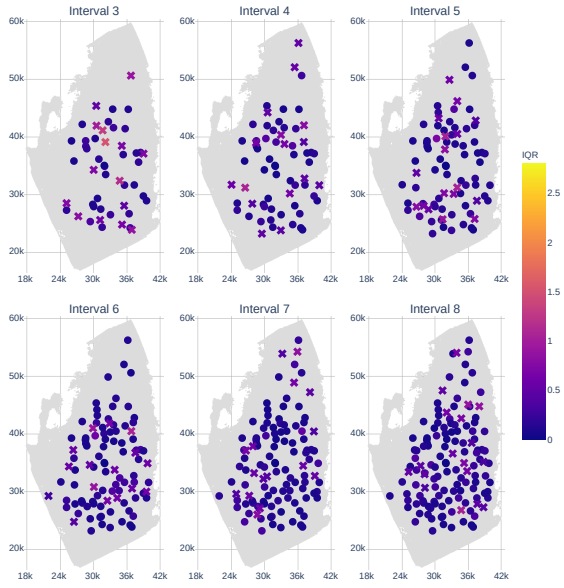
We here present further analyses to illustrate the application of Phy-RK in situations where the wells schedule is more complex than the one of the *reference* case, illustrated in Subsection 1. We thus introduce two problems, where the reservoir properties are the same of the *reference* case, with the only change in the wells schedule. The first example, in Subsection 5.1, deals with a variable number of wells drilled in each interval with same time interval length, whereas the second one, in Subsection 5.2, presents a realistic schedule with irregular time intervals in the drilling schedule.



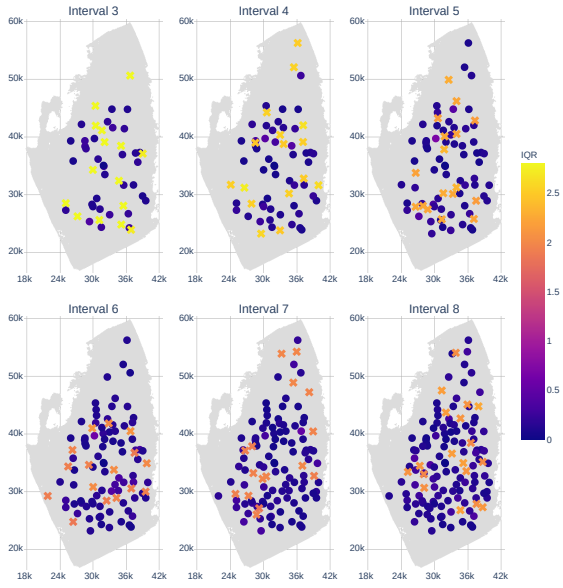
(a) Median Phy-ROK prediction



(b) IQR FlowNet predictions



(c) IQR Phy-ROK / kriged residuals



(d) IQR Phy-ROK / simulated residuals

Figure 10: Median and IQR of the time-averaged *Multiple* Phy-ROK predictions with one residual. We indicate with crosses the wells that are open in the corresponding interval and with circles wells already operating

5.1 Variable number of wells drilled in each interval

Hereafter, we analyze a case where the wells are drilled in 8 intervals, for a total period of 20 years, but their number differs among the intervals. In particular, the number of wells being drilled in

each interval is respectively 74, 14, 10, 3, 19, 4, 6 and 6. This choice, compared to the *reference* case, alters both the observed rates, shown in Figure 11 and obtained with a new OPM simulation with the updated schedule, and the FlowNet physical model. Here, the same prior distributions of the parameters are kept, but the history matching is repeated because the oil rates in the first interval simulated with *variable number* schedule differ from the *reference* case. Phy-RK is then applied to obtain the predictions in each interval, without any modification to the approach, except for the different schedule.

We present only the results obtained with the *Average* Phy-RK, as the *Multiple* Phy-RK produces similar results, when selecting the number of residuals as in the *Average* case. We then report the relative prediction errors of the four different methods (OK, Phy, Phy-ROK, Phy-RUK) in Figure 12, whereas the number of residual terms for Phy-ROK and Phy-RUK has been selected according to Figure 13. As in the *reference* case, Phy-ROK and Phy-RUK result to be the best predictors, showing lower errors than in the *reference* case, probably due to the higher number of observations available from the beginning in interval a_1, which is beneficial for both FlowNet history matching and Phy-RK.

