

Functionalized Data Operator Model for System Analysis and Forecasting

George Danko

University of Nevada, Reno Reno, Nevada, USA

Email: danko@unr.edu

How to cite this paper: Danko, G. (2022) Functionalized Data Operator Model for System Analysis and Forecasting. *Applied Mathematics*, 13, 988-1021.
<https://doi.org/10.4236/am.2022.1312062>

Received: November 2, 2022

Accepted: December 27, 2022

Published: December 30, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A deconvolution data processing is developed for obtaining a Functionalized Data Operator (FDO) model that is trained to approximate past and present, input-output data relations. The FDO model is designed to predict future output features for deviated input vectors from any expected, feared of conceivable, future input for optimum control, forecast, or early-warning hazard evaluation. The linearized FDO provides fast analytical, input-output solution in matrix equation form. If the FDO is invertible, the necessary input for a desired output may be explicitly evaluated. A numerical example is presented for FDO model identification and hazard evaluation for methane inflow into the working face in an underground mine: First, a Physics-Based Operator (PBO) model to match monitored data. Second, FDO models are identified for matching the observed, short-term variations with time in the measured data of methane inflow, varying model parameters and simplifications following the parsimony concept of Occam's Razor. The numerical coefficients of the PBO and FDO models are found to differ by two to three orders of magnitude for methane release as a function of short-time barometric pressure variations. As being data-driven, the significantly different results from an FDO versus PBO model is either an indication of methane release processes poorly understood and modeled in PBO, missing some physics for the pressure spikes; or of problems in the monitored data fluctuations, erroneously sampled with time; or of false correlation. Either way, the FDO model is originated from the functionalized form of the monitored data, and its result is considered experimentally significant within the specified RMS error of model matching.

Keywords

Artificial Intelligence, Machine Learning, Matrix Operator, Numerical Functionalization, Convolution Model, Deconvolution, Occam's Razor

1. Introduction

The scope of the application is modeling time-dependent gas concentration and temperature variations in airways in buildings and subsurface openings based on data. A vast literature deals with processor systems model identification from data using classic methods. Reference to the basic work is in order to recognize the contribution of deFigueiredo, Dwyer, Eykhoff, Kalman, Volterra, Weiner and Zadek, among many others, in non-linear system identification [1] [2]. Machine Learning (ML) with Neural Network (NN) and other method applications are fast evolving for patterns and features evaluation from data, reviewed in [3]. Deep Learning (DL) can efficiently uncover targeted signal patterns using multi-layer, representation-learning methods composed of simple, non-linear modules, each transforming the representation from one level into a higher, more abstract level, whereas the features attributes are also recognized from the input data. Deep, convolutional neural networks (ConvNet) are showing promising results in multi-dimensional tasks, such as image processing and speech recognition. A common feature of ConvNet is the use of a kernel function at each representation layer as integrator of the data evaluation reviewed in [4] for pattern recognition in image analysis.

The kernel function may be viewed as a fundamental key to integrate the observed output at each network layer for object function evaluation. The kernel function is also related to using fundamental, indicial solution for predicting the outcome of a system with relatively simple response behavior but under the perturbation of complex boundary input influences. DL of such system is shown in [5], searching for a combination of indicial solutions and effectively realizing deep understanding of the observed phenomenon. The lesson learned from [5] and similar other applications such as [6] leads to distinguish DL from Deep Understanding (DU) methods. Specifically, the DL process only reaches an accurate but constructed mapping of the data to match an expected outcome; whereas the DU process interprets the outcome by the identification of the fundamental, indicial system response function or functions, related to a physical process model.

Kernel functions in convolutional process and system models have been long used before the bonanza of ML. The convolutional integral with a kernel function as an indicial solution is evolved from Duhamel's superposition integral [7] for time-dependent heat and mass transport processes in three dimensional domains. Furthermore, a treatise of solutions of partial differential equations for a variety of variable, initial and boundary conditions is provided for heat conduction problems in [7].

Input data for ML may be provided by measurements or from numerical model simulations. Dynamic process models are available for computing the advective, convective, and diffusive mass, as well as energy and momentum transport problems. Numerical simulation solutions in discretized form are available from CFD models [8], porous media codes [9], and network solvers

[10] [11]. CFD models are often used for generating ‘synthetic data’ for ML method tests, e.g., in [6].

A ML method for identifying kernel functions from simulated data from porous media code is shown in [12] for the specific goal of representing the complex, linear, or non-linear numerical solution with a convolutional, functional operator or a Volterra series of operators. The method called Numerical Transport Code Functionalization (NTCF), has been used in many numerical applications for its computational efficiency, allowing to import the solution of a numerical simulation into analytical-type matrix-vector equations.

The function-function called functional in matrix (tensor) operator form is suitable for model building from time-series data analysis. The convolution integral form of Duhamel’s superposition theorem provides the basis for the development of the operator linearized for an applicable domain. A deconvolution procedure is sought for the identification of the functional operator from monitored, uncontrolled data streams for both system input and response output vectors obtained over past time intervals.

The inspiration for mathematically “inverse-engineering” and extracting predictive, functional models of the measured system from monitored data comes from the success of the NTCF procedure. In the current study, the input data is assumed to be measured as an un-controlled time series, different from the tasks described in [12], where controlled, input signals are used for generating system responses for NTCF system identification.

The goal of the study is to extract the indicial, kernel functions solutions to the time-honored governing equations of transport processes employed in [9] [10] [11]. Mathematical models for the simulation of time-dependent output process parameters such as gas concentration and temperature are described in [13] for a given geometry, site properties, initial conditions, and the input variation of boundary conditions. The governing mathematical models have rather simple forms in partial differential equations. Furthermore, similarity exists between the governing equations and their solutions in the time-dependent heat and mass transport processes, that is, in the partial differential equation known as Fourier’s and Fick’s second laws, written for solids in the form as follows:

$$\frac{\partial f}{\partial t} = D \cdot \Delta f \quad (1)$$

where f may be temperature or species concentration; and t , D and Δ are time, diffusivity and the Laplace operator, respectively. For moving fluids, the partial derivative on the left side of (1) is replaced by the substantial derivative as

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f,$$

where \mathbf{v} and ∇ are the advective velocity vector and the del (nabla) operator, respectively:

$$\frac{\partial f}{\partial t} = -\mathbf{v} \cdot \nabla f + D \cdot \Delta f \quad (2)$$

The challenge is to find algorithms for extracting the indicial solutions to expected transport models with governing equations of type (1) and (2) with unknown properties from data outputs in response to the inputs, both time-dependent, obtained from real-time measurements. The abstracted, output operator model and its kernel function(s) from the NTCF processor in [12] characterizes the hidden, indicial solutions of the partial differential equations of the governing, transport process equations. By analogy, it is expected that model building by an NTCF-like processor from monitored input and output data will characterize the response of the physical, measured system in the present study.

The heart of the solution, therefore, is to find the input-output model relations in the form of indicial kernel function(s) in the convolution integral, that may be expressed in a closed, matrix-vector equation form. The desired form of the FDO model for a linearized, interval-averaged time domain may be a single, matrix operator equation, the focus of the present study. Generalization of the FDO for strongly nonlinear systems may follow the method of Volterra series expansion [12] leading to a set of FDO equations, a task left for future work.

2. A Relativistic, Input-Output Model

The convolution integral expresses a “relativistic” relationship between a time-variable input function, $x(t)$, and a system-response output function, $y(t)$, in which values at time t can only be a function of past values of input $x(\tau)$, $\tau < t$, as follows:

$$y(t) = \int_{\tau=0}^t \frac{\partial}{\partial t} x(\tau) A(t-\tau) d\tau + y(0) \quad (3)$$

Equation (3) expresses Duhamel’s superposition theorem in an integral equation form. It assumes linearization within the model domain, where $A(t)$ is the unknown, indicial kernel function with a physical meaning of being a system-response output function to a unit step-change input function to a stationary transport system. The goal is to construct a system model from (3) between a known, given input function, $x(t)$, and a sampled output response function, $\hat{y}(t)$, in such a way that $y(t)$ is a best approximation of $\hat{y}(t)$ in a Least Square (LSQ) sense. The initial, constant output value may be assumed zero for simplicity in notations, $y(0) = 0$.

Note that most NN model’s training processes are un-relativistic, as all input samples for a closed training time window are commonly used to correlate with all output samples, allowing to use all input information at $x(\tau)$ [$0 \leq \tau \leq t$] with response outcome $y(\tau)$ [$0 \leq \tau \leq t$] for evaluating model constants. Such evaluation would result in a functional operator model represented by a full matrix, such as used in [13]. In contrast, the present goal is to find a relativistic, physically sound model in the form of a lower-triangular matrix operator, similarly to the model obtained from the NTCF processor [12].

The goal is to determine $A(t)$ of the unknown transport system from the known, $x(t)$, and $\hat{y}(t)$ functions, sampled over a sufficiently long time pe-

riod. Once $A(t)$ is known, often regarded as the “essential solution” to the system in the process identification literature, the solution for any input boundary condition function, $x(t)$, may be obtained from the integral in (3).

Input and output vectors in (3) are defined for the representation of $x(t)$ and $y(t)$ by sampled values taken at representative time divisions $t_1, t_2, \dots, t_i, \dots, t_N$, not necessarily agreeing with the measurement time periods. The temporal discretization of t and the selection of N samples defines an N -dimensional space in which the $x \in R^N$ and $y \in R^N$ vectors may be related by an operator model, $M \in R^N \times R^N$. Sampling with unequal time periods of an evenly spaced time series data is advantageous for data compression. Quantum vector representation of sampled, monitored data with a refining time discretization from past to present from an even-sampled, real-time data flow is described in [13], in which old, history data are sparsely represented, $\Delta t \ll t_1$, while the most recent reading reaches the finest sampling time interval, Δt , as $\Delta t = t_{N-1} - t_N$. With quantum (Q) vector representation, the number of data samples are compressed and reduced. However, even unaveraged, unfiltered data in vectorized form taken at t_i time intervals is advantageous as the vector representation of the values of $x(t_i)$ and $y(t_i)$ taken at t_1, \dots, t_N instants allows for parallel processing in which each element is best fit between model and observation. For simplicity of notations, the vector-processed form of the sampled, monitored data with a refining time discretization from past to present from an even-sampled, real-time data flow is referred to as a data Q-vector.

For simplicity of presentation of operator model development between vectorized input and output variables, the quantum or unfiltered data Q-vector components are assumed to be determined from the measured data at time $t_1, t_2, \dots, t_i, \dots, t_N, \dots$ instants of the continuous time at irregular but monotonously variable sampling time intervals within the sliding time window between t_1 and t_N . Furthermore, it is stipulated that the $t_1, t_2, \dots, t_i, \dots, t_N$ time base is unchanged during resampling of the measured data for a sliding window with $\Delta t = t_{N-1} - t_N$ as the finest sampling interval.

It is advantageous to interpolate $x(t)$ and $y(t)$ with piecewise-continuous function between base points with constant $m_i = \partial x / \partial t$ in each $x \in [x_i, x_{i-1}]$ interval, where:

$$m_i = \frac{x_i - x_{i-1}}{t_i - t_{i-1}} \tag{4}$$

Applying the integral in (3) for $y(t_i) (i = 1, \dots, N)$ as the sum of the cumulative time interval components; factoring out $m_i (i = 1, \dots, N)$ from each component; and applying a variable transformation $\lambda = t_i - \tau$ yields:

$$y(t_1) = m_1 \int_{t_0}^{t_1} A(\lambda) d\lambda \tag{5a}$$

$$y(t_2) = m_1 \int_{t_2-t_1}^{t_2} A(\lambda) d\lambda + m_2 \int_0^{t_2-t_1} A(\lambda) d\lambda \tag{5b}$$

$$y(t_i) = m_1 \int_{t_{i-1}}^{t_i} A(\lambda) d\lambda + m_2 \int_{t_i-t_2}^{t_i-t_1} A(\lambda) d\lambda + \dots + m_i \int_0^{t_i-t_{i-1}} A(\lambda) d\lambda \tag{5c}$$

$$y(t_N) = m_1 \int_{t_N-t_1}^{t_N} A(\lambda) d\lambda + \dots + m_i \int_{t_N-t_i}^{t_N} A(\lambda) d\lambda + \dots + m_N \int_0^{t_N-t_{N-1}} A(\lambda) d\lambda \quad (5d)$$

Substituting m_i from (4) in (5a) and (5b) and rearranging gives:

$$y(t_1) = x_1 \frac{1}{t_1 - t_0} \int_0^{t_1} A(\lambda) d\lambda \quad (6a)$$

$$y(t_2) = x_1 \left[\frac{1}{t_1 - t_0} \int_{t_2-t_1}^{t_2} A(\lambda) d\lambda - \frac{1}{t_2 - t_1} \int_0^{t_2-t_1} A(\lambda) d\lambda \right] + x_2 \frac{1}{t_2 - t_1} \int_0^{t_2-t_1} A(\lambda) d\lambda \quad (6b)$$

Since the integral mean values of $A(t)$ appear in (6a) and (6b), a new notation is in order to simplify the equations. With the introduction of row index i and column index j for matrix notations, an integral mean value expression is recognized from observing (6a) and (6b):

$$M_{i,j}^A = \frac{1}{t_j - t_{j-1}} \int_{t_i-t_{j-1}}^{t_i-t_j} A(\lambda) d\lambda \quad (7)$$

According to (7), the averaging time intervals in the $M_{i,j}^A$ integral mean values does not coincide with the time division intervals, unless the time division is equidistant. The simplified notation of (7) yields the simplified forms of the following convolutional model equations:

$$y(t_1) = x_1 M_{1,1}^A \quad (8a)$$

$$y(t_2) = x_1 [M_{2,1}^A - M_{2,2}^A] + x_2 M_{2,2}^A \quad (8b)$$

$$y(t_i) = x_1 [M_{i,1}^A - M_{i,2}^A] + x_2 [M_{i,2}^A - M_{i,3}^A] + \dots + x_i [M_{i,i}^A] \quad (8c)$$

$$y(t_N) = x_1 [M_{N,1}^A - M_{N,2}^A] + \dots + x_{j-1} [M_{N,j-1}^A - M_{N,j}^A] + \dots + x_N [M_{N,N}^A] \quad (8d)$$

Equations (8a)-(8d) may be written in matrix-vector form:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = M \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, \quad (9)$$

where:

$$M = \begin{bmatrix} M_{1,1}^A & 0 & \dots & 0 \\ M_{2,1}^A - M_{2,2}^A & M_{2,2}^A & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ M_{N,1}^A - M_{N,2}^A & M_{N,2}^A - M_{N,3}^A & \dots & M_{N,N}^A \end{bmatrix}. \quad (10)$$

The M matrix operator defined in (10) is a function of the time division vector only, $[x_1, x_2, \dots, x_N]$, and the $A(t)$ indicial function. Matrix M has zeros above the main diagonal, assuring that only earlier input may influence a later output. Therefore, the model satisfies common sense and consistent with the theory of relativity, helpful for matching data from a physical system with a mathematical model.

