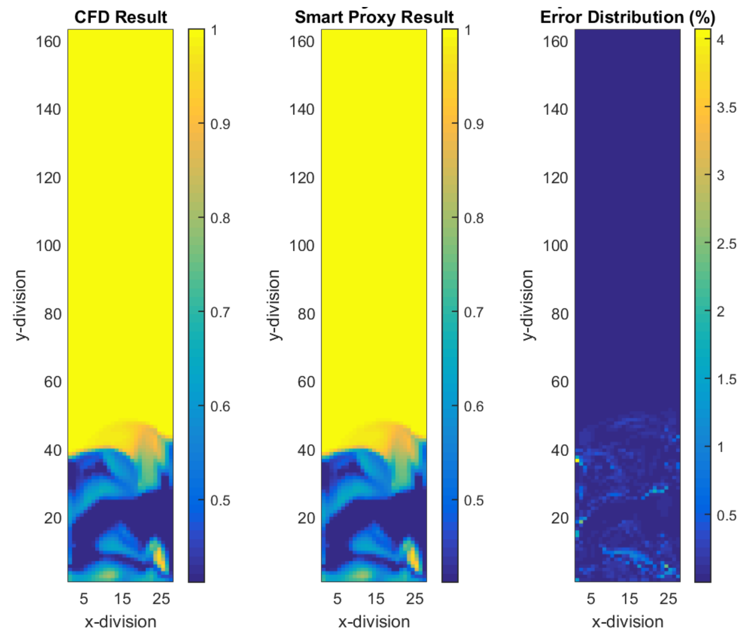


Data Driven Smart Proxy for CFD

Application of Big Data Analytics & Machine Learning in Computational Fluid Dynamics

Report One: Proof of Concept

November 2017

Office of Fossil Energy

NETL-PUB-21574

Disclaimer

Cover Illustration: A 2D slice of voidage contours of a fluidized bed. CFD results (left), AI-based smart proxy (middle) and percent error (right).

Suggested Citation: Ansari, A., Mohaghegh, S., Shahnam, M., Dietiker, J. F., Takbiri Borujeni, A., Fathi, E., *Data Driven Smart Proxy for CFD: Application of Big Data Analytics & Machine Learning in Computational Fluid Dynamics, Part One: Proof of Concept*; NETL-PUB-21574; NETL Technical Report Series; U.S. Department of Energy, National Energy Technology Laboratory: Morgantown, WV, 2017.

Data Driven Smart Proxy for CFD

Application of Big Data Analytics & Machine Learning in Computational Fluid Dynamics

Part One: Proof of Concept

Ansari, A.[1], Mohaghegh, S.[1,2], Shahnam, M.[3], Dietiker, J. F.[3,4], Takbiri Borujeni, A.[1], Fathi, E.[1]

[1] **Petroleum & Natural Gas Engineering Department, West Virginia University**
[2] **ORISE Faculty Program**
[3] **Energy Conversion Engineering Directorate, Research and Innovation Center, U.S. Department of Energy, National Energy Technology Laboratory**
[4] **West Virginia University Research Corporation, Morgantown, WV**

NETL-PUB-21574

November 2017

NETL Contacts:

Mehrdad Shahnam, Principal Investigator

Jonathan Lekse, Technical Portfolio Lead

David Alman, Executive Director, Research and Innovation Center

This page intentionally left blank

Table of Contents

List of Figures

List of Tables

Acronyms, Abbreviations, and Symbols

Term	Description
AI	Artificial Intelligence
ANN	Artificial Neural Network
CFD	Computational Fluid Dynamics
CSV	Comma Separated Value
DM	Data Mining
EIA	Energy Information Administration
IGCC	Integrated Coal Gasification Combined Cycle
KPI	Key Performance Indicator
MFIX	Multiphase Flow with Interphase eXchanges
MSE	Mean Square Error
PDE	Partial Differential Equation
RMSE	Root Square of Mean Square Error
VTU	Visualization Toolkit Unstructured points data

Acknowledgments

Professor Mohaghegh acknowledges the support provided by an appointment to the National Energy Technology Laboratory (NETL) Faculty Research Participation Program, sponsored by the U.S. Department of Energy and administered by the Oak Ridge Institute for Science and Education. This work is performed as part of NETL research for the U.S. Department of Energy's Cross Cutting Program.

EXECUTIVE SUMMARY

Simulation technologies can reduce the time and cost of the development and deployment of advanced technologies and allow rapid scale-up of these technologies for fossil fuel based energy systems. However, to ensure their usefulness in practice, the credibility of the simulations needs to be established with Uncertainty Quantification (UQ) methods. National Energy Technology Laboratory (NETL) has been applying non-intrusive UQ methodologies to categorize and quantify uncertainties in CFD simulations of gas-solid multiphase flows. To reduce the computational cost associated with gas-solid flow simulations required for UQ analysis, techniques commonly used in the area of Artificial Intelligence (AI) and Data Mining (DM) are used to construct smart proxy models, which can reduce the computational cost of conducting large number of multiphase CFD simulations.

The feasibility of using AI and machine learning to construct a smart proxy for a gas-solid multiphase flow has been investigated by looking at the flow and particle behavior in a non-reacting rectangular fluidized bed. NETL's in house multiphase solver, MFiX, has been used to generate simulation data for the rectangular fluidized bed. The CFD data is then used to train a smart proxy that can reproduce the CFD results with reasonable error (about 10%). MATLAB neural network toolbox has been used for the current development effort.

1. INTRODUCTION

Fossil fuel continues to be a reliable source of energy for power generation in the United States and worldwide. Technologies, such as chemical looping and gasification, aim to reduce the carbon emission of fossil fuel based power plants. Simulation technologies can reduce the time and cost of the development and deployment of such advanced technologies and allow rapid scale-up of these technologies. Simulation can be used to test new designs to ensure reliable operation under a variety of operating conditions. However, to ensure their usefulness in practice, the credibility of the simulations needs to be established with Uncertainty Quantification (UQ) methods. To this end, National Energy Technology Laboratory (NETL) has been applying non-intrusive UQ methodologies to categorize and quantify uncertainties in CFD simulations of gas-solid multiphase flows, which are encountered in fossil fuel based energy systems [1, 2, 3, 4]. Gas-solid flows are inherently highly unsteady and chaotic flows, where sharp discontinuity can exist at the interface between the phases. The challenge in CFD simulation of gas-solid flows is to adequately resolve the structures that exist at different spatial and temporal scales in an inherently transient flow. Additionally, in reacting gas-solid flow simulations, small time steps are needed in order to not only resolve the temporal scales of the flow, but also ensure numerical stability of the solution. A rule of thumb for adequate spatial resolution is for the grid spacing to be about 10 times the particle diameter [5]. The grid requirement for maintaining such a ratio of grid size to particle diameter for smaller size particles makes such simulations computationally costly and impractical [4]. Recent work at NETL [4] has shown the number of simulations, which is required for uncertainty quantification, can easily exceed many tens of simulations. The spatial and temporal resolution requirements for multiphase flows makes CFD simulations computationally expensive and potentially beyond the reach of many design analysts.