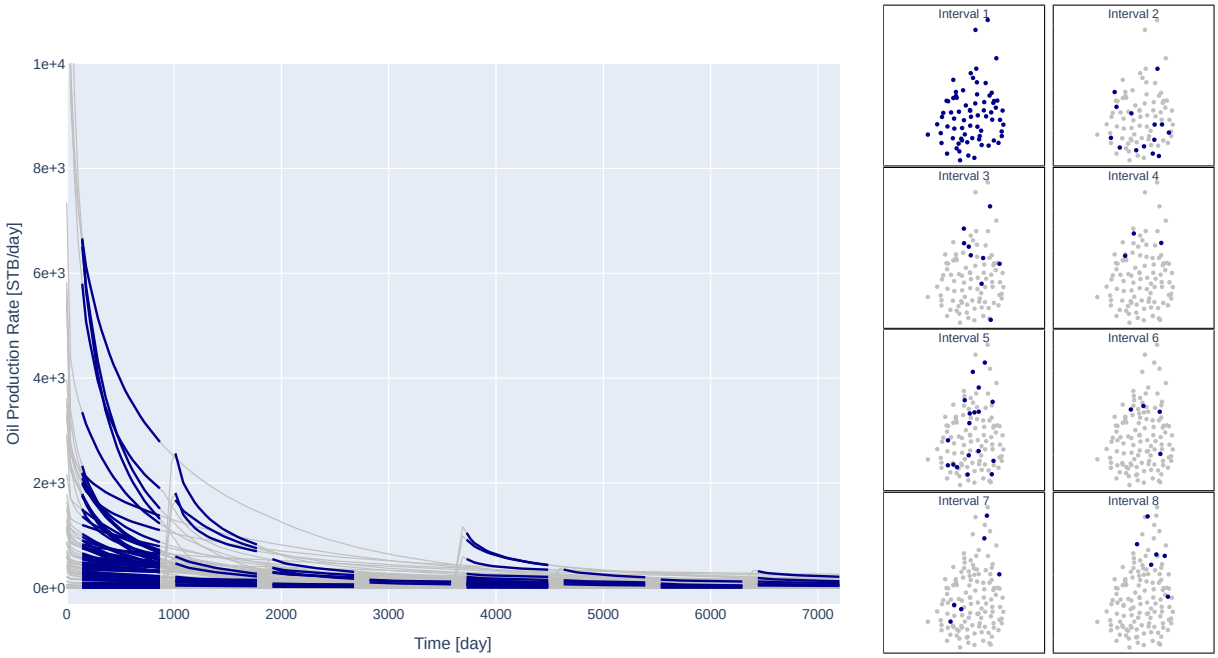


Figure 11: Production rates and schedule for the *variable number* scenario described in Subsection 5.1. Left panel: blue lines represent the production rates corresponding to newly drilled wells in each interval, gray lines correspond to wells already operating. Right panel: blue circles represent the newly drilled producers and gray circles the already operating wells

5.2 Realistic schedule

We now show how the Phy-RK predictor can be applied in presence of a realistic schedule, where time intervals have different lengths and an arbitrary number of wells is drilled at their starting points. In the *realistic* schedule, after an initial ramp up of two years that results in the opening of about 50% of the wells, a period of 10 years without new wells follows and, after that, infilling wells are put into production. In this phase, three groups of wells are drilled every two months, followed by a period of eight months without new wells. This pattern is repeated six times, with a