The relationship between M and $A(t)$ is shown in the structure of (10) as M can be uniquely obtained using (7) with a given $A(t)$ indicial function using, for

example, a piecewise linear interpolation in $x(t)$, for a given time division vector. For any given $A(t)$ indicial function, matrix M can be computed from (7), to be used in (9) for obtaining an output vector $[y(t_i)]$ as the system's response to an arbitrary input vector $[x(t_i)]$.

The $A(t)$ indicial function may be obtained analytically for simple geometry and linear system's governing equations such as in (1) and (2). However, for complex, anisotropic geometry, inhomogeneous and nonlinear systems, only computational, discretized, numerical models may be used for obtaining an essential, step-change response solution for approximating a locally linearized $A(t)$ indicial function for constructing a fast-predicting, analytical-type matrix model for repeated use. Such a transformational, NTCF procedure [12] requires controlled boundary conditions specifications for the numerical transport code runs to obtain a matrix functional. Such controlled, boundary conditions are not possible to specify from a stream of real time, monitored data. To find $A(t)$ and M from monitored data, a deconvolution procedure is needed.

3. Deconvolution Functionalization of Data for a Relativistic, Input-Output Model

For deconvolution of data for the unknown $A(t)$ and M , consider first the invertibility problem for $A(t)$ from a given M extracted from the measured values of $x(t)$, and $y(t)$. It is straightforward to find N number of $A(t)$ directly from (9) and the definition of $A(t)$ as a system response solution for a unit step change in x :

$$\begin{bmatrix} A(t_1) \\ A(t_2) \\ \vdots \\ A(t_N) \end{bmatrix} = M \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (11)$$

However, far more sampled values of A are needed to evaluate its integral mean values in (7) for the components of the M matrix in (10). Therefore, while it is assuring to know that there exists an $A(t)$ behind M , an explicit expression or representation of $A(t)$ is not practical nor necessary. A system model may be more convenient in a matrix operator form, as M is directly usable as a linear operator for analytical, algebraic solutions. Therefore, the desired form of the relativistic, input-output model is given in (9) with components defined in (7) and (10). The deconvolution procedure in the inverse solution of the convolution integral may be considered completed by finding the M matrix.

3.1. Determination of M from Observed, Sampled Values of \hat{x} and \hat{y}

The $M_{i,j}$ elements of matrix M are to be determined from (9). It is convenient to evaluate each row of M separately. The solvability of M from N input samples relates to the size of the x and y . Rearrangement of (9) for elements y_n in each

row, $n \in [1, N]$, yields:

$$y_n = [x_1 \ x_2 \ \cdots \ x_n] \begin{bmatrix} M_{n,1} \\ M_{n,2} \\ \vdots \\ M_{n,n} \end{bmatrix} \quad (11)$$

Since n number of $M_{n,i}$, $i \in [1, n]$ are unknown in (11), then a minimum of n equations are needed to solve for $M_{n,i}$. Therefore, repeated acquisition is needed for each $n \in [1, N]$ element of the Q-vector values of y , taken at shifted time intervals by step δ , $\{y_n^{0\delta}, y_n^{1\delta}, \dots, y_n^{(n-1)\delta}\}$. Likewise, repeated n element Q-vectors need to be processed, taken at shifted time intervals by step δ , such as $\{[x_1 \ x_2 \ \cdots \ x_n]^{0\delta}, [x_1 \ x_2 \ \cdots \ x_n]^{1\delta}, \dots, [x_1 \ x_2 \ \cdots \ x_n]^{(n-1)\delta}\}$.

The shifted time intervals in the sliding windows are conveniently processed using the finest $\Delta t = t_{N-1} - t_N$ time steps. Considering noisy data, there exist a large-enough number of k samples ($k \geq n$) to be sufficient for the solution of the set of equations:

$$Y_n^k = X_n^k \begin{bmatrix} M_{n,1} \\ M_{n,2} \\ \vdots \\ M_{n,n} \end{bmatrix} \quad (12)$$

where

$$Y_n^k = \begin{bmatrix} y_n^{0\delta} \\ y_n^{1\delta} \\ \vdots \\ y_n^{(k-1)\delta} \end{bmatrix} \quad \text{and} \quad X_n^k = \begin{bmatrix} [x_1 \ x_2 \ \cdots \ x_n]^{0\delta} \\ [x_1 \ x_2 \ \cdots \ x_n]^{1\delta} \\ \vdots \\ [x_1 \ x_2 \ \cdots \ x_n]^{(k-1)\delta} \end{bmatrix} \quad (13)$$

The predicted vector Y_n^k on the left side of (12) must be matched with the corresponding, measured Q data vector, $\hat{Y}_n^k = \{\hat{y}_n^{0\delta}, \hat{y}_n^{1\delta}, \dots, \hat{y}_n^{(k-1)\delta}\}$, for the measured values where $n \leq k$, in an LSQ fit sense:

$$\sum_{i=0}^{k-1} (\hat{y}_n^{i\delta} - y_n^{i\delta})^2 = \min \quad (14)$$

It is straightforward to use measured input data vectors also for predicting output vector from the operator model, substituting Q-vectors as the elements of \hat{X}_n^k in the right side of (12) as $\{[\hat{x}_1 \ \hat{x}_2 \ \cdots \ \hat{x}_n]^{0\delta}, [\hat{x}_1 \ \hat{x}_2 \ \cdots \ \hat{x}_n]^{1\delta}, \dots, [\hat{x}_1 \ \hat{x}_2 \ \cdots \ \hat{x}_n]^{(n-1)\delta}\}$. It can be shown that substitution of (13) into (14), using the measured input data and taking the partial derivatives of the equation with respect to $M_{n,i}$, $i = 1, \dots, n$, leads to a matrix equation as follows:

$$(X_n^k)^T \hat{Y}_n^k = (\hat{X}_n^k)^T \hat{X}_n^k \begin{bmatrix} M_{n,1} \\ M_{n,2} \\ \vdots \\ M_{n,n} \end{bmatrix}, \quad (15)$$

where $(\hat{X}_n^k)^T$ denotes the transpose of matrix \hat{X}_n^k containing measured values. The unknown elements of $M_{n,j}$, $j < n$ and $1 \leq n \leq N$ is expressed from (15) explicitly by formally left-multiplying it with the inverse of $(\hat{X}_n^k)^T \hat{X}_n^k$ i.e., $\left[(\hat{X}_n^k)^T \hat{X}_n^k \right]^{-1}$:

$$\begin{bmatrix} M_{n,1} \\ M_{n,2} \\ \vdots \\ M_{n,n} \end{bmatrix} = \left[(\hat{X}_n^k)^T \hat{X}_n^k \right]^{-1} (\hat{X}_n^k)^T \hat{X}_n^k \hat{Y}_n^k \tag{16}$$

If $\left[(\hat{X}_n^k)^T \hat{X}_n^k \right]^{-1}$ ($1 \leq n \leq N$) is computable, matrix elements $M_{n,i}$,

($1 \leq i \leq n$) can be evaluated for each row independently for a pre-selected N quantum vector and model size. The solvability of (16) depends on the condition of $(\hat{X}_n^k)^T \hat{X}_n^k$, related to the quality and the size of the input data vector as well as to the selected time base vector and the construction of the quantum vector of X_n^k from sampled measurement data. If matrix $(\hat{X}_n^k)^T \hat{X}_n^k$ is singular, (16) does not provide a solution; if $(\hat{X}_n^k)^T \hat{X}_n^k$ is rank-deficient or badly conditioned, the solution will not satisfy (14). The rank of a matrix is the number of linearly independent rows and columns, computed as the number of singular values that are larger than a tolerance. The tolerance for matrix A calculated as the $\max[\text{size}(A)] * \text{eps}[\text{norm}(A)]$, where eps for double precision calculation is 2^{-52} .

Increasing the number of data samples, k , relative to the size of the model, N is one remedy for reaching a meaningful solution to (16). However, for a given input data set, the solvability of (16) is an issue, affecting the existence of the operator model.

3.2. Definition of the Condition of Solvability of M from the Input Data

Let matrix \hat{X}_N^k be formed from consecutively sampled data Q-vectors by equal-distance steps $i\delta$, $i \in [0, k-1]$ arranged backward over a sliding time window of $k\Delta t$, where $\hat{x}_N^{(k-1)\delta}$ is the last sampled, most recent Q-vector data element, $\hat{x}(t_N)$; and $\hat{x}_1^{0\delta}$ is the earliest Q-vector element in the first sampled sliding window:

$$\hat{X}_N^k = \begin{bmatrix} [\hat{x}_1 \ \hat{x}_2 \ \dots \ \hat{x}_N]^{0\delta} \\ [\hat{x}_1 \ \hat{x}_2 \ \dots \ \hat{x}_N]^{1\delta} \\ \vdots \\ [\hat{x}_1 \ \hat{x}_2 \ \dots \ \hat{x}_N]^{(k-1)\delta} \end{bmatrix} \tag{17}$$

The condition of solvability for M is defined by the existence of the inverse matrix, $iX = \left[(\hat{X}_n^k)^T \hat{X}_n^k \right]^{-1}$. If matrix iX is well conditioned with all real values under and in the main diagonal (and zeros in the upper triangular), the model

identification task for M is well conditioned and solvable.

3.3. Occam's Razor

The solvability of (16) depends on the condition of matrix $\left[\left(\hat{X}_n^k \right)^T \hat{X}_n^k \right]^{-1}$. The maximum value of N may be limited, depending on the maximum number of available sample points, k , of \hat{X}_N^k and \hat{Y}_N^k . The maximum value of N and oversampling number k may be determined by trial-and-error, calculating the rank of $\left[\left(\hat{X}_n^k \right)^T \hat{X}_n^k \right]^{-1}$ before model fitting, that is, the number of linearly independent rows or columns in the input data and keeping it sufficiently high for a given error of computation. The maximum number of oversamples, k , is limited by the number of measurements on $\hat{x}(t_i)$ and $\hat{y}(t_i)$ taken over time for establishing their quantum vectors for each sliding time window of t_1^k, \dots, t_N^k . The smaller the model size N , the fewer the necessary oversamples, k , for reaching a unique solution to Equation (16).

Reducing size N of the matrix operator in an AI-based model relates to its complexity for grasping the underlying processes which control the outputs for given inputs. The concept of Occam's Razor [14] is the principle of parsimony with the minimization of N , a virtue worth considering for improving the model's quality and providing solvability. The subjectivity of such a decision may hereby be replaced by the condition of solvability of (16) with a maximum value of N for the maximum number of oversamples k , available or desirable for model identification.

Reducing the complexity of the model without reducing N for gaining better predicting power also calls for Occam's Razor. The model's size, N , affects the grasp of the history effects and the delay mechanism of substance transport between cause and response. Once the FDO model is identified, simplification may be done by eliminating the history effect not from the model, but from the input data, equating the past variations with the the most recent, measured value. This means a step-change input (as if the same input has persisted from time zero) for finding the response function from the FDO model at every fine time steps. Such a model simplification is executed by replacing M with a diagonal matrix, M_D as follows:

$$M_D(j, j) = \begin{cases} \sum_{i=1}^j M(j, i), & j \in [1, N], i \in [1, j]. \\ 0 & i \neq j \end{cases} \quad (18)$$

3.4. Hypothesis for the Learnability of an Operator Model of a Physical Process from Uncontrolled Input-Output Data Streams

For given sets of sampled Q-vector input

$$\left\{ \left[\hat{x}_1 \hat{x}_2 \dots \hat{x}_N \right]^{0\delta}, \left[\hat{x}_1 \hat{x}_2 \dots \hat{x}_N \right]^{1\delta}, \dots, \left[\hat{x}_1 \hat{x}_2 \dots \hat{x}_N \right]^{k\delta} \right\} \text{ and Q-vector output}$$

$\{[\hat{y}_1 \ \hat{y}_2 \ \dots \ \hat{y}_N]^{0\delta}, [\hat{y}_1 \ \hat{y}_2 \ \dots \ \hat{y}_N]^{1\delta}, \dots, [\hat{y}_1 \ \hat{y}_2 \ \dots \ \hat{y}_N]^{k\delta}\}$ obtained from a data streams over a monitored time interval T with available maximum number of sampled values of $nT = T/\Delta t$, there exists a best combination of $N < nT$ and $k < nT$ which provides an M with a minimum error of fit in an LSQ metric for modeling the physical process for the output data.

The hypothesis in 3.4 is intuitive, supported by experimentation with real word data, FDO model identification, model fit tests to input data, and model prediction power measured from resilience against input variations in different shapes from those used in model training.

The hypothesis supports incentives to experimentations with Occam's Razor to cut the size of M to the smallest but still powerful for predictions and low in RMS fitting error.

4. Application Example of FDO Model for Methane Inflow and Concentration Variation under Barometric Pressure Changes

4.1. Physics-Based Model Test for Pressure-Driven Methane Inflow Prediction

Monitored data for atmospheric conditions from operating mine for 327 days under normal operating conditions are taken from a previous study [13] for the subject DU learning analysis. The monitored parameters are air flow rate in the face drift (Qa), incoming methane (CH_4) gas concentration at the main gate (c_{MG}), exiting Methane concentration at the tail gate (c_{TG}), and barometric pressure Pb , all sampled at 5-minute time intervals.

The nature of the methane mass influx due to pressure-driven, diffusive Darcy's flow with time was studied using the NTCF model technique [12] in a previous report, to be published separately from the current work. For physics-based, methane flux calculation, the following matrix equation is used:

$$Q_M = M_M \cdot A \cdot (P_{st} - \hat{P}) + Q_0, \quad (19)$$

where $Q_M, M_M, A, P_{st} - P$, and Q_0 are respectively the methane mass flux Q-vector, [kg/s]; the admittance matrix, [kg/s/m²/Pa]; the surface area, [m²]; the pressure driving force Q-vector [Pa] from monitored barometric pressure data at the working face; and the initial methane emission mass flux Q-vector, Q_0 [kg/s], assumed to be kept at zero. The $M_M \cdot A$ term may be considered a calibrated, but an *a priori* PBO in (19).