It's clear that a paradigm shift in simulation technology is needed in order to make reacting gas-solid flow CFD simulations with appropriate grid resolution more practical for design and optimization purposes during design scale up. To accelerate the design and analysis process, high fidelity surrogate models that can capture the flow behavior of the design under consideration can be utilized. Surrogate models are increasingly used in design exploration, optimization and sensitivity analysis. Advances in big data analytic and machine learning allows for creation of data-driven metamodels, which can faithfully duplicate the behavior of the data that was used for their construction. This new technology has been successfully applied in the upstream petroleum industry [6] [7] [8] [9]. Smart Proxy modeling takes advantage of pattern recognition capabilities of artificial intelligence and machine learning to build powerful tools to predict the behavior of a system with far less computational cost compared to traditional CFD simulators.

The goal of this research project is to build a smart proxy model at the cell level, which is constructed from simulation data generated by high fidelity CFD models to, in effect, replace the use of computationally expensive CFD for the design space under study for further analysis, optimization and uncertainty quantification. The goal of this portion of research project outlined in this report is to establish proof of concept for the application of this technology to Computational Fluid Dynamics. A smart proxy model, which is constructed from simulation data generated by high fidelity CFD models can in effect replace the use of computationally expensive CFD for the design space under study and further analysis and optimization. The smart proxy can be used to perform uncertainty quantification analysis in order to quantify errors and uncertainties that are inherent in any simulation and to quantify uncertainties in the output variables in the model that result from the uncertainties in the input variables. The smart proxy could potentially allow the user to explore the performance of the design, well beyond the CFD simulation time window. In other word, few hundred seconds of CFD simulation time can be used to construct a smart proxy, which can be used to explore the design performance of the unit after many hours of performance. The uniqueness of this approach is in:

1. Developing a unique engineering-based data preparation technology that optimizes the training of the neural networks. This innovative technique incorporates supervised fuzzy cluster analysis to:
    a. Identify the most influential parameters for the training process, and
    b. Identify the optimum partitioning of the data for training, calibration and validation.
2. Using an "ensemble-based" approach to building the smart proxy, taking advantage of multiple neural networks and intelligent agents to accomplish the objectives of the project.

1.1 STRUCTURE OF THE WORK

The research and development concentrating on the CFD Smart Proxy modeling will be presented in multiple reports. Each report will concentrate on a major portion of the research work and accomplishments that are useful to the general research community. The report presented in this document is regarding the proof of concept of using the Smart Proxy technology for replicating a CFD simulation model. This report includes four chapters. In chapter one (this chapter), the problem was defined, and the final objective of the research was articulated.

In chapter two, a brief definition of multiphase flow and its governing equations are provided to lay the groundwork for understanding the engineering and scientific details associate with the CFD model being studies. Also, the literature about the use of AI and Machine Learning related to fluid dynamics problems is reviewed.

Chapter three discusses the methodology and the machine learning method which is used in this research. The artificial neural network with all the required information is introduced in this chapter. The network architecture with all input and output system are discussed.

Results and discussions are presented in chapter four. The conclusions and recommendations for the next phase of the research are presented in chapter five

2. BACKGROUND

This section of the report is dedicated to providing some background information on MFiX multiphase flow software and machine learning.

2.1 MFIX

Multiphase flows, both reacting and non-reacting, are part of many processes in power generation and chemical processing industries. As expressed earlier, CFD is a valuable tool in design and optimization of processes and reactors used in these industries. NETL has been in the forefront of developing CFD modeling tools that can help engineers and designers in improving the performance of processes such as gasification, chemical looping. The MFiX (Multiphase Flow with Interphase eXchanges) suite of CFD software [10] is an open-source, general purpose multiphase CFD software suitable for modeling the hydrodynamics, along with heat transfer and chemical reaction for a wide spectrum of flow conditions (dilute to dense). Multiphase flows can be modeled either in a continuum (Eulerian) framework or a Lagrangian framework. The two frameworks can be summarized as follows:

- Continuum (Eulerian): Both solid phase and gas phase are treated as interpenetrating continua (Two Fluid Model, TFM). Multiple solid phases can be used to describe multiple solid particles of different sizes and properties (Multi Fluid Model, MFM). Continuum approach is computationally less intensive, but it cannot capture all the flow complexities, especially in multiphase flow where interaction between particles plays a major role [11].
- Discrete Particle (Lagrangian): Track each particle in the fluid by using Newton's Law of motion. This method is more straightforward to apply, even in multiphase flow, but the computational costs is high [11].

There are several approaches to modeling multiphase gas-solid flows. Depending on the application, either the gas phase or the solid phase or both phases can be modeled in Eulerian or Lagrangian framework [11] [12] [13]. Table 2-1 shows the different modeling approaches to gas-solid multiphase flow modeling.

In the present work, the MFiX-TFM is used to model a rectangular 3D fluidized bed. MFIX-TFM, which is based on kinetic theory of granular flow (KTGF) models both the gas phase and particulate phase as interpenetrating continuous phases. The governing equations employed for the conservation of mass and momentum for each phase (m = g for gas phase and m = s for solid phase) are

$$\frac{\partial}{\partial t}(\varepsilon_m\rho_m) + \nabla.(\varepsilon_m\rho_m\vec{v}_m) = \sum_{\substack{n=1\\ n\neq m}}^{N_m} R_{mn} \tag{2-1}$$

$$\frac{\partial}{\partial t}(\varepsilon_m\rho_m\vec{v}_m) + \nabla.(\varepsilon_m\rho_m\vec{v}_m\vec{v}_m) = \nabla.\left(\bar{\bar{S}}_m\right) + \varepsilon_m\rho_m\vec{g} - \sum_{\substack{n=1\\ n\neq m}}^{M} I_{mn} \tag{2-2}$$

Where

ε_m is the phase volume fraction

ρ_m is the phase density

$\vec{v}_m$ is the phase velocity vector

R_{mn} is mass transfer between phases

$\bar{\bar{S}}_m$ is the phase stress tensor

I_{mn} is the interaction force representing the momentum transfer between the phases

The closure terms for the solid phases are obtained through kinetic theory of granular flow. Detailed information on the constitutive relationships used to model momentum exchange between the phases along with the solid stress model incorporated in MFiX-TFM can be obtained from MFiX online documentations [14] [15]. Equations 2-1 and 2-2 form a system of nonlinear partial differential equations. An iterative algorithm is used in MFiX to solve this system of PDEs. Figure 2-1 illustrates the solution sequences used in MFiX for solving the equations 2-1 and 2-2. As it is discussed in the next section, it is crucial to follow the same sequence in constructing the smart proxy.