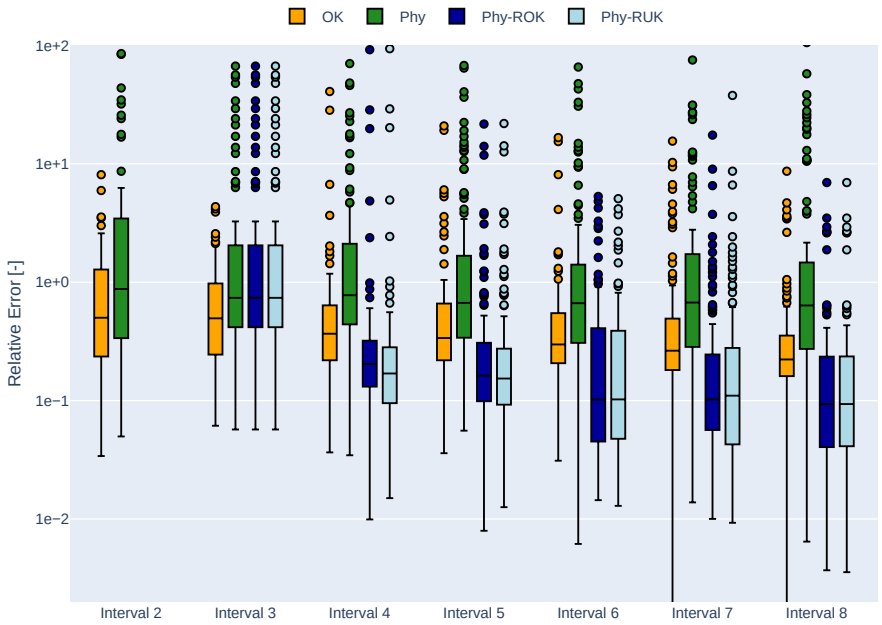


Figure 12: Relative errors boxplots of *variable* scenario for *Average* Phy-RK produced by OK, Phy, Phy-ROK and Phy-RUK, with selected number of residuals. Note that the y-axis is represented on a log scale

final longer time window of roughly three years. The schedule is depicted in Figure 14 for clarity. The history matching period of the FlowNet model spans up to the beginning of the infilling phase.

Whenever the prediction intervals have different lengths, Phy-RK, as described in Section 2, cannot be directly applied, as it requires a common domain length for observed and predicted oil rates. Therefore, we here propose an extension of Phy-RK to overcome this limitation. We distinguish two cases: (i) if the prediction interval a_i is shorter than the observed one a_{i-1} and (ii) if the prediction interval a_i is longer than a_{i-1}. In the former case, Phy-RK can be directly applied, by truncating the prediction to fit the domain. In the latter case, one needs a way to extend the prediction, that is naturally produced by Phy-RK only in the initial part of interval a_i whose length corresponds to the one of a_{i-1}. We propose to subdivide the interval a_i into subintervals $\{b_j^i\}_{j=1}^J$, where J is the upper part of the ratio between the length of a_i and a_{i-1}, i.e. $J = \lceil |a_i|/|a_{i-1}| \rceil$, whose length is equal to a_{i-1}, except for the last than could be shorter. Limiting the analysis to first order residuals, in each subinterval the Phy-RK prediction is given by

$$\hat{\mathcal{Y}}_{\boldsymbol{s}_0, b_j^i} = \mathcal{F}_{\boldsymbol{s}_0, b_j^i} + \hat{\mathcal{X}}^{(1)}_{\boldsymbol{s}_0, b_j^i}, \tag{8}$$

where $\hat{\mathcal{X}}^{(1)}_{\boldsymbol{s}_0, b_j^i}$ is given by functional Universal Kriging applied on the residuals observed in a_{i-1}. Thus, the prediction in each subinterval is given by the sum of the physical model prediction in that subinterval and of the predicted residual, which is not updated among the subintervals. The predictions are possibly truncated in the last subinterval b_J^i, as this can be shorter than a_{i-1}. Whenever subintervals are introduced, the resulting Phy-RK prediction in a_i are discontinuous, exhibiting jumps at the interfaces between subintervals. Therefore, residuals of order greater than one would propagate these discontinuities, implying highly irregular predictions. Hence, the case including higher order residuals is not studied further here.

We now apply this approach to the aforedescribed *realistic* schedule. We introduce an approx-