The Q-vectors in (19) may be formed from a time-series string of data according to the concept introduced in [13]. For simplicity in presentation, these are all assumed to be simple, re-sampled quantities taken at monotonously variable, uneven time intervals, $\tau_i, i \in [1, N]$ of the time-variable values (either measured or calculated) over a sliding time window in N steps, introduced in [13]. For example, a Q-vector sampling curve is shown in **Figure 1** (left side) for $N = 35$ for a full time span of 327 days sampled monotonously, refining from

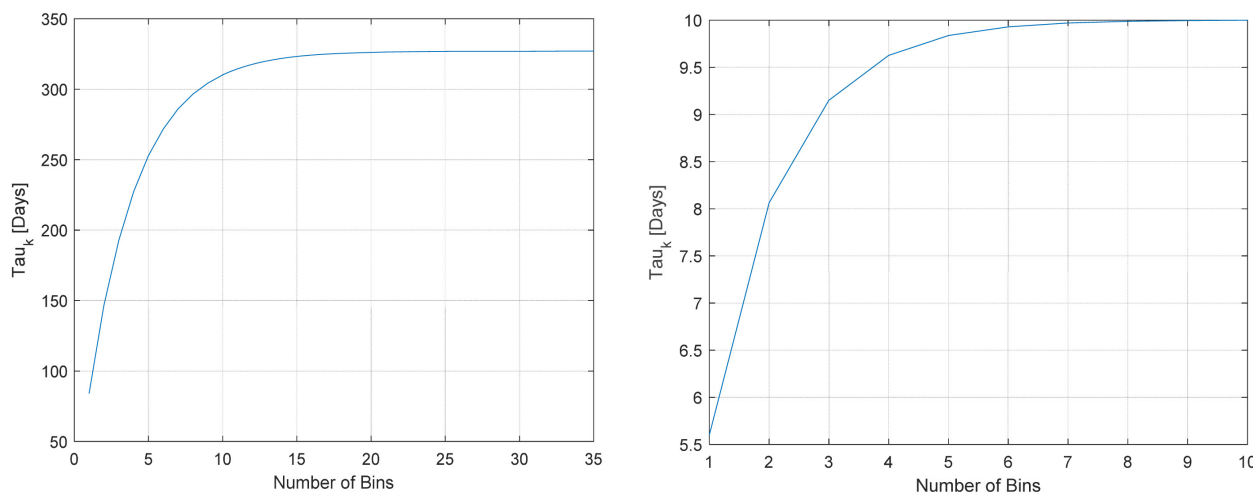


Figure 1. Q-vectors τ time divisions for 327 days with $N=35$ (left); and for 10 days with $N=10$ (right).

coarse to the finest, with the last step of $\Delta t = (\tau_N - \tau_{N-1}) = 1$. **Figure 1** (right side) shows a sampling curve for a reduced model size and sampling number of $N=10$, and also a reduced time span of 10 days for the sweeping, sliding time window, with the finest time step kept at $\Delta t = 1$.

In the $N=35$ example, each step is 5 minute (300 s) long, progressing over the real time scale between the beginning of day 1 to day 327 in 5 minute intervals, giving 94,176 steps for the calculation time domain, irrespective of the model and Q-vector size, N . While the time steps are equal, the Q-vectors in (19) are sampled elements, carrying the full history from coarse to fine steps for the full time period of 327 days if desired.

The M_M methane transport admittance matrix for the mine's working face is calculated from an analytical-computational model, based on the indicial admittance function and its structure shown in (7)-(10). The resulting, lower-diagonal M_M matrix is plotted in **Figure 2** in two views for the positive and negative domains. As shown, the close-to diagonal elements under the main diagonal dominate over the far off-diagonal elements.

For an assumed $P_{st} = 4 \times 10^5$ [Pa] strata gas pressure and the measured barometric pressure, \hat{P} , the methane-driving pressure difference variation is given in **Figure 3** for all time steps (left); and for an enlarged range (right). The methane flux Q-vector, Q_M in [kg/s], is calculated from (19) using a calibrated, methane transferring surface area, A , in [m²], which provides an optimum match for the observed \hat{Q}_M , in [kg/s] from monitored concentration data over the entire 327 days:

$$\sum_j^{nT} \left(\sum_i^N \left[A \cdot M_M \cdot (P_{st} - P(i, j)) - \hat{Q}_M(i, j) \right]^2 \right) = \min \quad (20)$$

In (20), the observed methane flux Q-vector, \hat{Q}_M , is calculated from monitored data as follows:

$$\hat{Q}_M = \hat{Q}a \cdot (\hat{c}^{TG} - \hat{c}^{MG}), \quad (21)$$

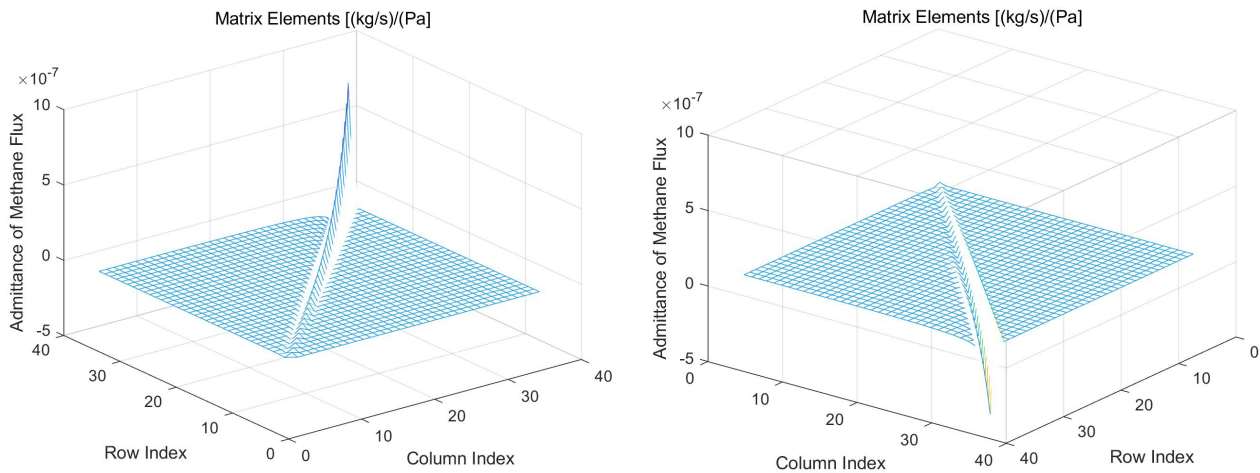


Figure 2. M_M matrix for pressure-driven methane transport in the positive domain (left) and negative domain (right).

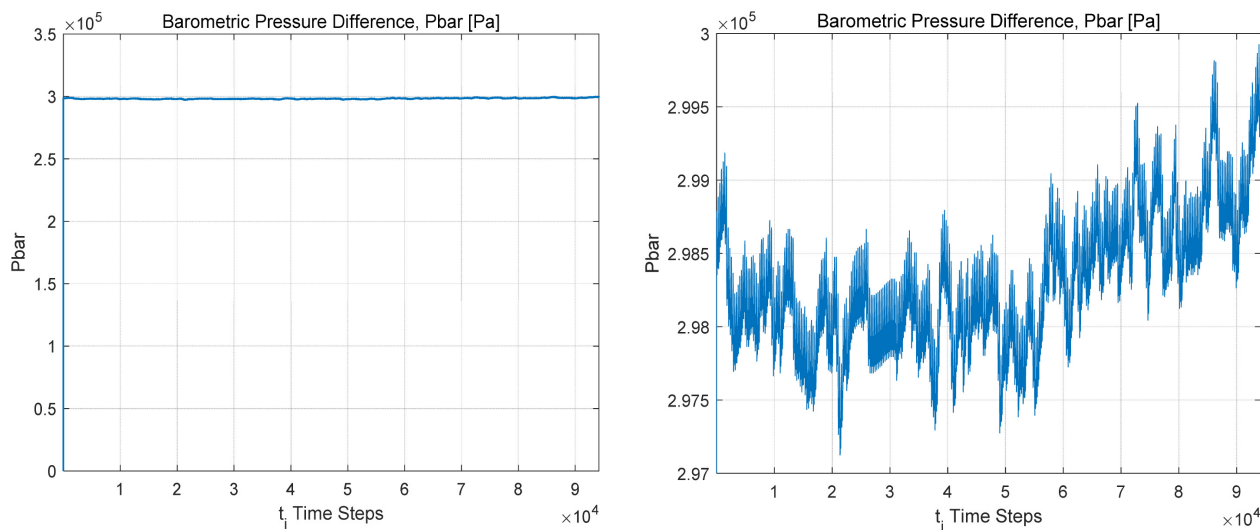


Figure 3. The methane-driving pressure variation $P_{st} - P$ over 327 days (94,176 steps of 5-min length, left); and zoom-in (right).

where \hat{Q}_a , \hat{c}^{TG} , and \hat{c}^{MG} are the monitored air flow rate, [kg/s]; monitored concentration at the tail gate [kg/kg]; and monitored concentration at the main gate [kg/kg], all repeatedly sampled into Q-vector forms in size of $N \times nT$.

The last (newest) elements of observed methane flux Q-vector, $\hat{Q}_M(N, j)$, $j \in [1, nT]$ are plotted in **Figure 4**, together with the last element of the predicted $Q_M(N, j)$ values from (19) for comparison. As shown, the calibrated, analytical model can approximate well the measured \hat{Q}_M trend only in an averaged sense over 327 days, but poorly on the fine time scale, being insensitive in \hat{Q}_M to the fine, low-amplitude pressure variations with time.

4.2. FDO Model Test for Pressure-Driven Methane Inflow Prediction

As the physics-based, methane flux model cannot explain the measured inflow fluctuations with time, the FDO model identification method is used for directly

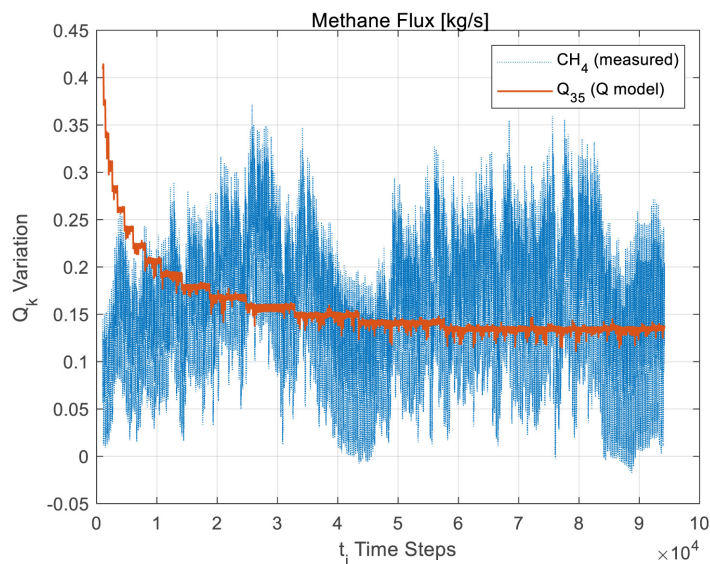


Figure 4. Measured and modeled methane flux Q-vectors: measured, $\hat{Q}_{35}(35, j)$; and calibrated, analytical PBO model, $Q_{35}(35, j)$.

obtaining matrix operator M_{MA} from observed data to satisfy observation:

$$\hat{Q}_M = M_{MA} \cdot (\hat{P}_{\max} - \hat{P}), \quad (22)$$

where \hat{Q}_M , M_{MA} , and $\hat{P}_{\max} - \hat{P}$, and Q_{q_0} are respectively the methane mass flux Q-vector, [kg/s]; the admittance matrix, [kg/s/Pa]; and the pressure driving force Q-vector [Pa] from the fine-scale pressure variation in the air flow at the face.

Note that the large pressure difference between the hydrostatic pressure of the strata and the air, $P_{st} - \hat{P}$, in (19) is greatly reduced to the scale of the fine variations withing the atmospheric pressure, assuming as the short time driving force for methane release. The pressure driving force used in (22) is processed from the difference between the minimum, observed barometric pressure readings over the 327 days, minus the time-variable, monitored and sampled pressure values, \hat{P} . **Figure 5** shows the last, most recent elements of the Q-vectors, $\hat{P}_{\max} - \hat{P}$ over all time steps (left); and in an enlarged range of 4 days (right).

For finding M_{MA} in (22), the deconvolution procedure described in the foregoing with (11)-(16) is used first, embracing the entire time period of 327 days. $N = 35$ is selected for the model and Q-vectors size, a single “sliding window” size of 327 days, and $k = 300 \times 288 = 86400$ samples at 5-minute time steps for repetition in the LSQ solution.

The model identification procedure converges in a few seconds, resulting in an M_{MA} operator, plotted in **Figure 6** in two views for the positive and negative domains. As shown, M_{MA} is a lower triangular matrix with variable elements under the main diagonal, dominating in the lower-right half, representing the newest data, over the upper-left half, representing the past, history data in the input and output Q-vectors.

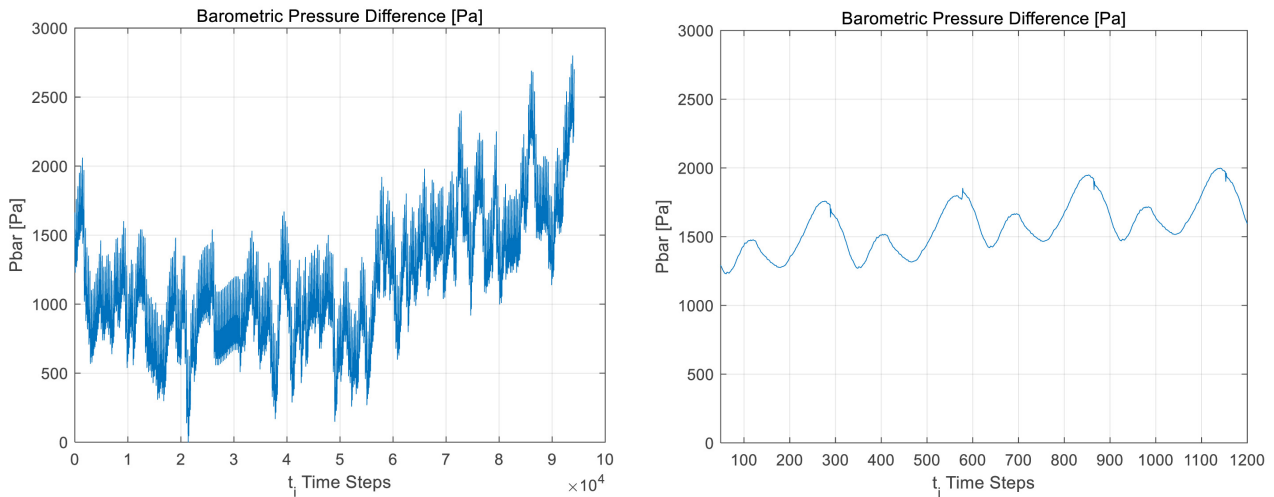


Figure 5. The methane-driving pressure variation $P_{\max} - P$ over 327 days (94,176 steps of 5-min length, left); and zoom-in (right).

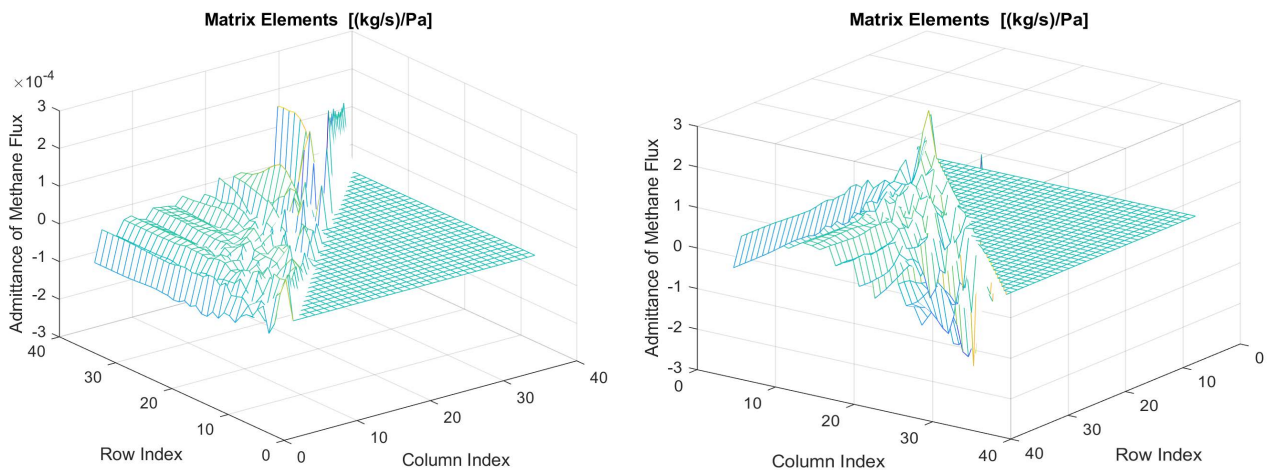


Figure 6. M_{MA} matrix for pressure-driven methane transport in the positive domain (left) and negative domain (right).