Table 2-1 Multiphase flow modeling approaches [11]

	Name	*Gas Phase*	*Solid Phase*	*Coupling*	*Scale*
1	*Discrete bubble model*	*Lagrangian*	*Eulerian*	*Drag Closure for bubbles*	*10 m*
2	*Two Fluid Model*	*Eulerian*	*Eulerian*	*Gas-Solid drag closure*	*1 m*
3	*Unresolved Discrete particle model*	*Eulerian*	*Lagrangian*	*Gas-particle drag closure*	*0.1 m*
4	*Resolved Discrete particle model*	*Eulerian*	*Lagrangian*	*Boundary condition at particle surface*	*0.01 m*
5	*Molecular Dynamics*	*Lagrangian*	*Lagrangian*	*Elastic collisions at particle surface*	*<0.001 m*

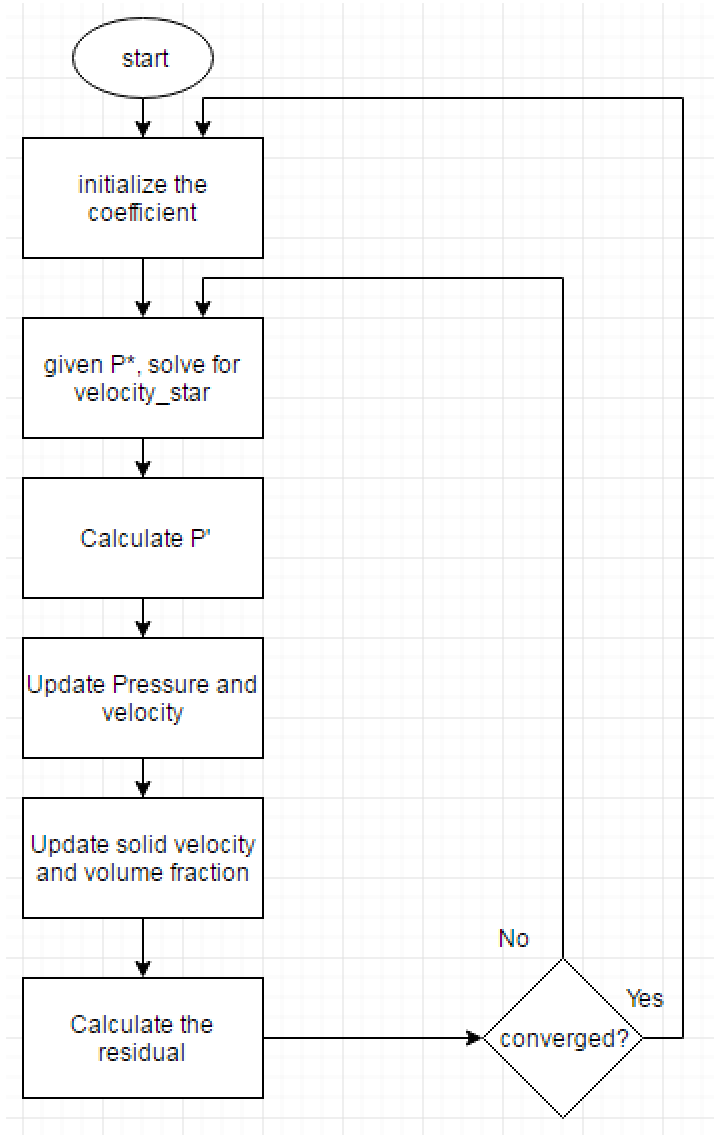


Figure 2-1 MFiX solution algorithm

2.2 MACHINE LEARNING

Based on the definition presented by Arthur Samuel [16], *"Machine learning is a field of study that gives computers the ability to learn without being explicitly programmed."* Machine learning is a process through which computer will learn from data to find a possible pattern in the data set. This process encompasses three main components:

- Learning algorithm
- Data
- Patterns in the data

If these three components are present, a successful learning process can be achieved based on the capability of the learning algorithm. There are two major type of Machine Learning: supervise learning and unsupervised learning [17].

2.2.1 Artificial Neural Network

One of the popular machine learning processes is Artificial Neural Network (ANN). The idea of ANN came from the neurons of the brain and the way they are communicating with each other to solve a problem. Each artificial neural network consists of an input layer, one or more hidden layers, and an output layer. The number of neurons (processing elements) in the output and the input layers are chosen based on the nature of the problem being solved and the properties which are going to be predicted. Figure 2-2 shows a typical ANN with three input neurons and two output neurons. ANN has one or more hidden layers and each layer has a specific number of neurons [18]. In order to have a well-trained network, proper parameters should be introduced to the network. If improper data are used to train the network there is no guarantee to have a well-trained network that lead to correct predictions, in other words, *"Garbage in, Garbage out."* In the upcoming sections of this report, a smart way of selecting parameters will be introduced.

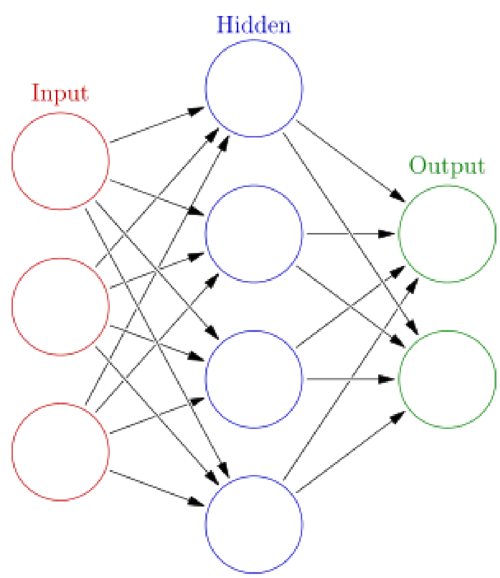


Figure 2-2 Artificial Neural Network schematic

The number of hidden layers and the neurons in each hidden layer depends on the complexity of the problem, number of parameters, and number of records. Experience also plays an important role in this decision making. But generally, there is no solid rule for them. As a rule of thumb, the number of neurons in the first hidden layer shouldn't be less than the number of input parameters.