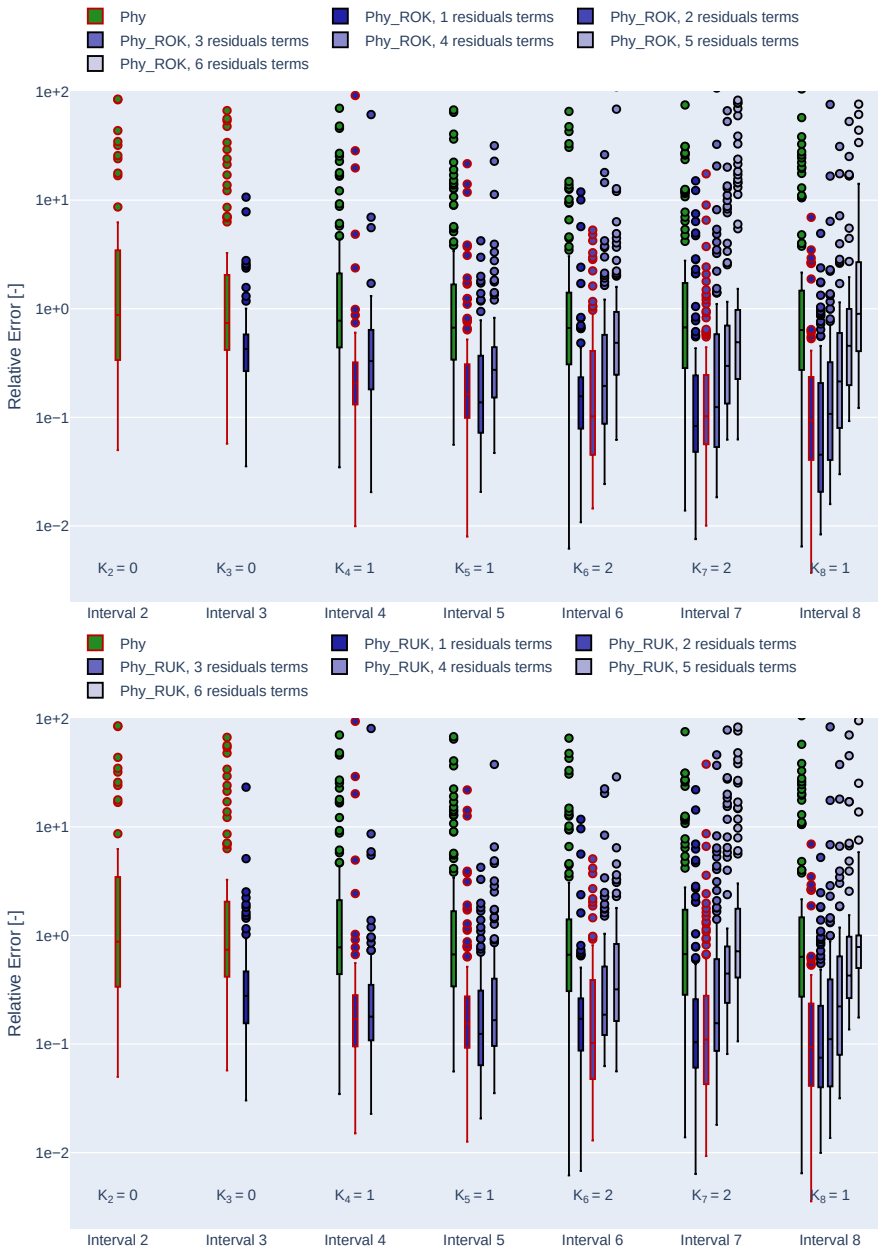


Figure 13: Relative prediction error boxplots for *variable* scenario produced by *Average* Phy-ROK (left panel) and *Average* Phy-RUK (right panel) with different numbers of residual terms for the *Average* Phy-RK approach. The boxplots with red outlines are the ones corresponding to the selected residuals number

imation to avoid too much discontinuity in the predictions, pretending that groups of wells drilled at distance of two months are actually drilled at the same moment. This approximation is used only in the prediction step, where production rates of each triplet of wells groups (dashed blue lines in Figure 14) are aligned to the corresponding starting point of the interval (previous solid blue line). Thanks to the alignment, Phy-RK is applied in each of the six intervals, which result to be equally longer, except for the last one, where the prediction is extended following the aforedescribed procedure.

Once the predictions are computed, predicted oil rates are aligned back to the original schedule. We depict the predictions in Figure 15a and the corresponding prediction errors in Figure 15b. In this case, Phy-RK exhibits prediction errors similar to OK, with both methods that are much better than FlowNet predictions.