The shape of M_{MA} in **Figure 6** is unlike that of M_M in **Figure 2**, revealing fundamental differences in the physical processes, assumptions in the methane-driving pressure difference force, and the foundations of the operator model, moving from speculative to be data-driven. The experimental foundation of M_{MA} lends its priority over M_M that comes from simplified assumptions. The high value of A necessary for matching calibration in **Figure 4** indicates that the measured methane flux variation may respond to fast pressure variation from a shallower strata layer of much larger surface area than that for the large-amplitude and slower pressure change during coal extraction from the long wall.

The performance of M_{MA} can be checked using the monitored $\hat{P}_{\max} - \hat{P}$ data, and predicting Q_{35} from (22), and comparing it with the measured data, $\hat{Q}_{M,35}$. The results are in **Figure 7**, showing that the FDO model covers well the variation of the monitored methane flux data. The RMS error of match, divided

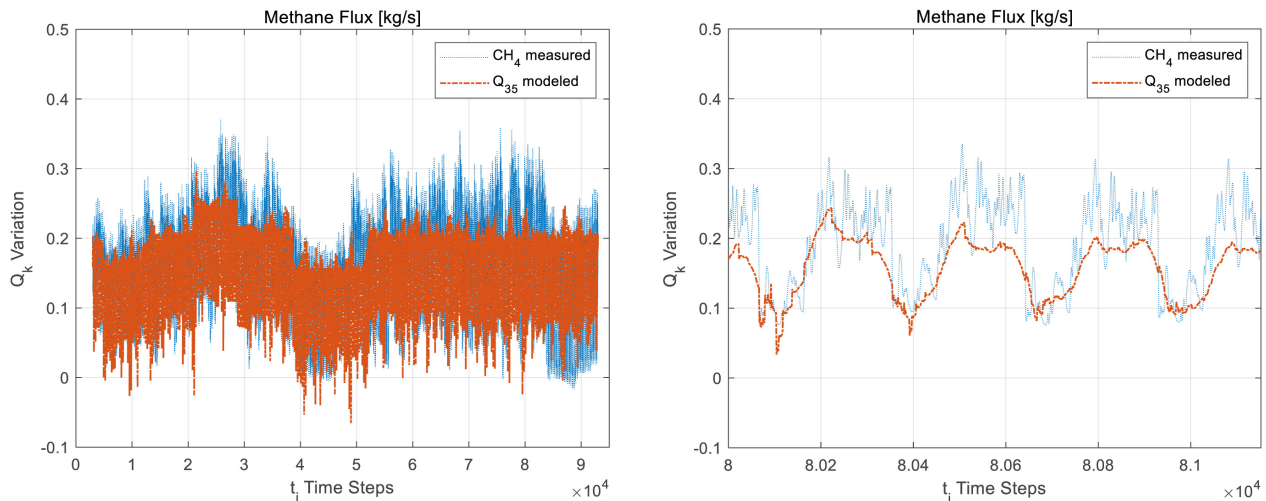


Figure 7. Measured and modeled methane flux variation over 327 days: $\hat{Q}_{35}(35, t_i)$, and $Q_{35}(35, t_i)$ from FDO(35) model (left); enlargement for four days (right).

by the arithmetic average of the methane flux is 0.29, that is, 29% for the entire 327 days with a single FDO model, comprising of a lower-triangular, 35×35 matrix with only 630 non-zero elements.

The error of fit of the FDO model and its optimization procedure are characterized by its probability density function and histogram, shown in **Figure 8**. The histogram matches very closely the normal distribution graph, proving the success of the LSQ procedure for random, white-noise error minimization.

4.3. Simplified FDO Models for Real-Time Analysis and Forecasting

Although a single FDO model may be obtained in seconds from a past, monitored data over 327 days, the acquisition time interval for the data series is long and exhausting. For real-time data analysis and forecasting, the concept of Occam's Razor is used to reduce the size of N as well as to revert to a simplified model according to (18), converting the full matrix to diagonal elements for eliminating model complexity. Experimentation is necessary to find the best combination of N (size of operator and Q-vectors), nT (history time span, that is, the sliding time window with N divisions), and k (number of repetitions of the Q-vectors at Δt successions). According to Hypothesis 3.4, there should be a best combination for each task, however difficult it may be to find it.

4.3.1. The First Occam's Razor Example for an FDO Model

Is given with a reduced-size selection of $N = 10$, $nT = 10$ days (2880 time steps); $k = 10$ days (2880 time steps) for Q-vectors processing; and FDO model identification for each 5-minute, real time instant over the allowable time period that stretches from day 10 (step 2880) to day 316 (approx. step 91,000). Each of the approx. 87,000 model is identified and may be used for analyzing of the sub-system around the subject time instant or forecasting its behavior in the

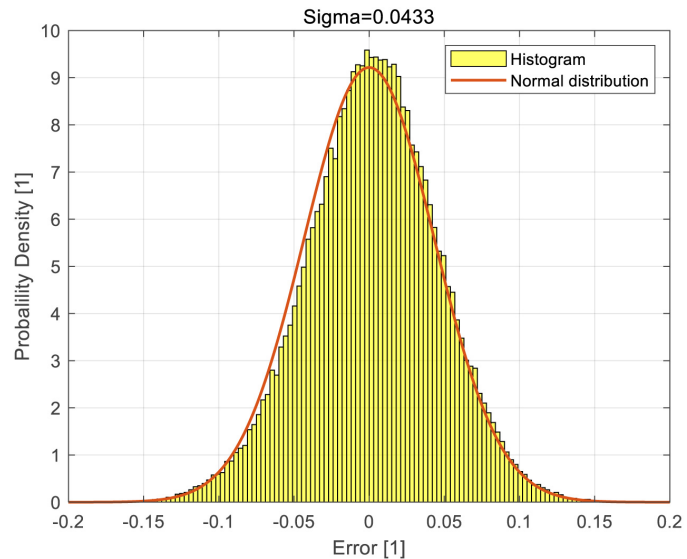


Figure 8. Probability density function and histogram of the error of fit of FDO model ($N = 35$, $RMS = 0.0433$).

immediate future before the next data and sub-model is already available. Each 10×10 lower triangle, sparse matrix is stored in a compressed form arranging the non-zero elements row-continuously into a single, 55-element column vector for further processing. **Figure 9** shows all such vectors plotted along the time steps showing high variations between the consecutive models.

The measured and the modeled methane flux variation are shown in **Figure 10** for over 300 days (left), and for a four-day time period as an enlargement (right). **Figure 10** depicts three overlaid graphs such as: 1) the measured Q-vector's last element, $\hat{Q}_{10}(10, t_i)$; 2) the last element of $Q_{10}(10, t_i)$ from the full FDO model using prediction model (22); and 3) the last element of $Q_{10}(10, t_i)$ from the simplified, diagonal FDO model using diagonalization (18) and prediction model (22). As shown, graph (1) of the measured methane flux is followed well with graphs (2) and (3) from the two models. As expected, graph (2) from the full FDO model shows a closer match that graph (3) from the diagonal FDO model to graph (1). Graph (3) is of a simplified curve, reacting faster, yet less sensitively to sudden jumps in its driving pressure difference input, as it will be shown in following applications.

Figure 11 shows the probability density function and histogram of error of fit of the FDO model. The match is close to a normal distribution graph, but not as good as that found in **Figure 8** for the full model.

4.3.2. The Second Occam's Razor Example for an FDO Model

Is given with a reduced-size selection of $N = 15$, $nT = 15$ days (4320 time steps); $k = 50$ days for Q-vectors processing; and FDO model identification for each 5-minute, real time instant over the allowable time period that stretches from day 15 to day 312 (approx. step 86,000). Each of the approx. 78,000 model is identified and may be used for analyzing of the sub-system around the

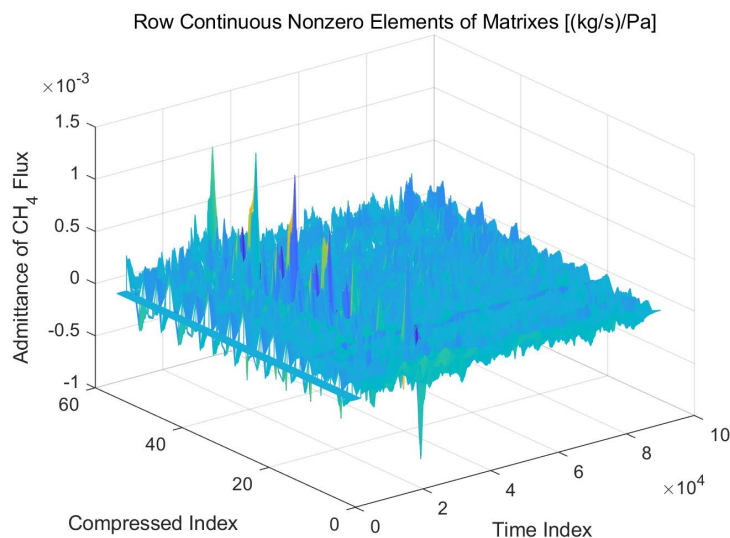


Figure 9. The variation of the 55 nonzero elements of the compressed, row-continuous matrixes of FDO over all time steps ($N = 10$).

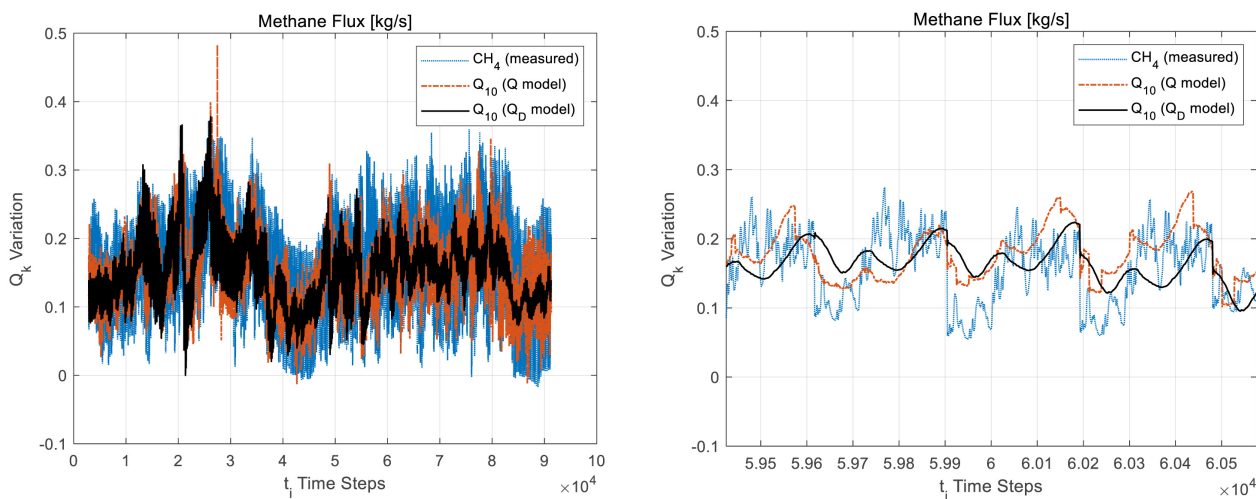


Figure 10. Measured and modeled methane flux variation over 300 days: $\hat{Q}_{10}(10, t_i)$, measured; $Q_{10}(10, t_i)$, predicted by the averaged, full Q model; and $Q_{10}(10, t_i)$, from the Q_D model, (left); enlargement for four days (right).

subject time instant. The 15×15 lower triangle, sparse matrixes are stored in a compressed form arranging the non-zero elements row-continuously into single, 120-element column vectors for further processing. **Figure 12** shows all such vectors plotted along the time steps.

The measured and the modeled methane flux variation are shown in **Figure 13** for over 300 days (left), and for a four-day time period as an enlargement (right). **Figure 13** shows three overlaid graphs such as 1) the measured Q-vector's last element, $\hat{Q}_{15}(15, t_i)$; 2) the last element of $Q_{15}(15, t_i)$ from the full FDO model using prediction model (22); and 3) the last element of $Q_{15}(15, t_i)$ from the simplified, diagonal FDO model using diagonalization (18) and prediction model (22). As shown, graph (1) of the measured methane flux is followed

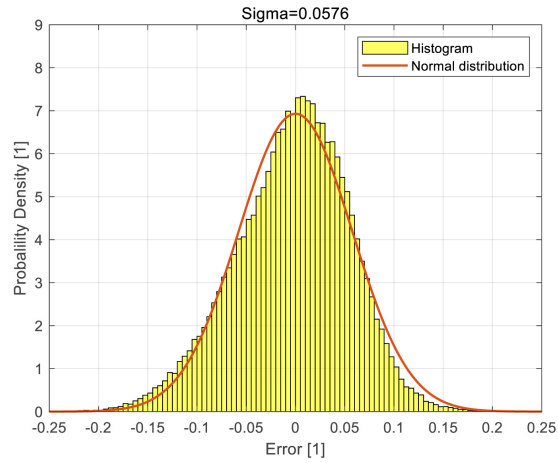


Figure 11. Probability density function and histogram of the error of fit of FDO model ($N = 10$, $RMS = 0.0576$).

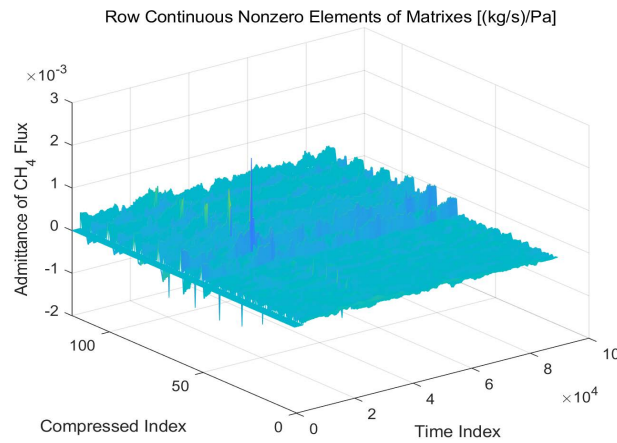


Figure 12. The variation of the 120 nonzero elements of the compressed, row-continuous matrixes of FDO over all time steps ($N = 15$).