2.2.1.1 Objective Function

Regardless of the learning method, each machine learning process needs an optimization procedure that helps the process reduce the error as much as possible. The very common and simple objective function in supervised learning is the summation of all the differences between predicted values by the learning method and the actual values of the output. Since summation of positive and negative errors can reduce the size of the overall error, the objective function is defined as the square of the difference between actual and predicted values [18], as shown by equation 2-3.

$$J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(y_{actual} - y_{predicted}\right)^2$$

2-3

During the learning process, the learning algorithm tries to assign different weights to each of the connection between neurons in Figure 2-2, in a way that the global error of the objective function becomes minimum. Also, a blind calibration is done simultaneously to stop the learning process. We will discuss the training set, calibration set, and validation set in more depth in sections 3.2.4.

2.3 PREVIOUS WORK

The idea of using Artificial Intelligence in petroleum engineering was first introduced by Mohaghegh and Ameri [19]. They took advantage of ANN for predicting the permeability of the formation based on geological well logs. Mohaghegh and Ameri [19] showed that neural network is capable of making the task of permeability determination automated rather than doing it over and over by log analyst. They also stated that neural network can handle far more complex tasks. Mohaghegh et al. [20] used ANN for predicting gas storage well performance after hydraulic fracture his later investigations.

Alizadehdakhel et al. [21] successfully used ANN to predict the pressure loss of a two-phase flow in the 2-cm diameter tube. Gas and liquid velocities and the pressure drop along the pipe were the three input parameters to ANN, with average pressure drop being the output. They utilized 8 different networks with different number of neurons to find out the optimum number neurons. Mean Squared Error and R-square were used as a criterion to pick the best network design. They also obtained the most efficient transfer function between Log-Sigmoid, Hyperbolic-Tangent Sigmoid, and linear.

Shahkarami et al. [9] used ANN to model the pressure and saturation distribution in a reservoir which was used for CO_2 sequestration. This problem required a large number of time steps for simulation of CO_2 injection and storage using a commercial software. They ran 10 different cases in CMG (commercial reservoir simulator) and then the results were used as input for ANN. The output of the ANN was pressure distribution, water saturation, and CO_2 mole fraction. 80% of the data coming from the CMG simulation runs were used to train the network while 10% were used for the calibration. The remaining 10% of data was used for validation process. They have shown that ANN can be used as a powerful tool for multiphase flow simulation in oil and gas industry.

Esmaili et al. [22] incorporated a newly developed AI-based reservoir modeling technology known as Data-Driven Reservoir Modeling [23] in order to model fluid flow in shale reservoirs using detail well logs, completion, and production data. By understanding the behavior of the shale reservoir, conducting the hydraulic fracture could be much easier. Moreover, this method has the ability to perform the history matching on the production data. Kalantari-Dehghani et al. [24] coupled numerical reservoir simulator with AI methods to develop a shale proxy model that is able to regenerate numerical simulation results in just a few seconds. They introduced three different well-based tier systems to achieve a comprehensive input data for the ANN. In another work, Kalantari-Dehghani et al. [25] showed that data-driven proxy models at the hydraulic fracture cluster level could be used separately as a reservoir simulator especially in low permeability reservoir such as shale which has a nonlinear behavior.

3. METHODS

In this section of the report, the solution methodology both for MFiX CFD simulations and neural network training will be discussed in detail. Additionally, the steps needed to create input to the neural network, which is the most important step of communicating with the learning algorithms is discussed.

3.1 CFD SIMULATION SETUP

A schematic of the rectangular fluidized bed, used in this study is shown in Figure 3-1. The fluidized bed, which is 0.12 x 0.72 x 0.12 m in X, Y and Z directions has an initial bed height of 0.12 m, and initial bed voidage of 0.42. The bed material has a density of 2000 kg/m^3 and a diameter of 400 µm. Air velocity is set to 0.6 m/s and is uniformly distributed across the inlet. The spatial grid resolution is 4.4 mm (11 particle diameters) in all directions and is based on a grid resolution study that was carried out for four different grid levels, as shown in Table 3-1. This is in line with Fullmer and Hrenya [5] work that a grid spacing as small as 10 particle diameters is needed for numerical accuracy. Simulation of the fluidized bed is carried out for 30 seconds. MFiX outputs all relevant information such as gas and solid velocities, voidage and pressure field for the entire domain. Figure 3-2 shows the instantaneous voidage contours at a point in time, during the simulation. The output data from MFiX is used as the input and output data for the training, calibration and validation of the Artificial Neural network (ANN). Since MFiX reports the results based on the location of each control volume in the grid, the order and the exact location of each grid becomes extremely important for ANN. In MFiX, each control volume is represented by its X, Y and Z location (I, J and K indices). An additional single index, IJK, is defined in MFiX that is unique to each control volume (which is defined by it I, J and K indices). Figure 3-3 shows the order at which the IJK index is used. MFiX outputs all the data following the IJK index order. ParaView, which is an open-source visualization software is used to extract the required data from MFiX files at each time step, for use in ANN training.