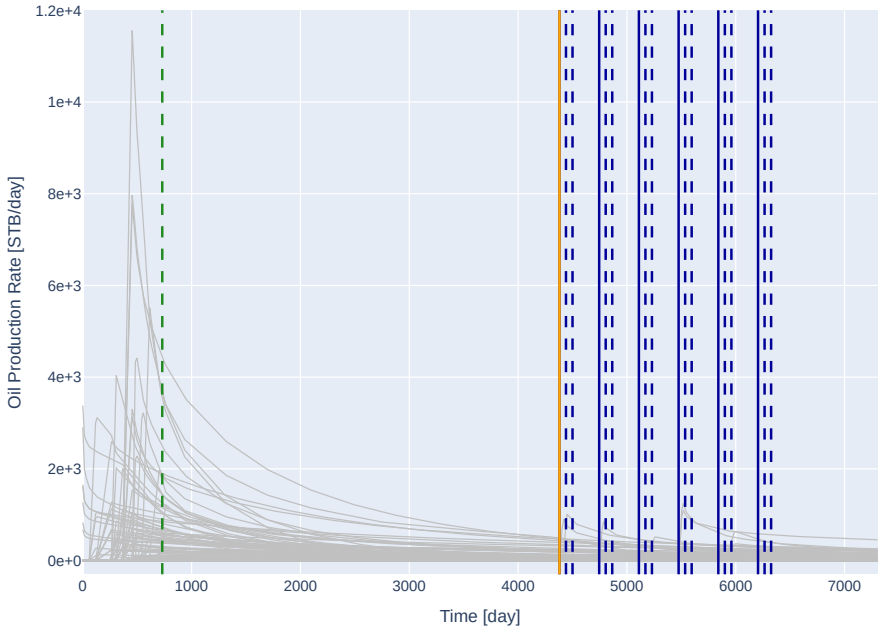
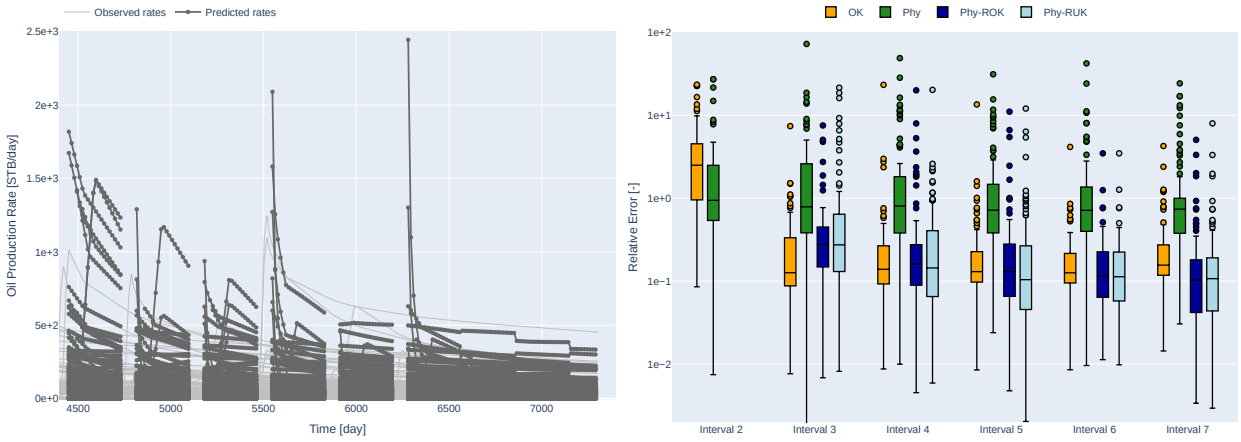


Figure 14: Oil production rates for the *realistic* scenario described in Subsection 5.2. Vertical lines represent the well schedule: green dashed line delimits the end of the initial ramp up, orange line delimits the history matching period, and blue lines the infilling phase



(a) (b)

Figure 15: (a) *Average* Phy-RK predictions and observed oil production rates in the *realistic* case. Zoom over the last six intervals. (b) Relative errors boxplots for *Average* Phy-RK produced by OK, Phy, Phy-ROK and Phy-RUK, with selected number of residuals in the *realistic* case.

6 Conclusions and future perspective

In this work, we present a novel surrogate model to forecast oil production rates in a mature single-phase hydrocarbon reservoir. Our approach allows one to take advantage from past production data to build a data-driven predictor, which exploits also a physical modelization of the reservoir. In particular, the recently developed Physics-based Residual Kriging predictor is considered, including a physical term model through FlowNet, the latter representing a very cutting edge network-based reservoir model. Thanks to the integration of these two components, an accurate yet computationally efficient predictor is derived, as it does not require a 3D full-physics reservoir modelization. Furthermore, we here describe how ensembles of surrogate predictors can be obtained and exploited, even quantifying the uncertainty associated with the predictions.