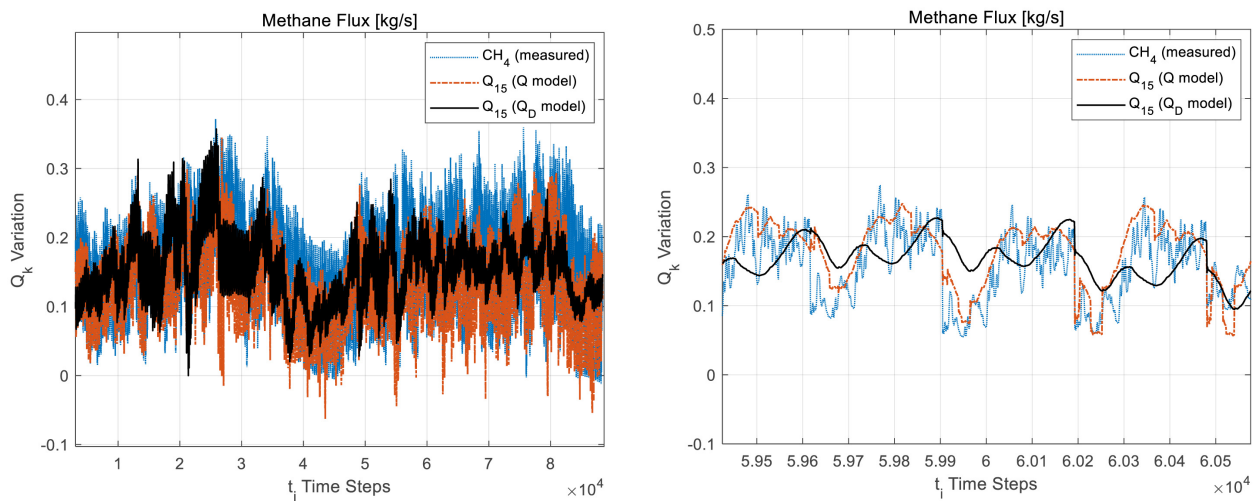


Figure 13. Measured and modeled methane flux variation over 300 days: $\hat{Q}_{15}(15, t_i)$, measured; $Q_{15}(15, t_i)$, predicted by the averaged, full Q model; and $Q_{15}(15, t_i)$, from the Q_D model, (left); enlargement for four days (right).

well with graphs (2) and (3) from the two models. As expected, graph (2) from the full FDO model shows a closer match to graph (1) than graph (3) from the diagonal FDO model. The FDO models results in the second Occam's Razor example are favorably compare with those in the first example, showing that increased training time and model size are still advantageous. However, longer training time delays the availability of the FDO model for real-time analysis and forecasting.

The error of fit of the FDO model to input data and the optimization procedure are characterized by the probability density function and histogram, depicted in **Figure 14**, showing a close to normal error distribution with a slight offset at the lowest part of the histogram curve, indicating a process bias that may be caused by an unknown side-effect, trended as systematic error by the DU model identification.

The methane concentration at the tail gate, c^{TG} , the ultimate parameter of interest for work safety, is calculated back based on (21) from the full FDO, as well as for the simplified, diagonal model for methane mass flux, Q_M , as:

$$c^{TG} = \hat{c}^{MG} + Q_M / \hat{Q}a. \quad (23)$$

As seen in (23), c^{TG} may be affected by three components, \hat{c}^{MG} , Q_M from the FDO model affected by \hat{P} , and $\hat{Q}a$. Therefore, the prediction of Q_M from the FDO model becomes part of a three-variable input-output algebraic model, \mathcal{F} , for methane concentration analysis and forecast as:

$$c^{TG} = \mathcal{F}(\hat{c}^{MG}, \hat{Q}a, \hat{P}). \quad (24)$$

Using the available data, the quality of (23) and (24) is checked, comparing the calculated results from (23) with the measured results, \hat{c}^{MG} , for the full and the simplified, diagonal models for Q_M , depicted in **Figure 15** with the three

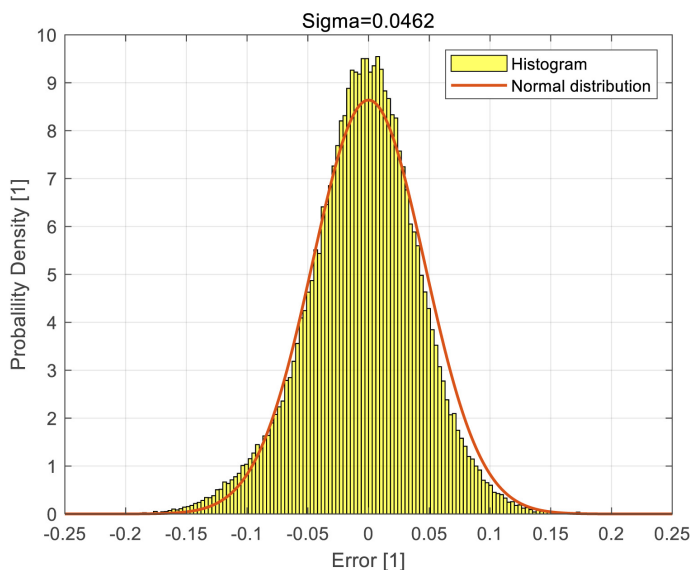


Figure 14. Probability density function and histogram of the error of fit of FDO model for methane flux ($N= 15$, $RMS = 0.0941$).

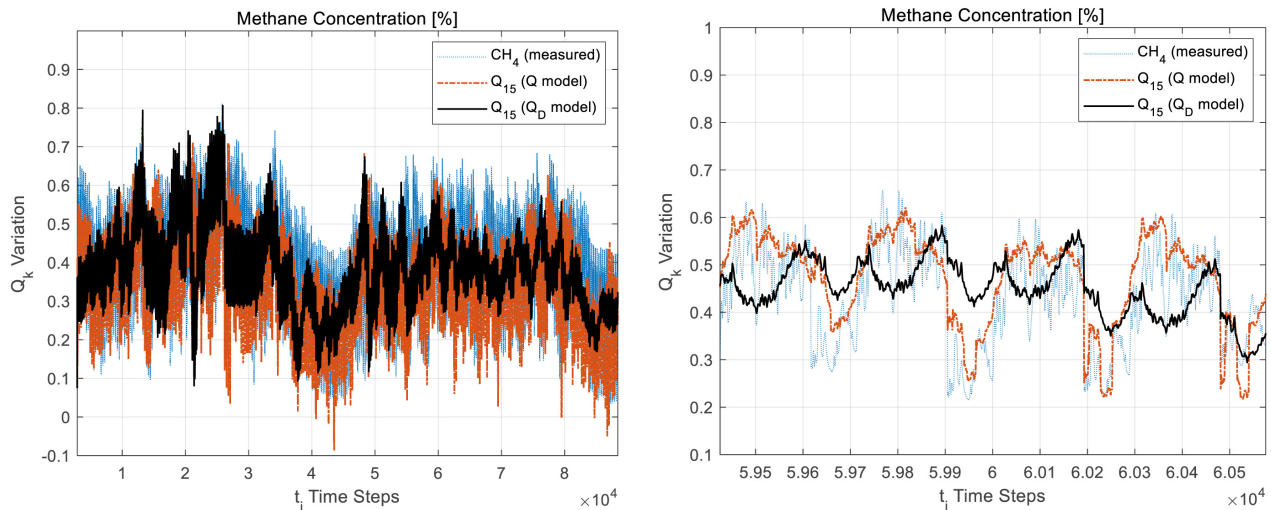


Figure 15. Measured and modeled methane concentration variation (converted into per cent from mass fraction) over 300 days: $\hat{Q}_{15}(15, t_i)$, measured; $Q_{15}(15, t_i)$, predicted by the averaged, full Q model; and $Q_{15}(15, t_i)$, from the Q_D model, (left); enlargement for four days (right).

curves, (1)-(3), similarly defined to those in **Figure 13**. As shown, the concentration trends and matches from (23) are similar to those presented in **Figure 13** for Q_M .

4.3.3. The Third Occam's Razor Example for an FDO Model

Is derived from the second by time-averaging the 78,000 individual FDO models. Averaging further simplifies the individual FDO models, replacing them with a single matrix for the entire time period. The mesh of the averaged, restored, full matrix operator of the single FDO is shown in **Figure 16** in two views: in the positive (right), and negative (left) domains. The averaged matrix should be comparable to the single matrix of M_{MA} obtained in 4.2.

The measured and modeled methane flux variations are shown in **Figure 17** (left) over 300 days, depicting $\hat{Q}_{15}(15, t_i)$, measured, $Q_{15}(15, t_i)$, predicted by the averaged, full FDO model, and $Q_{15}(15, t_i)$, predicted by the diagonal FDO model, on the left. Enlargement of the curves for a four days' time period is also shown in **Figure 17** (right).

Figure 18 shows the error of fit of the averaged FDO model to input data and its optimization procedure, comparing the probability density function and histogram of the error. As depicted, the error is greater for the averaged FDO model (RMS = 0.0627) than that for the individual model's case in 4.2 (RMS = 0.0433); but is still close to the normal distribution.

4.4. Brief Discussion of the FDO Models in 4.2 and 4.3

The FDO models described in 4.2 and 4.3 have shown simulation results using monitored data under normal operations for both methane mass flux liberated from the strata and the resulting methane concentration variations in the air, correlated with the variation of barometric pressure. The variation of the sampled

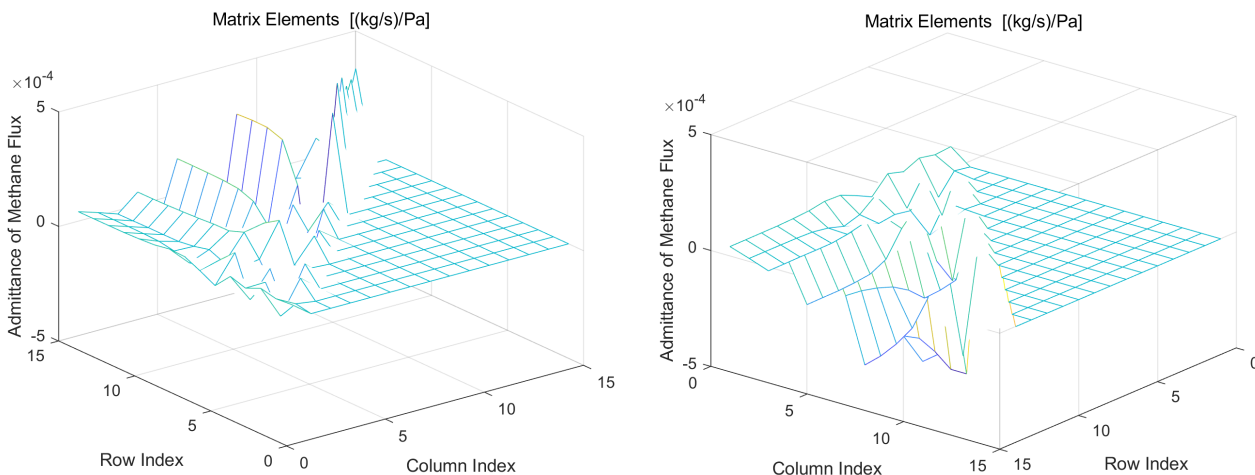


Figure 16. Average of the individual matrices for the FDO ($N = 15$) of pressure-driven methane transport in the positive domain (left) and negative domain (right).

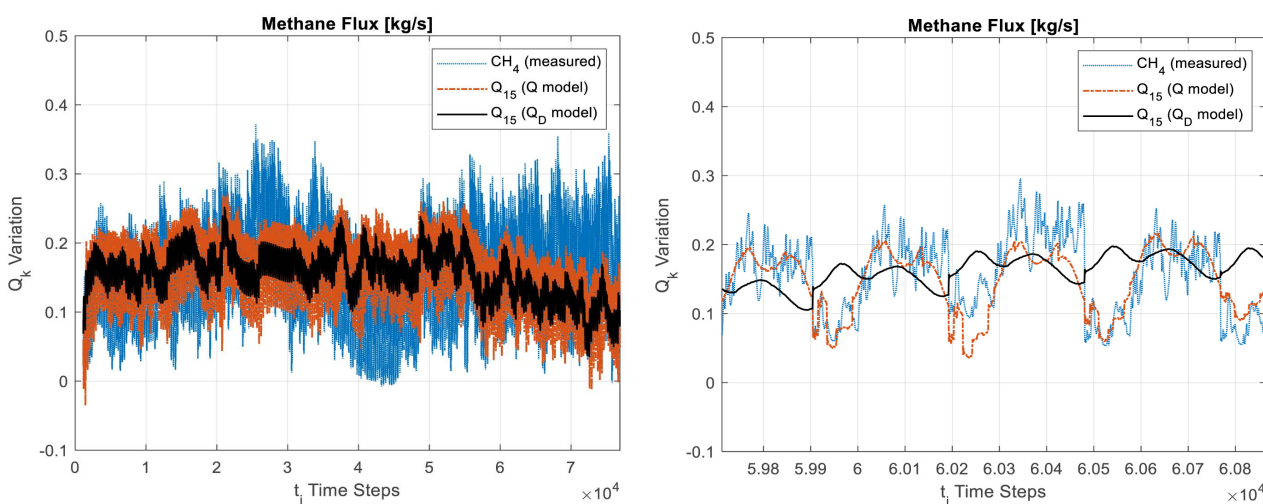


Figure 17. Measured and modeled methane flux variations over 300 days: $\hat{Q}_{15}(15, t_i)$, measured, $Q_{15}(15, t_i)$, predicted by the averaged Q model; and $Q_{15}(15, t_i)$, from the Q_D model, (left); enlargement for four days (right).

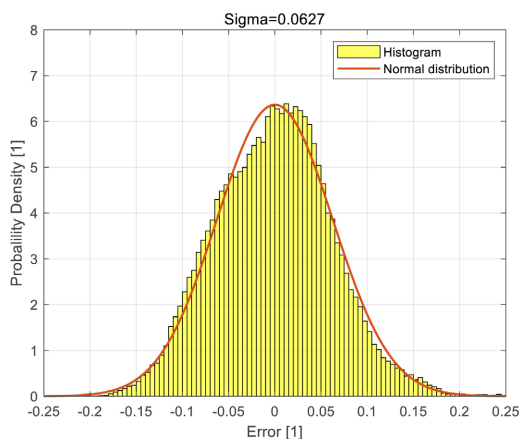


Figure 18. Probability density function and histogram of the error of fit of the averaged FDO model ($N = 15$, $RMS = 0.0627$).

data with time for both methane mass flux and barometric pressure under normal operation are moderate and acceptable for the safety of work at the longwall face. The variations may be considered normal, attributed to the minute-wise, hourly, and daily, periodically changing working conditions, the movement of the shearer, the progression of the moving longwall face and the shields, and the variations of the incoming air flow rate and its pressure, affected the upstream part of the ventilation system. As shown, the FDO model predictions performed remarkably well against monitored data from normal operations, especially those with simplifications following the philosophy of Occam's Razor. All FDO models are obtained by the DU method of the methane transport system. Therefore, they are ready for analyzing the effect of extreme changes in the input variables, such as in \hat{c}^{MG} , $\hat{Q}a$, and \hat{P} listed in (24), affecting the highest methane concentration, c^{TG} , at the working face.