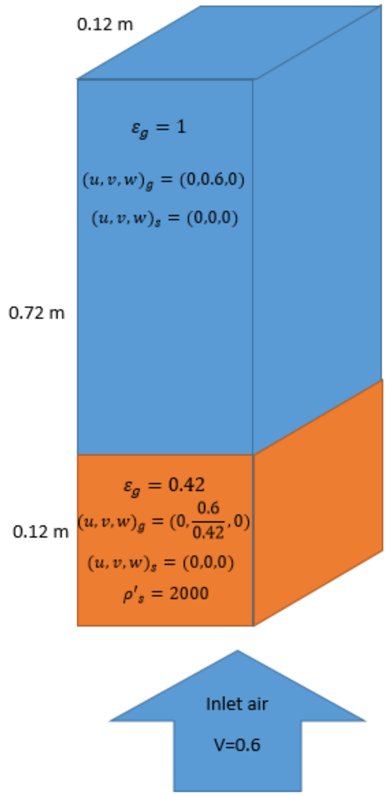


Figure 3-1 Geometry and initial condition of the problem

Table 3-1 Different grid size and the number of cells

Grid Classification	*Cell size*	*No. of Cells*	*No. of Nodes*
Coarse	*8*48*8 (15 mm)*	*3,072*	*3,969*
Medium	*12*72*12 (10 mm)*	*10,368*	*12,337*
Fine	*18*108*18 (6.6 mm)*	*34,992*	*39,349*
Very Fine	***27*162*27 (4.4 mm)***	***118,098***	***127,792***

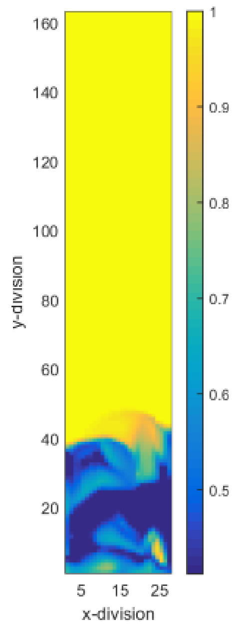


Figure 3-2 Snap shot of voidage contour predicted by MFiX

3.2 ARTIFICIAL NEURAL NETWORK SETUP

Table 3-2 shows the 9 MFiX output parameters used for ANN training. Once the output files of MFiX are converted to *.csv file they are reorganized to serve as the input for the ANN. Every time-step has one *.csv file that contains 9 columns and 118,098 rows. Each column represents one output parameter and each row corresponds to one cell. The input to the ANN is all the data at time-step *t* while the output will be one or more parameters at time-step *t+1*. In this approach, the network will learn what the output should be, given a set of input data. When the learning process is completed, the deployment process (prediction) will be performed.

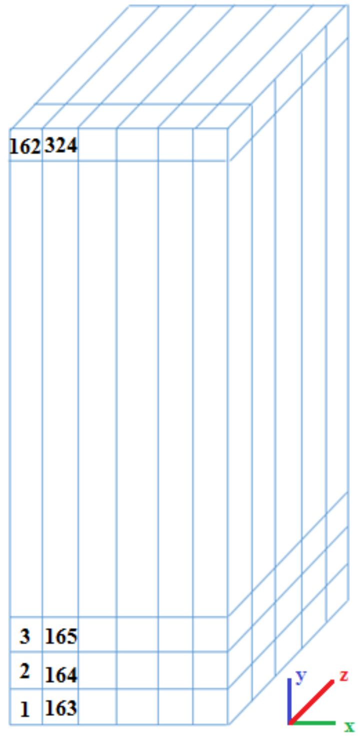


Figure 3-3 MFiX numbering order

Table 3-2 MFiX output variables used in ANN training

Symbol	*Description*
ε_g	*Gas volume fraction*
P	*Gas Pressure*
P_s	*Solid Pressure*
u_g	*Velocity of gas in x direction*
v_g	*Velocity of gas in y direction*
w_g	*Velocity of gas in z direction*
u_s	*Velocity of solid in x direction*
v_s	*Velocity of solid in y direction*
w_s	*Velocity of solid in z direction*

3.2.1 Tier System

In order to communicate all the required information with the ANN so that it can have a reasonable understanding of the state of the process, and to learn in an effective manner, a tier system was developed. Each cell is in contact with 26 surrounding cells; 6 of them have the surface contact with the original cell, 12 have line contact with the original cell, and the remaining 8 have point contact with the original cell.

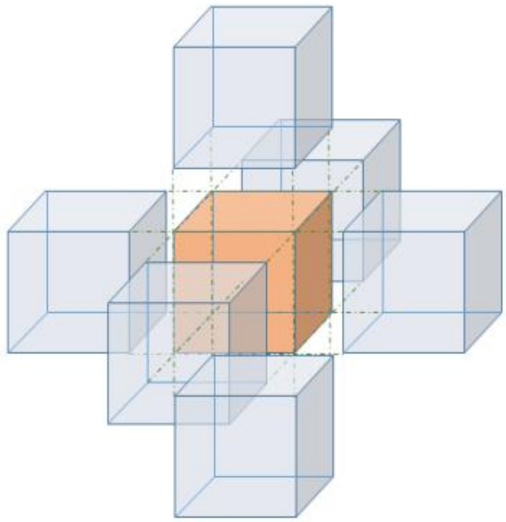


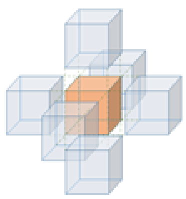
Figure 3-4 The tier system with the 6 cell in surface contact with the focal cell

Like any numerical method, the value of each variable in each cell is correlated to the variable value in the surrounding blocks. With that idea in mind, the ANN will not only learn from the 9 parameters (Table 3-2) in a cell, it will also learn from the surrounding cells which are called *"Tier"* cells. Each cell has several tiers. Tier 1, tier 2 and tier 3 are the surrounding cells that are in surface contact, line contact and point contact respectively. Depending on the complexity of the problem, different tier system will be used as input to ANN. Figure 3-4 shows a tier 1 structure, where the main (focal) cell is surrounded by its 6 neighboring cells. For this case, the 9 parameters of the original cell and 9 parameters of the tier 1 cells make 63 parameters ((6+1) * 9), which are the input for the ANN. Depending on the complexity of the problem and spatial and temporal correlations between different tiers and the center cell more or less input parameters might be required.

3.2.2 Input Matrix

It is not enough to consider only the values of each parameter in a focal cell and the related tiers in the input matrix, but for the network to learn the behavior of the process and perform pattern recognition, the location of each cell in the geometry is also crucial. Adding the location as an input helps the system understand the spatial correlation between different parameters, as well. On the other hand, walls (boundary conditions) have important impact on the flow pattern, therefore, the location of walls with respect to the focal cell should also be somehow included into the inputs for the training of the ANN. To accommodate these ideas, six different distances to the wall confinements (top, bottom, east, west, north, and south) are considered in order to define the exact location of each focal cell and the parameters associated with each cell. By adding these 6 distances to the previous 63 parameters (9 parameters of 7 cells - the focal cell plus six tier one cells), the total number of parameters used as input becomes 69, as shown in Figure 3-5. So, the dimension of input matrix is 69 by 118,098 (i.e., number of parameters multiply by the number of cells).