We apply the derived surrogate model to a chalk reservoir, analyzing three different scenarios in terms of the drilling schedule of the wells. Furthermore, we propose an extension of the Phy-RK predictor that allows one to apply it in presence of arbitrary wells schedule. In all the cases, Phy-RK outperforms the pure geostatistical and pure physical model counterparts in terms of prediction error, and it proves to be a promising predictor for oil rates in mature reservoirs.

As future research, the presented approach could be extended to two-phase problems, where oil and water production rates have to be jointly predicted. In this context, the Phy-RK predictor can still exploit the FlowNet physical model, as it implements all the functionalities of multi-phase reservoir simulators. Being the underlying physics more complex, the FlowNet calibration in the history matching phase would require particular attention to ensure accurate predictions.

Acknowledgments

The authors gratefully thank Alberto Cominelli and Ernesto Della Rossa (Eni S.p.A) for their precious contribution to this work.

Riccardo Peli, Alessandra Menafoglio and Piercesare Secchi gratefully acknowledge the financial support of Eni S.p.A [memorandum of agreement number 4400007601] and Eni Corporate University S.p.A.[grant number 7299/2018].

References

Ahmed, T. and Meehan, N. (2012). *Advanced Reservoir Management and Engineering*. Gulf Professional Publishing, Houston.

Alfonzo, M. and Oliver, D. (2019). Evaluating prior predictions of production and seismic data. *Computational Geosciences*, 23:1–17.

Aziz, K. and Settari, A. (1979). *Petroleum Reservoir Simulation*. Springer Netherlands.

Caballero, W., Giraldo, R., and Mateu, J. (2013). A universal kriging approach for spatial functional data. *Stochastic Environmental Research and Risk Assessment*, 27:1553–1563.

Coutinho, E. J. R., Dall'Aqua, M., and Gildin, E. (2021). Physics-aware deep-learning-based proxy reservoir simulation model equipped with state and well output prediction. *Frontiers in Applied Mathematics and Statistics*, 7:49.

Dake, L. (1978). *Fundamentals of Reservoir Engineering*. Elsevier Science, Amsterdam.

Emerick, A. and Reynolds, A. (2012). History matching time-lapse seismic data using the ensemble kalman filter with multiple data assimilations. *Computational Geosciences*, 16:639–659.

Emerick, A. and Reynolds, A. (2013). Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55:3–15.

Equinor (2021). FlowNet. `https://github.com/equinor/flownet`.

Flemisch, B., Flornes, K., Lie, K.-A., and Rasmussen, A. (2011). OPM: the Open Porous Media initiative. *AGU Fall Meeting Abstracts*.

Ghasemi, M., Ibrahim, A., and Gildin, E. (2014). Reduced Order Modeling In Reservoir Simulation Using the Bilinear Approximation Techniques. *SPE Latin America and Caribbean Petroleum Engineering Conference*, Day 1 Wed, May 21, 2014. D011S007R003.

Gilman, J. R. and Chet, O. (2013). *Reservoir simulation : history matching and forecasting*. Society of Petroleum Engineers, Richardson Texas.

Guo, Z. and Reynolds, A. (2019). Insim-ft in three-dimensions with gravity. *Journal of Computational Physics*, 380.

Guo, Z., Reynolds, A. C., and Zhao, H. (2017). A Physics-Based Data-Driven Model for History Matching, Prediction, and Characterization of Waterflooding Performance. *SPE Journal*, 23(02):367–395.

Han, D. and Kwon, S. (2021). Application of machine learning method of data-driven deep learning model to predict well production rate in the shale gas reservoirs. *Energies*, 14(12).

Jin, Z., Liu, Y., and Durlofsky, L. (2020). Deep-learning-based surrogate model for reservoir simulation with time-varying well controls. *Journal of Petroleum Science and Engineering*, 192:107273.

Kiærr, A., Lødøen, O. P., De Bruin, W. J., Barros, E., and Leeuwenburgh, O. (2020). Evaluation of a data-driven Flow Network model (FlowNet) for reservoir prediction and optimization. *Conference Proceedings, ECMOR XVII*, 2020(1):1–18.

Menafoglio, A., Grujic, O., and Caers, J. (2016a). Universal kriging of functional data: Trace-variography vs cross-variography? application to gas forecasting in unconventional shales. *Spatial Statistics*, 15.