4.5. Pressure-Driven, Extreme Methane Liberation Analysis and Forecasting

The FDO models described in 4.2 and 4.3 are all tested for analyzing the effect of sudden, extreme barometric pressure variations in the working face. The synthetic, extreme barometric pressure variation is created from the monitored \hat{P} data steam by superimposing two, negative, absolute barometric pressure spikes starting at the 40,000 and 60,000 time steps, shown in **Figure 19** as positive difference steps of $\hat{P}_{\max} - \hat{P}$ over all time steps. The absolute pressure spikes include 10 steps of linear descent by 5000 [Pa], 10 steps of the negative plateau, and 10 steps of linear ascent to the undisturbed value, all shown as positive pressure difference changes in **Figure 19**.

4.5.1. ORP1: Perturbed Input Pressure Prediction Using the First Occam's Razor's FDO Model

The model prediction result for methane flux using the subject FDO with the

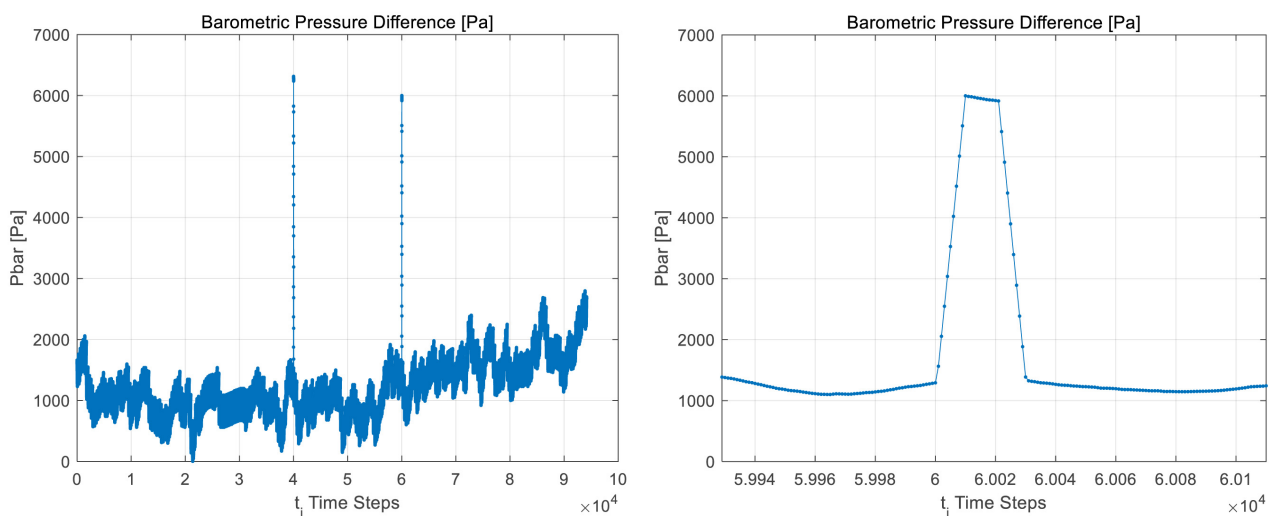


Figure 19. Assumed barometric pressure difference variation with two synthetic, extreme, barometric pressure spikes at 40,000 and 60,000 time steps (left); and enlargement of one spike (right).

hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 20**. The modeled methane flux variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; 2) $Q_{10}(10, t_i)$, perturbed, predicted by the full Q model; and 3) $Q_{10}(10, t_i)$, from the Q_D model, (left). **Figure 20** (on the right) shows enlargement for four days of curves (1)-(3).

The model prediction result for the percentage of methane concentration using the subject FDO with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 21**. The modeled methane concentration variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; 2) $Q_{10}(10, t_i)$, perturbed, predicted by the full Q model;

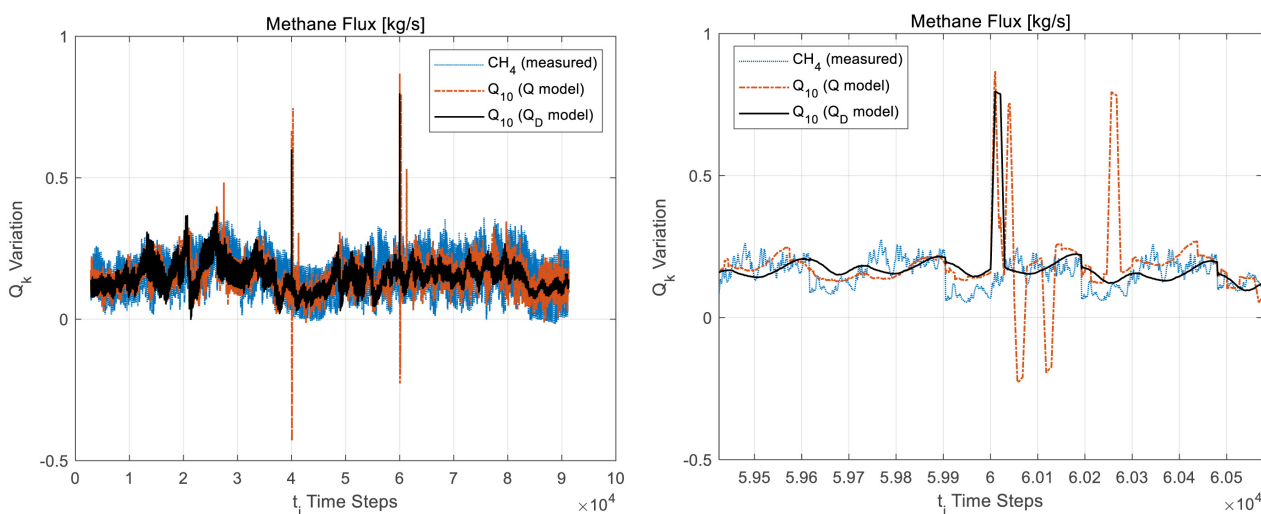


Figure 20. (ORP1) Modeled methane flux over 300 days: $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; $Q_{10}(10, t_i)$, perturbed, predicted Q model; and $Q_{10}(10, t_i)$, Q_D model, (left); enlargement for four days (right).

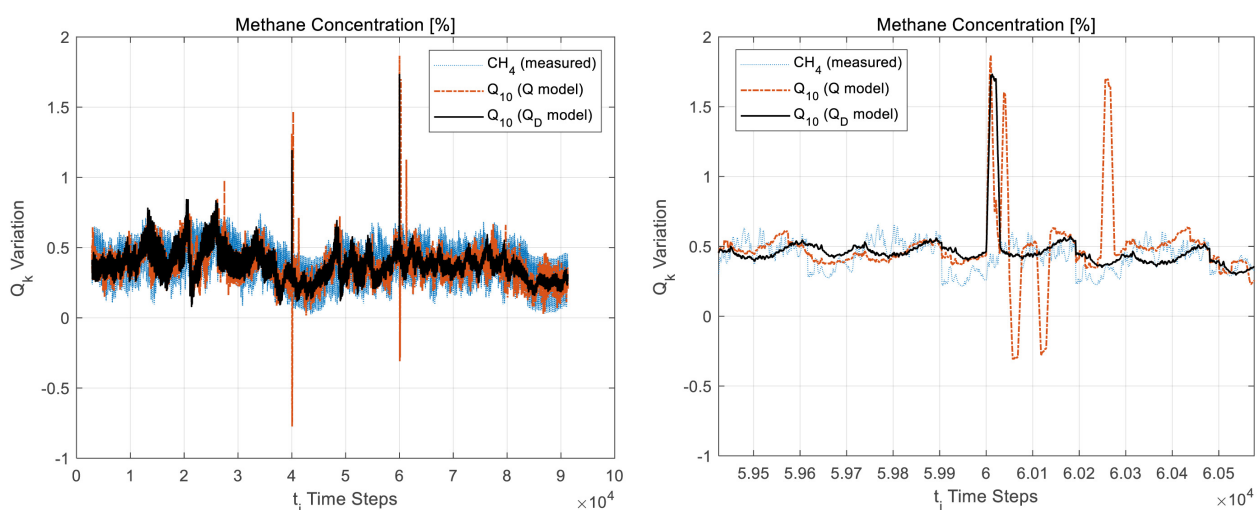


Figure 21. (ORP1) Modeled methane concentration over 300 days: $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; $Q_{10}(10, t_i)$, perturbed, predicted Q model; and $Q_{10}(10, t_i)$, Q_D model, (left); enlargement for four days (right).

and 3) $Q_{10}(10, t_i)$, from the Q_D model, (left). **Figure 21** (on the right) shows enlargement for four days of curves (1)-(3).

4.5.2. ORP2: Perturbed Input Pressure Prediction Using the Second Occam's Razor's FDO Model

The model prediction result for methane flux using the subject FDO with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 22**. The modeled methane flux variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; 2) $Q_{15}(15, t_i)$, perturbed, predicted by the full Q model; and 3) $Q_{15}(15, t_i)$, from the Q_D model, (left). **Figure 22** (on the right) shows enlargement for four days of curves (1)-(3).

The model prediction result for the percentage of methane concentration using the subject FDO with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 23**. The modeled methane concentration variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; 2) $Q_{15}(15, t_i)$, perturbed, predicted by the full Q model; and 3) $Q_{15}(15, t_i)$, from the Q_D model, (left). **Figure 23** (on the right) shows enlargement for four days of curves (1)-(3).

4.5.3. ORP3: Perturbed Input Pressure Prediction Using the Third, Averaged-Type Occam's Razor's FDO Model with the 10×10 Model Size

The model prediction result for methane flux using the subject, averaged FDO with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 24**. The modeled methane flux variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; 2) $Q_{10}(10, t_i)$, perturbed, predicted by the averaged Q model; and 3) $Q_{10}(10, t_i)$, from the Q_D model, (left). **Figure 24** (on the right) shows enlargement for four days of curves (1)-(3).

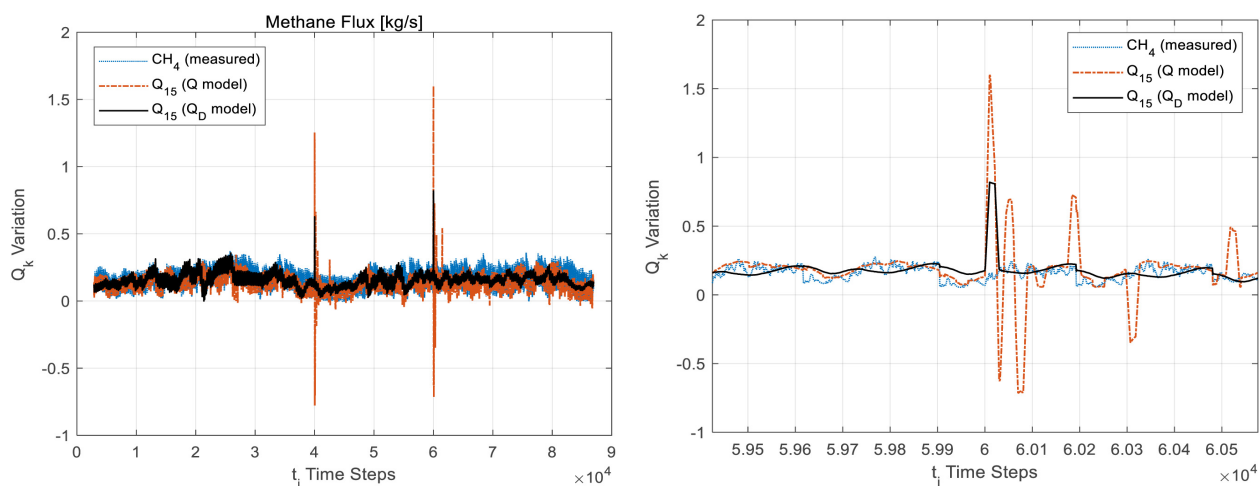


Figure 22. (ORP2) Modeled methane flux over 300 days: $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; $Q_{15}(15, t_i)$, perturbed, predicted Q model; and $Q_{15}(15, t_i)$, Q_D model, (left); enlargement for four days (right).

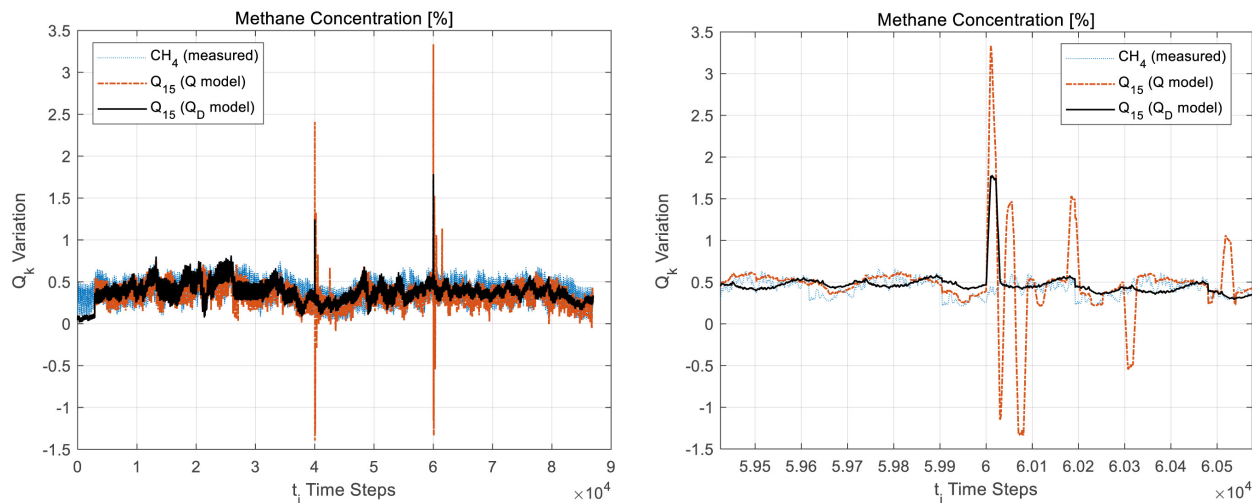


Figure 23. (ORP2) Modeled methane concentration variations over 300 days: $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; $Q_{15}(15, t_i)$, perturbed, full Q model; and $Q_{15}(15, t_i)$, Q_D model, (left); enlargement for four days (right).