$Location\ of\ Cell_{ijk}$					
D_1	D_2	D_3	D_4	D_5	D_6



$Cell_{ijk}\ and\ 6\ tiers$								
ε_g	P	P_s	u_g	v_g	w_g	u_s	v_s	w_s

Figure 3-5 69 parameters of ANN

3.2.3 Neural Network Architecture

Each artificial neural network consists of an input layer, one or more hidden layers, and an output layer. The input and output parameters are chosen based on the nature of the problem and the property which is going to be predicted. In the last section, it was described how the number of input parameters were selected to be 69. The output of the ANN could be one or more parameters. There will be different scenarios to compare different ANN with different number of output parameters. There is no clear guideline on how many hidden layers and neurons are required at each layer for a given problem. A rule of thumb indicates that the number of neurons in the first hidden layer shouldn't be less than the number of input parameters. For the first try, one hidden layer with 100 neurons is considered where 69 parameters are used as input and only one parameter is selected as output, as shown in Table 3-3. The network characteristics are shown in Table 3-4. Feed-forward back propagation method is used for the training. The transfer function for the hidden layer and the output layer was chosen to be TANSIG, as shown in Figure 3-6.

Table 3-3 Important numbers in Neural Network Model

Number of Inputs	*69*
Number of hidden layers	*1*
Number of Hidden Neurons	*100*
Number of records	*118,098*
Number of Output	*1*

Table 3-4 Neural Network characteristics

Network Type	*Feed-forward Back propagation*
Training Function	*Levenberg-Marquardt*
Adaption Learning Function	*LEARNGDM*
Performance Function	*MSE*
Transfer Function	*TANSIG*

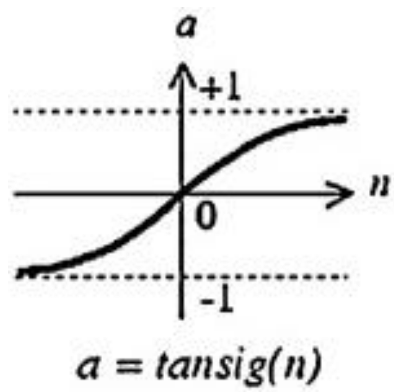


Figure 3-6 Nueral Network transfer function (TANSIG)

3.2.4 Data Partitioning

A good ANN is a model that learns the pattern in the given data-set while it is able to predict the behavior of a new given dataset, this model is called "Just Right". If the ANN does not learn the pattern in the data very well the model is called "Under-fit". If the ANN learns the pattern of the data very well with a very small error but it is not able to predict the behavior of a new given data-set the model is called "Over-fit". Under-fitting occurs for so many reasons such as lack of information (the model should have more parameters and more examples). Overfitting occurs when the network learns to mimic almost all the data points exactly but when it comes to the prediction, the model performs poorly for a new given data, in other words the model memorizes all the data points. Figure 3-7 shows these 3 states of training.

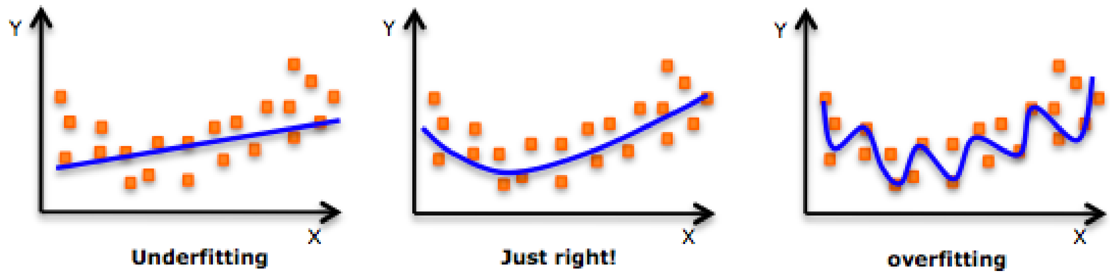


Figure 3-7 Underfitting and overfitting of the data

To overcome the overfitting problem only a portion of the data is used to train the network and the rest of data is kept outside of the training as a criteria to stop the training process when the model is "Just Right". The remaining data points which the model has not seen in the training process, are further divided into two sub groups; calibration and validation.

Training is an iterative process where in each iteration the optimization algorithm tries to move toward the lower error. Calibration data-set is used while the training is being carried out. The error in both training data set and calibration dataset usually decreases at the beginning of the training process, however somewhere along the training process, the error in calibration data set stops decreasing while the error continues to decrease in the training data set. The model at this point is usually the best model because it has provided the lowest possible error for the calibration data set (blind data set) and while it has an acceptable error for the training data set.

The validation data set is used upon the completion of the training process when the best ANN is achieved. Although both calibration and validation data sets are blind but having an ANN model with a low calibration error does not mean that the ANN is a good predictor (because the best model is already picked when the calibration error is minimum) unless the ANN error in validation data set is also acceptable. The percentage of the data partitioning used for the preliminary study of this project is shown in Table 3-5. It is important to mention that this partitioning is the preliminary one and a deeper study will be conducted on the percentage of the data as it will be described in the upcoming sections of this report.

Table 3-5 Original data partitioning

Data	*Training*	*Calibration*	*Validation*
Percentage of data (%)	*70*	*15*	*15*

3.3 SPATIO-TEMPORAL DATABASE

MFiX simulation results of the rectangular fluidized bed described in section 3.1 is used as the training, calibration and validation data to build a Spatio-Temporal data base for the construction of fluidized bed smart proxy (Figure 3-8).

Figure 3-8 Spatio-Temporal Database and optimized database

Initially, the Spatio-Temporal database was created based on the data from one single time step. This database included the location and the properties (listed in Table 3-2) of each cell and the properties of tiers (totally 69 parameters). Later in this report (section 3.4.5), additional time steps will be added to the database to have a general model for all the time steps. The elapsed time should be also added to the data base as a new parameter, therefore the total number of parameters becomes 70. The Spatio-Temporal database treats each cell as a separate record, so the model has 118,098 records which is equal to the number of cells. This database is then optimized for a more efficient storage and access.

3.4 SOLUTION SCENARIOS

Different scenarios are considered in order to reach the final goal of this part of the project, which is to illustrate the feasibility of constructing a smart proxy for a fluidized bed, based on data generated from CFD. The term *"Different scenarios*" refers to having different input and output structures and also using different time-steps for the training of ANN, while the training technique is kept the same.