Menafoglio, A., Grujic, O., and Caers, J. (2016b). Universal Kriging of functional data: trace-variography vs cross-variography? Application to gas forecasting in unconventional shales. *Spatial Statistics*, 15:39–55.

Menafoglio, A., Guadagnini, A., and Secchi, P. (2016c). Stochastic simulation of soil particle-size curves in heterogeneous aquifer systems through a bayes space approach. *Water Resources Research*, 52(8):5708–5726.

Menafoglio, A., Secchi, P., and Dalla Rosa, M. (2013). A Universal Kriging predictor for spatially dependent functional data of a Hilbert Space. *Electronic Journal of Statistics*, 7:2209–2240.

Oliver, D. S. (2020). Diagnosing reservoir model deficiency for model improvement. *Journal of Petroleum Science and Engineering*, 193:107367.

Oliver, D. S., Reynolds, A. C., and Liu, N. (2008). *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press.

Pathak, A. K. (2021). *Petroleum Reservoir Management: Considerations and Practices*. CRC Press, Boca Raton.

Peaceman, D. W. (1977). *Fundamentals of numerical reservoir simulation*. Elsevier Scientific Pub. Co., New York.

Peli, R., Menafoglio, A., Cervino, M., Dovera, L., and Secchi, P. (2022). Physics-based residual kriging for dynamically evolving functional random fields. *Stochastic Environmental Research and Risk Assessment*, pages 1–18.

Ramsay, J. O.; Silverman, B. W. (2005). *Functional data analysis*. Springer, New York, second edition.

Tang, H., Volkov, O., Tchelepi, H. A., and Durlofsky, L. J. (2021). Reduced-Order Modeling in a General Reservoir Simulation Setting. *SPE Western Regional Meeting*, Day 2 Wed, April 21, 2021. D021S008R001.

Wanderley, R., Gildin, E., Jensen, J., Lake, L., and Kabir, S. (2018). A state-of-the-art literature review on Capacitance Resistance Models for reservoir characterization and performance forecasting. *Energies*, 11:33–68.

Yoon, S., Alghareeb, Z. M., and Williams, J. R. (2016). Hyper-Reduced-Order Models for Subsurface Flow Simulation. *SPE Journal*, 21(06):2128–2140.

Zhao, H., Kang, Z., Zhang, X., Sun, H., Cao, L., and Reynolds, A. (2015). INSIM: A data-driven model for history matching and prediction for waterflooding monitoring and management with a field application. In *SPE Reservoir Simulation Symposium 2015*, volume 1.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

43/2022 Zappon E.; Manzoni A.; Gervasio P.; Quarteroni A.
A reduced order model for domain decompositions with non-conforming interfaces

41/2022 Arnone, A.; Ferraccioli, F.; Pigolotti, C.; Sangalli, L.M.
A roughness penalty approach to estimate densities over two-dimensional manifolds

40/2022 Fumagalli, A.; Patacchini, F. S.
Well-posedness and variational numerical scheme for an adaptive model in highly heterogeneous porous media

42/2022 Gatti, F.; Fois, M.; de Falco, C.; Perotto, S.; Formaggia, L.
Parallel simulations for fast-moving landslides: space-time mesh adaptation and sharp tracking of the wetting front

39/2022 Ferro, N.; Perotto, S.; Gavazzoni, M.
A new fluid-based strategy for the connection of non-matching lattice materials

37/2022 Boon, W. M.; Fumagalli, A.
A Reduced Basis Method for Darcy flow systems that ensures local mass conservation by using exact discrete complexes

38/2022 Burzacchi, A.; Landrò, M.; Vantini, S.
Object-oriented Classification of Road Pavement Type in Greater Maputo from Satellite Images

35/2022 Perotto, S.; Bellini, G.; Ballarin, F.; Calò, K.; Mazzi, V.; Morbiducci, U.
Isogeometric hierarchical model reduction for advection-diffusion process simulation in microchannels

36/2022 Vaccaro, F.; Brivio, S.; Perotto, S.; Mauri, A.G.; Spiga, S.
Physics-based Compact Modelling of the Analog Dynamics of HfOx Resistive Memories

34/2022 Antonietti, P.F.; Vacca, G.; Verani, M.
Virtual Element method for the Navier-Stokes equation coupled with the heat equation