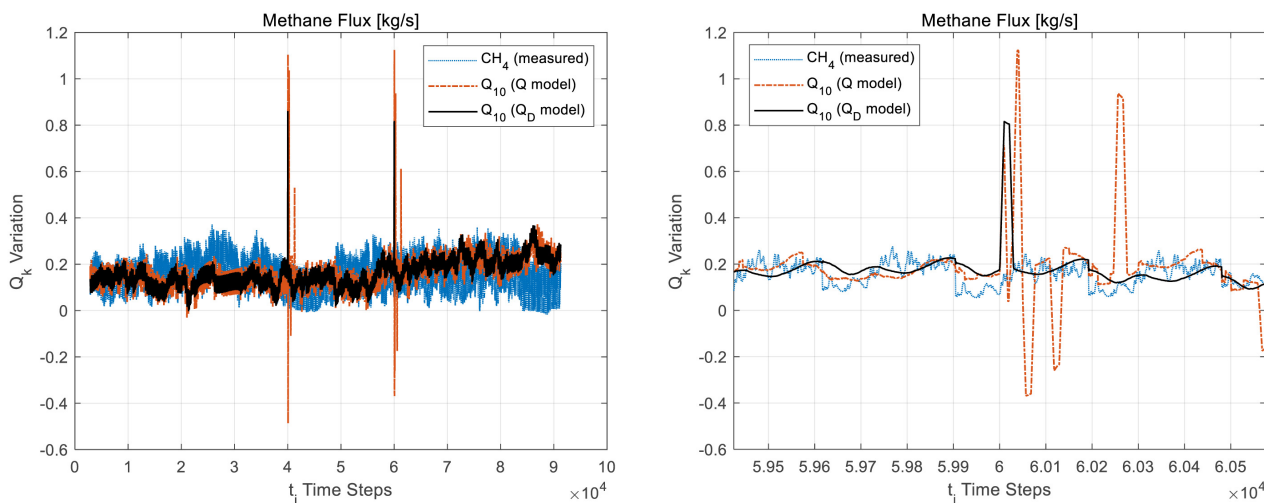


Figure 24. (ORP3) Modeled methane flux variations over 300 days: $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; $Q_{10}(10, t_i)$, perturbed, predicted by the averaged Q model; and $Q_{10}(10, t_i)$, from the Q_D model, (left); enlargement for four days (right).

The model prediction result for the percentage of methane concentration using the subject, averaged FDO with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 25**. The modeled methane concentration variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; 2) $Q_{10}(10, t_i)$, perturbed, predicted by the averaged Q model; and 3) $Q_{10}(10, t_i)$, from the Q_D model, (left). **Figure 25** (on the right) shows enlargement for four days of curves (1)-(3).

4.5.4. ORP4: Perturbed Input Pressure Prediction Using the Third, Averaged-Type Occam's Razor's FDO Model with the 15×15 Model Size

The model prediction result for methane flux using the subject, averaged FDO

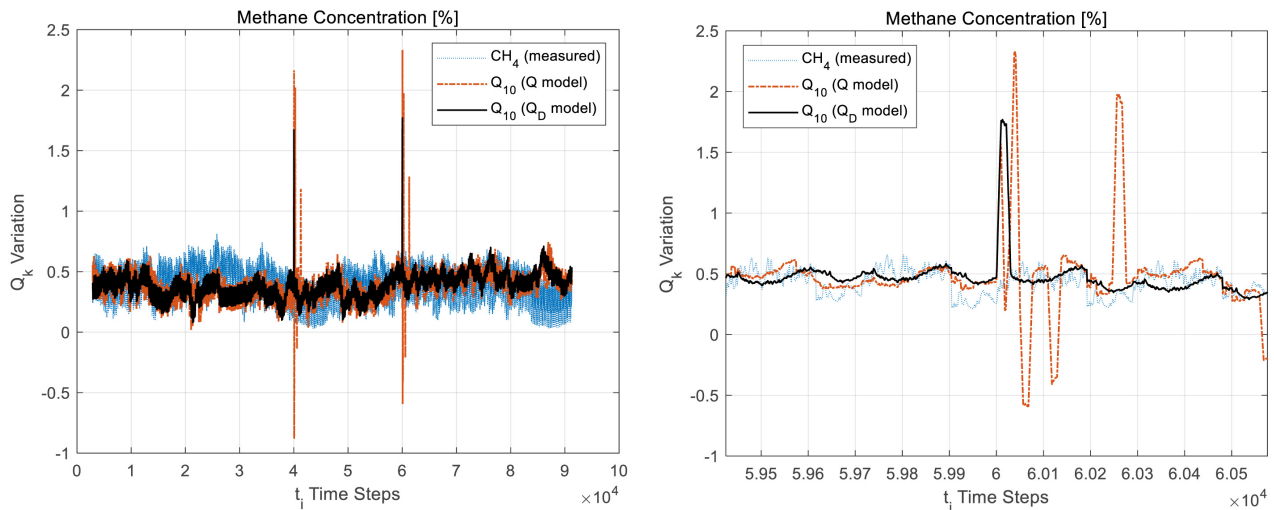


Figure 25. (ORP3) Modeled methane concentration over 300 days: $\hat{Q}_{10}(10, t_i)$, measured, unperturbed; $Q_{10}(10, t_i)$, perturbed, averaged Q model; and $Q_{10}(10, t_i)$, Q_D model, (left); enlargement for four days (right).

with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 26**. The modeled methane flux variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; 2) $Q_{15}(15, t_i)$, perturbed, predicted by the averaged Q model; and 3) $Q_{15}(15, t_i)$, from the Q_D model, (left). **Figure 26** (on the right) shows enlargement for four days of curves (1)-(3).

The model prediction result for the percentage of methane concentration using the subject, averaged FDO with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 27**. The modeled methane concentration variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; 2) $Q_{15}(15, t_i)$, perturbed, predicted by the averaged Q model; and 3) $Q_{15}(15, t_i)$, from the Q_D model, (left). **Figure 27** (on the right) shows enlargement for four days of curves (1)-(3).

4.5.5. ORP5: Perturbed Input Pressure Prediction Using the Full, Averaged FDO Model with 35×35 Model Size

The model prediction result for methane flux using the subject, averaged FDO, M_{MA} , from 4.2 with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 28**. The modeled methane flux variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{35}(35, t_i)$, measured, unperturbed; 2) $Q_{35}(35, t_i)$, perturbed, predicted by the averaged Q model; and 3) $Q_{35}(35, t_i)$, from the Q_D model, (left). **Figure 28** (on the right) shows enlargement for four days of curves (1)-(3).

The model prediction result for the percentage of methane concentration using the subject, averaged FDO with the hypothetical, synthetic, perturbed barometric pressure input is shown in **Figure 29**. The modeled methane concentration variations over 300 days are depicted (on the left) by three curves: 1) $\hat{Q}_{35}(35, t_i)$, measured, unperturbed; 2) $Q_{35}(35, t_i)$, perturbed, predicted by the

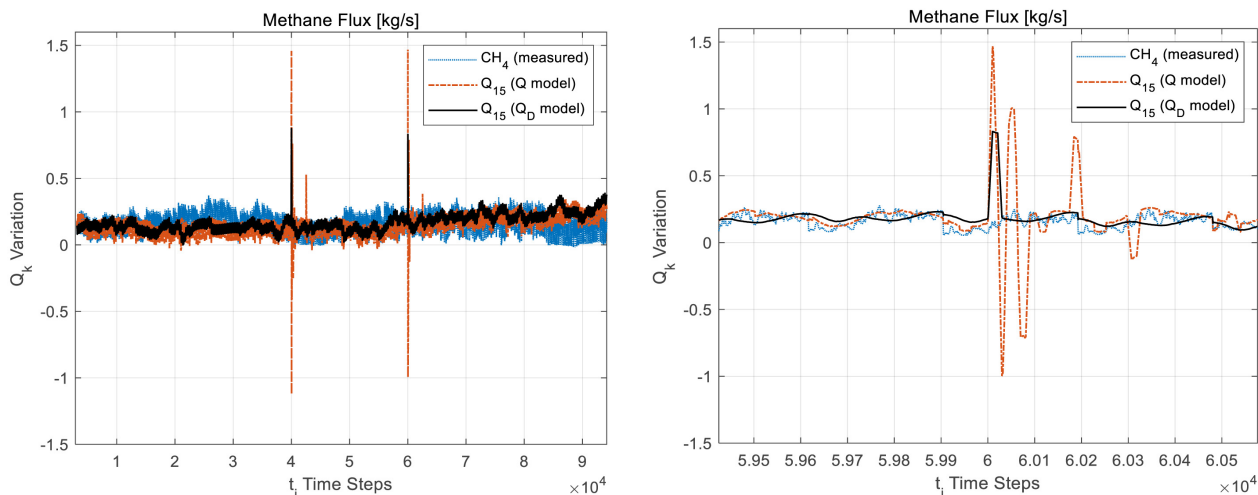


Figure 26. (ORP4) Modeled methane flux over 300 days: $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; $Q_{15}(15, t_i)$, perturbed, averaged Q model; and $Q_{15}(15, t_i)$, Q_D model, (left); enlargement for four days (right).

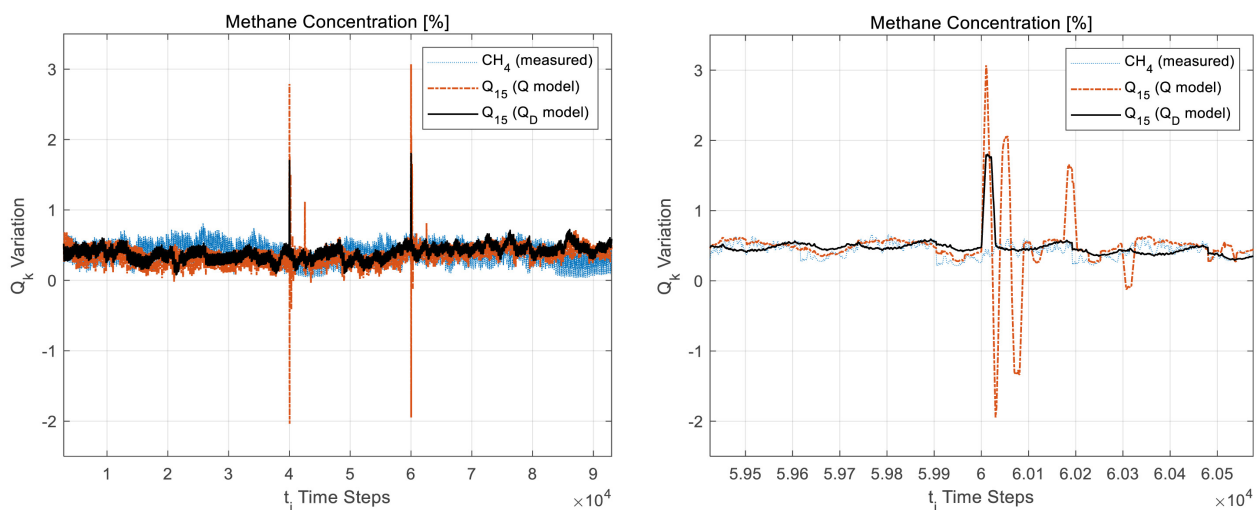


Figure 27. (ORP4) Modeled methane concentration over 300 days: $\hat{Q}_{15}(15, t_i)$, measured, unperturbed; $Q_{15}(15, t_i)$, perturbed, averaged Q model; and $Q_{15}(15, t_i)$, Q_D model, (left); enlargement for four days (right).

averaged Q model; and 3) $Q_{35}(35, t_i)$, from the Q_D model, (left). **Figure 28** (on the right) shows enlargement for four days of curves (1)-(3).

4.5.6. ORP6: Perturbed Input Pressure Prediction Using the PBO Model with 35×35 Model Size

The PBO model is used for predicting methane release as a response to the fine scale, perturbed, pressure difference input driving force of $\hat{P}_{max} - \hat{P}$, shown in **Figure 19**. A further calibration multiplier was necessary to re-calibrate the PBO matrix elements for matching the predicted, average rate of methane release for 300 days.

The simulated methane release rate variations over 300 days are shown in **Figure 30** by three curves: 1) $\hat{Q}_{35}(35, t_i)$, measured, unperturbed; 2) $Q_{35}(35, t_i)$,

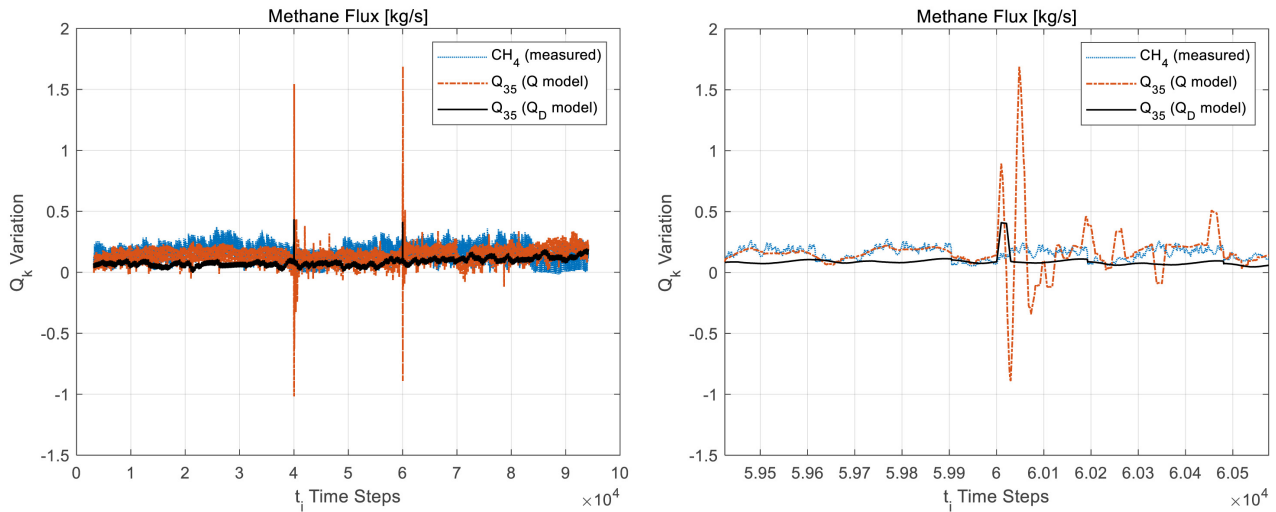


Figure 28. (ORP5) Modeled methane flux variations over 300 days: $\hat{Q}_{35}(35, t_i)$, measured, unperturbed; $Q_{35}(35, t_i)$, perturbed, predicted by the averaged Q model; and $Q_{35}(35, t_i)$, from the Q_D model, (left); enlargement for four days (right).