Each scenario has two parts, first is the training process and second is the deployment process. A pair of time-steps is used in the training process. The training process stops based on user defined criteria. This criterion could be the total number of iteration, the total time of training, or the number of calibration failure or as is the case in this work, the termination criteria is a combination of all of the above. The learning algorithm is such that the network learns more after each iteration, so to prevent overfitting or memorization, calibration error is always checked. If the calibration error increases for a predefined number of iterations, the training stops. Most of the time, calibration is the criterion which makes the training stop.

As mentioned in the section 3.2.2, at a given time step, 69 parameters are used as input to ANN during the training stage. Additionally, there are also 9 CFD parameters, Table 3-2, at the next time step in CFD, which should be included as additional input parameters to ANN during the training stage. We can train the ANN with either one or multiple CFD parameters at the next time step. It's important to reiterate that in the training stage, CFD output results for the variable that ANN is being trained for, are input to ANN, along with the static parameters (depicted in Figure 3-9).

Figure 3-9 Training stage flow chart

The trained network is then ready for the deployment stage, where data at a given time step is used as input and the trained network provides prediction at the next time step (Figure 3-10). The input of the ANN for each deployment could come from the CFD directly or from the ANN itself. Cascading and non-cascading deployment are defined based on what type of input is used for the network and it will be discussed in detail in the following sections.

Figure 3-10 Deployment stage flow chart

As it was stated earlier, this phase of the current research aims to show the feasibility of constructing a smart proxy, based on CFD results for a fluidized bed. As such, the scenarios outlined below show the systematic steps, which has been taken from the least complex scenario to the more complex scenarios, where more input parameters are used for training of the ANN. The scenarios followed in order of complexity are:

- ANN training with 69 input parameters at a given time step during the initial stage of fluidization and during the later time, section 3.4.1.
- ANN training with 69 input parameters at a given time step, using a cascading and non-cascading process, section 3.4.2.

- ANN training with 69 input parameters at a given time step and one output parameter compare to multiple output parameters, section 3.4.3.
- ANN training with 69 input parameters at a given time step with explicit temporal approach compared to implicit temporal approach, section 3.4.4.
- ANN training at multiple time steps which requires 70 input parameters, section 3.4.5.
- Reducing the size of the 70 input parameter data set, used for ANN training, section 3.4.6.

3.4.1 Early Time versus Late Time

Gas-solid flow in a fluidized bed is highly unsteady and chaotic. The gas-solid flow initially behaves like a slug flow, before instabilities set in and fluidization begins at a later time, as shown in Figure 3-11. As time goes by, changes in flow regime occur and flow becomes more chaotic and heterogeneous. It is therefore necessary to investigate how well an ANN can be trained, when the degree of heterogeneity in the flow increases. An ANN is trained based on flow encountered at the early stage of fluidized bed operation, Figure 3-12, and an ANN is trained based on the flow conditions at a later time, Figure 3-13. In both cases, the 69 inputs come from time-step *t* and the CFD output is from time-step *t+1* in Figure 3-12 and time step *t+2* in Figure 3-13. The larger time step used in Figure 3-13 is for expediting the training process. The CFD output could be one parameter (as is the case in Figure 3-14) or multiple parameters. Each time step used for training represents 1 millisecond of simulation time.

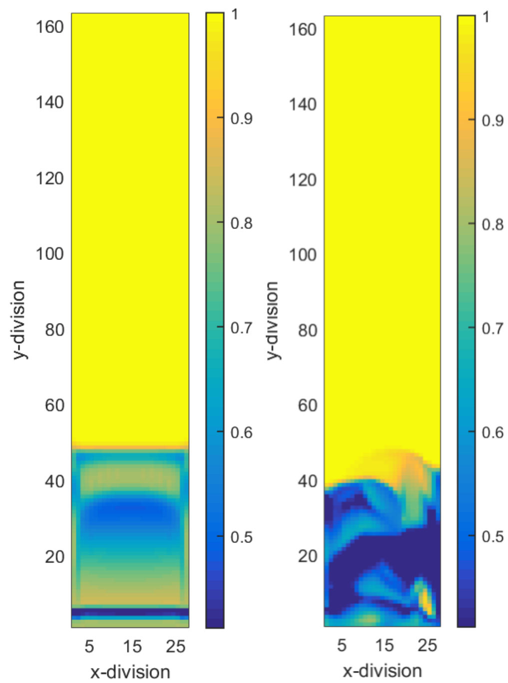


Figure 3-11 Gas volume fraction distribution initially in the bed (left) and later when it is fully fluidized (right)

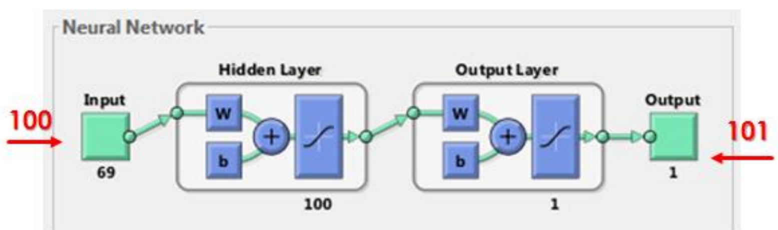


Figure 3-12 Input at time step 100 and CFD output at time step 101 used for training (early time, startup)

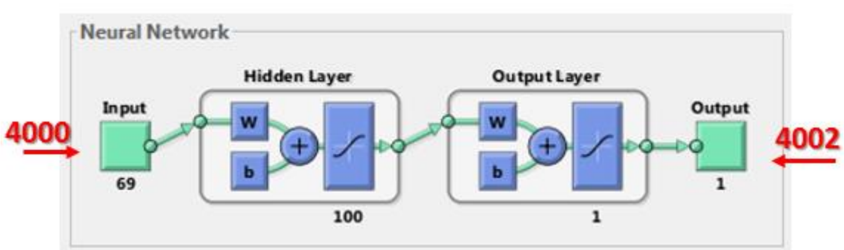


Figure 3-13 Input at time step 4000 and CFD output at time step 4002 used for training (later time, fully fluidized bed)

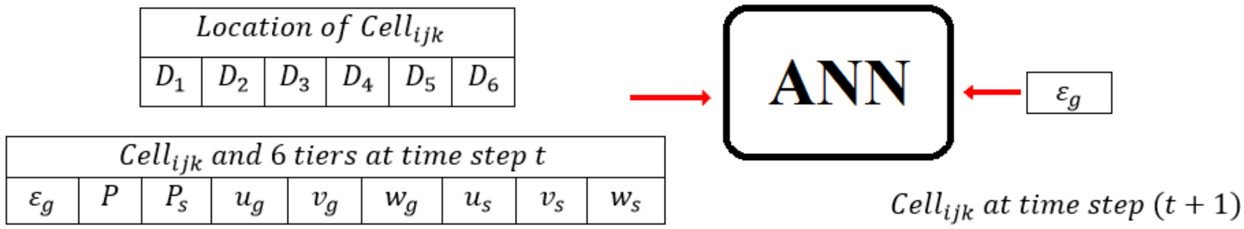


Figure 3-14 Input and CFD output parameters used for training

The purpose of this analysis is to show that the ANN is capable of capturing all the physics involved in different time-steps (different flow regimes). In the next section, complete results of this analysis will be presented and discussed in detail.