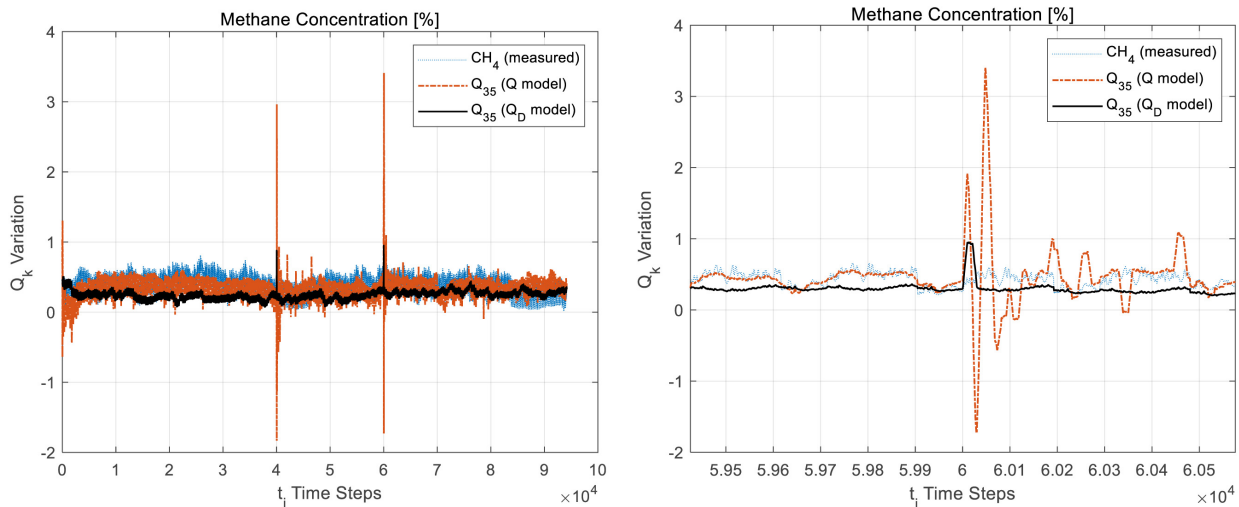


Figure 29. (ORP5) Modeled methane concentration variations over 300 days: $\hat{Q}_{35}(35, t_i)$, measured, unperturbed; $Q_{35}(35, t_i)$, perturbed, predicted by the averaged Q model; and $Q_{35}(35, t_i)$, from the Q_D model, (left); enlargement for four days (right).

perturbed, predicted by the re-calibrated PBO model; and 3) $Q_{35}(35, t_i)$, from the Q_D model, (left). **Figure 30** (on the right) shows enlargement for four days of curves (1)-(3). The re-calibration multiplier, determined by trial-and-error, is verified by **Figure 30**, showing a good match in the average rate of methane release between model and measured data.

Figure 30 shows high positive and negative, random methane release spikes. The positive, random spikes are in the realistic regime against monitored data. The negative spikes are considered linearized model artifacts, as discussed in the foregoing. The PBO model response in methane release increase to the synthetic, negative, absolute pressure spikes are definite and high, about 2 to 3 times higher than those experienced from the FDO models.

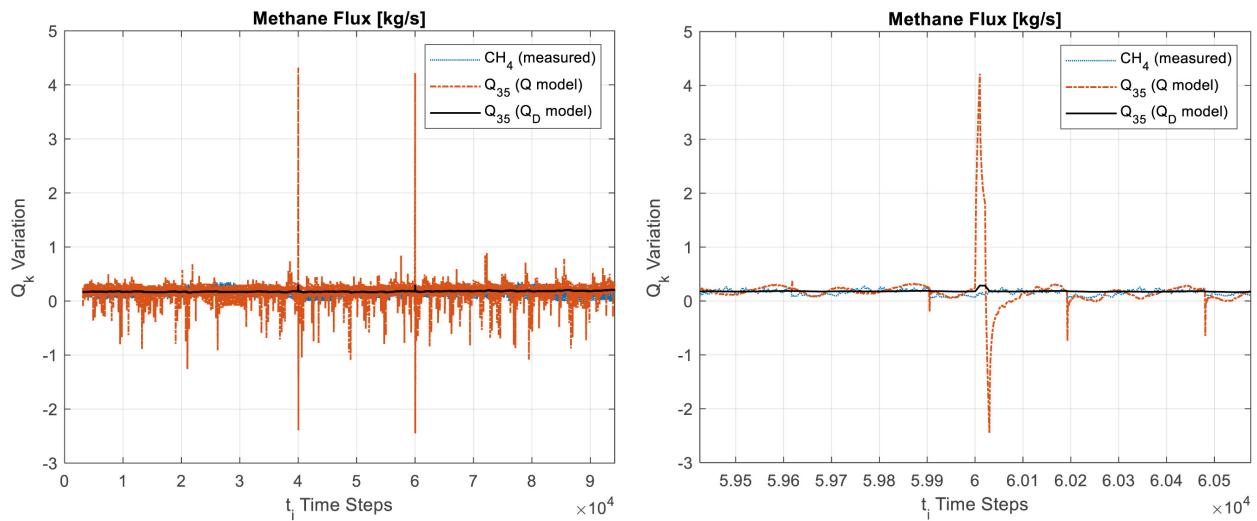


Figure 30. (ORP6) Modeled, perturbed methane flux variations together with unperturbed measurement results over 300 days: $\hat{Q}_{35}(35, t_i)$, measured, unperturbed; $Q_{35}(35, t_i)$, perturbed, predicted by the calibrated, analytical PBO Q model; and $Q_{35}(35, t_i)$, from the analytical PBO Q_D model, (left); enlargement for four days (right).

4.6. Brief Discussion of Pressure-Driven, Extreme Methane Liberation Analysis Results and Forecasting Applications

4.6.1. Interpretation and Application of the FDO Model for Analysis

Application of the FDO model with strongly perturbed, suddenly increased barometric pressure input, significantly different in shape and nature from the monitored data used in FDO model identification raises the risk of errors due to the potential presence of false correlation in the input and output data streams. Such false correlation may be suspected due to the common influence of mining activities and operations at the longwall face upon both monitored and recorded input pressure and output methane concentration data. Lacking proof to exclude it, a hypothesis must be made about the lack of such false correlation for accepting the results presented hereby for FDO model analysis forecasting.

As depicted in **Figures 20-29**, the model predictions for methane liberation due to pressure-depression pulses generally agree, all showing responses in methane flux pulses within a -1.5 to 1.5 [kg/s] range for the 300 m face drift. The larger the model size from $N = 10$ to 35, the higher the methane flux peaks in both the negative and the positive direction.

The extreme, negative values from a linear model must be interpreted as a physically invalid extrapolation, well outside the small, all-positive validity range, that is used in model identification. Only the positive methane flux, entering the air flow is physically plausible. Methane mass flux may pulsate within a positive domain up and down, such as observed in **Figure 4** (with only a minute violation of the range at a few points, indicating likely measurement error uncertainty in the input data) used as the input data for model identification. Methane flow direction from the air flow into the strata is inconceivable, as methane mixes with air rapidly and only a low-concentration mixture (but not pure methane) may be driven back to the porous and fractured strata under forcing

driving pressure.

However, the tendency to block methane inflow right after the start of the Q-vector of the negative pressure pulse reveals useful methane transport information about the linearized system. It would be counterproductive to consider forcefully eliminating the unwanted negatives from the results. The beginning part of the Q-vector refers to the oldest data in the sliding window, varying from about 5.5 days ($N=10$) to 70 days ($N=35$) as shown in **Figure 1**. The beginning part of the $\hat{P}_{\max} - \hat{P}$ pressure Q-vector may be initiating methane blockage from a farther distance in the methane-bearing strata and starting a “peristaltic wave” of methane flux into the airway, culminating in a series of positive and negative pulses toward the latest time, showing dynamic fluctuations and definite time delays in **Figures 20-29**.

The diagonal simplification of the FDO model in (18) replaces M with a diagonal matrix, M_D . The physical meaning of diagonalization is an assumed history variation change in the driving force's Q-vector (that is, in \hat{P}), forcing to replace the real, variable history (used during model identification) into an abstracted, step change variation in \hat{P} , constant from the beginning to the most recent value. The step change solution eliminates the delayed, “fly-wheel” effects in the prediction for methane flux with time in the simplified, diagonal model solution.

It is interesting to see that all diagonal, M_D model solutions (marked as Q_D curves), show acceptable (albeit somewhat simplified) match to the measured, unperturbed data, but robust, single, positive methane flux as well as methane concentration responses to pulse-type perturbations. Note that M_D inherits all information carried in the full M matrix of the FDO model, and it integrates, rather than forcefully eliminates unwanted response components.

It is convenient to use the modeled pulses for the methane flux rates and the methane concentrations from the M_D models for their apparent stability marked as Q_D curves in **Figures 19-29**. The small size, sliding window, full model with $N=10$ appears to be satisfactory for the response analysis to barometric pressure depression pulses. Each model is trained over a 10-day sliding time window, with past data sampled every 5 minutes for 2880 readings for processing the input Q vectors. Each model is evaluated within 10^{-3} sec on a regular laptop computer. Each model is updated at every 5 minutes, allowing to adjust to operational changes, and evaluate model consistence relative to previous models.

4.6.2. Comment on Checking the Prospect of False Correlation in the FDO Models

The PBO model is used for predicting methane release as a response to the the fine scale, perturbed, pressure difference input shown in **Figure 30**. An acceptable match is seen between the predicted, average rate of methane release from the re-calibrated PBO against the measured data for 300 days. The PBO results show high sensitivity with positive and negative, random methane release spikes.

Nevertheless, the PBO model response in methane release increase to the

synthetic, negative, absolute pressure spikes are definite and high, about 2 to 3 times higher than those experienced from the FDO models.

In spite of understandable differences from the FDO models, the PBO model verifies a definite, physically plausible, deterministic methane release due to negative absolute barometric pressure changes. Since the PBO model is free from false correlation effects, its general agreement with the FDO models makes the prospect of false correlations unlikely in the FDO models.

4.6.3. Application of the FDO Model for Forecasting

Each 5-minute FDO model is applicable for a much longer time period than its history time period, backward and forward, as shown in model matching examples in **Figures 20-27** where acceptable difference is seen in the model predictions between the instantaneous (full) and the averaged models over 300 day time periods. A slowly changing FDO model can be used for forecasting response outcome for an upcoming, conceivable, future variation of its input parameters, such as demonstrated in the examples for two, negative, barometric pressure pulses.

Similarly, for example, a sudden fan stoppage in the ventilation network may generate pressure as well as simultaneous air flow rate pulses. Such disturbance may be analyzed with the latest, undisturbed FDO model, and the input Q vectors, substituted in (22) and (23) for the instantaneous forecasting of methane flux liberation and future concentration variation.

For preventive interventions, it is advisable to incorporate the analysis of the mines' methane transport system by identifying their FDO models, and analyze their dynamic responses to any conceivable disturbances for understanding the safety and health risks. This way, the regular forecast information from weather service, or from mine operations risk analysis can be turned into quantitative model input and methane concentration warning data from the FDO models.

5. Concluding Remarks

A matrix operator model is developed and tested for a system model building and time-series data analysis. The deconvolution procedure identifies an FDO model from monitored, uncontrolled data streams using both system input and response output vectors obtained over past time intervals.

The FDO model is applicable to predict future output features for deviated input vectors from the expected, future input for hazard evaluation, for example, for methane inflow into the working face in an underground mine.

The results show that a complex gas inflow system in a mine example can be approximated well within reasons by a PBO model only for processes that are well understood, filtered, and simplified.

FDO models may be identified for matching the unexplainable, short-term variations with time in the measured data of methane inflow. The numerical coefficients of PBO and FDO model are found to differ by two to three orders of magnitude for methane release as a function of short-time barometric pressure

variations.

For real-time data analysis and forecasting, Occam's Razor is used to reduce the size of N and as well as to revert to a simplified model according to (18), converting the full matrix to diagonal elements for eliminating model complexity. Experimentation is necessary to find the best combination of N (size of operator and Q-vectors), nT (history time span, that is, the sliding time window with N divisions), and k (number of repetitions of the Q-vectors at Δt successions).

A hypothesis is formulated for the learnability of FDO of a physical process from uncontrolled input-output data streams in 3.4 intuitively, supported by experimentation with real word data, FDO model identification, model fit tests to input data, and model prediction power measured from resilience against input variations in different shapes from those used in model training.

As being data-driven, the significantly different results from an FDO versus the PBO models may shed light of methane release processes poorly understood and modeled in PBO, missing the relevant boundary conditions or input data interpretation for the modeled processes.

However, problems in the monitored and recorded data fluctuations, their quality and inter-dependence may still be another reason for differences between the PBO and FDO models. False input-output relations may be perceived if both inputs and outputs are caused by a third, common disturbing factor, not entered as an independent input in the model identification procedure. False correlation may appear in the application of the model which will be valid only when the third, common factor is also present.

An experimental proof, using a re-calibrated PBO model is shown to likely exclude the likelihood of false correlation in the presented study.

Acknowledgements

The financial support for the studies by the Alpha Foundation of Mine Safety and Health is gratefully acknowledged.

Conflicts of Interest

No conflict is reported.

References

- [1] Eykhoff, P. (1974) System Identification, Parameter and State Estimation. 2nd ed., John Wiley & Sons, Hoboken.
- [2] DeFigueiredo, R. and Dwyer, T. (1980) A Best Approximation Framework and Implementation for Simulation of Large-Scale Nonlinear Systems. *IEEE Transactions on Circuits and Systems*, **27**, 1005-1014. <https://doi.org/10.1109/TCS.1980.1084741>
- [3] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444. <https://doi.org/10.1038/nature14539>
- [4] Yamashita, R., Nishio, M., Do, R.K.G. and Togashi, K. (2018) Convolutional Neural Networks: An Overview and Application in Radiology. *Insights into Imaging*, **9**, 611-629. <https://doi.org/10.1007/s13244-018-0639-9>

-
- [5] Huang, W., Shen, Z., Huang, N.E. and Fung, Y.C. (1998) Use of Intrinsic Modes in Biology: Examples of Indicial Response of Pulmonary Blood Pressure to 6 Step Hypoxia. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 12766-12771. <https://doi.org/10.1073/pnas.95.22.12766>
- [6] de Miranda, S., Patruno, L., Ubertain, F. and Vairo, G. (2013) Indicial Functions and Flutter Derivatives: A Generalized Approach to the Motion-Related Wind Loads. *Journal of Fluids and Structures*, **42**, 466-487. <https://doi.org/10.1016/j.jfluidstructs.2013.08.009>
- [7] Carslaw, H.S. and Jaeger, J.C. (1986) *Conduction of Heat in Solids*. 2nd ed., Oxford Science Publications, New York, 30.
- [8] Ansys Fluent, Fluid Simulation Software (2022) <https://www.ansys.com>
- [9] Tough2, Numerical Simulation Code for Multi-dimensional Coupled Fluid and Heat Flow of Multiphase, Multicomponent Fluid Mixtures in Porous and Fractured Medium, (2022) <https://tough.lbl.gov>
- [10] Ventsim, 3D Mine Ventilation Software (2022) <https://www.howden.com>
- [11] Danko, G. (2017) *Model Elements and Network Solutions of Heat, Mass and Momentum Transport Processes*. Springer-Verlag GmbH Germany, 1-251. <https://doi.org/10.1007/978-3-662-52931-7>
- [12] Danko, G. (2006) Functional or Operator Representation of Numerical Heat and Mass Transport Models. *Journal of Heat Transfer*, **128**, 162-175. <https://doi.org/10.1115/1.2136919>
- [13] Danko, G. (2021) Quantum Operator Model for Data Analysis and Forecast. *Applied Mathematics*, **12**, 963-992. <https://www.scirp.org/journal/am> <https://doi.org/10.4236/am.2021.1211064>
- [14] Blumer, A., Ehrenfeucht, A., Haussler, D. and Warmuth, M.K. (1987) Occam's Razor. *Information Processing Letters*, **24**, 377-380. [https://doi.org/10.1016/0020-0190\(87\)90114-1](https://doi.org/10.1016/0020-0190(87)90114-1)