3.4.2 Cascading versus Non-cascading

Cascading and non-cascading refer to the source of input that is used for the deployment process. If the input comes directly from the CFD simulation model for each deployment stage, then the process is called "non-cascading", shown in Figure 3-15. If the input of the ANN for each deployment stage comes from the output of previous deployment, then the process is called "cascading", shown in Figure 3-16. Although the non-cascading deployment has little benefit, since it requires input from CFD solver at every time step, it should always be studied in order to confirm that the trained network is working properly. Eventually, every parameter should be predicted by cascading method, however, first non-cascading training should be performed. In the following sections, the results from both non-cascading and cascading deployment process will be shown for early and late time frames.

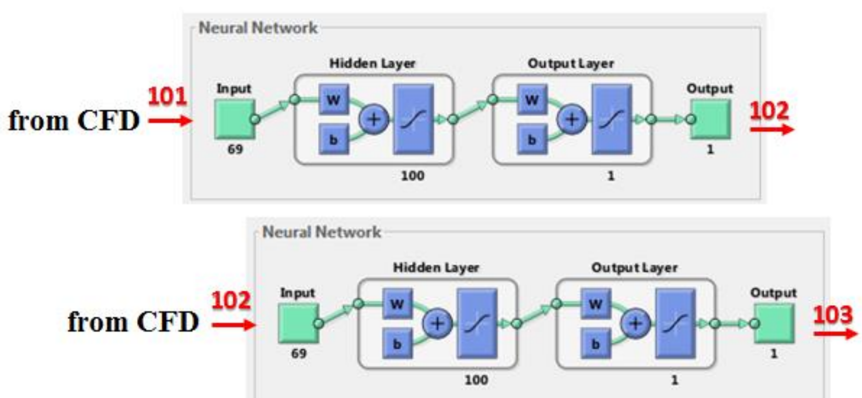


Figure 3-15 The process of non-cascading deployment

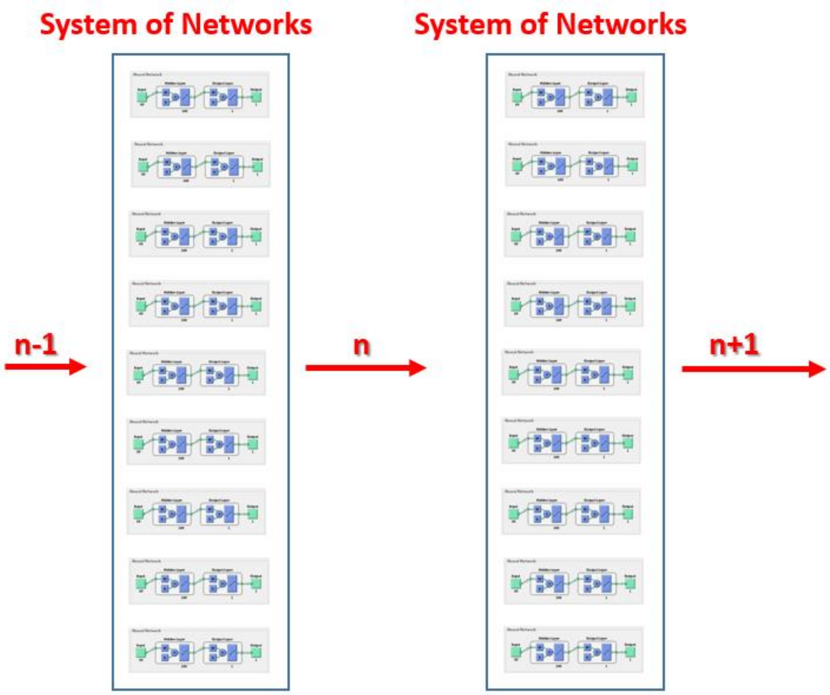


Figure 3-16 The process of cascading deployment

3.4.3 Single Output versus Multiple Output

As discussed earlier, ANN can have one output or have multiple outputs at the same time. Obviously, having multiple outputs simultaneously increases the training time, furthermore, the network has to fit multiple outputs with the same weights, so the network has less flexibility to learn from data. However, sometimes better results are obtained, especially if there are correlations and dependencies between the output parameters. This is a problem dependent issue that must be studied and decided upon. Figure 3-17 and Figure 3-18 show the input and output of the ANN when only one output and 3 outputs are used respectively.

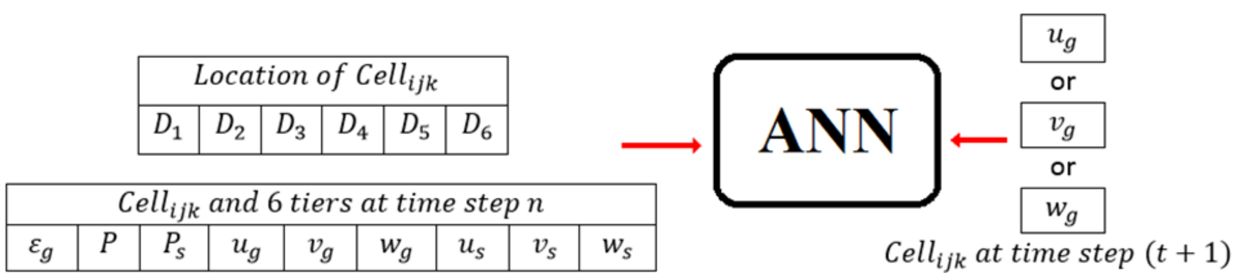


Figure 3-17 Traning with one output (one component of gas velocity)

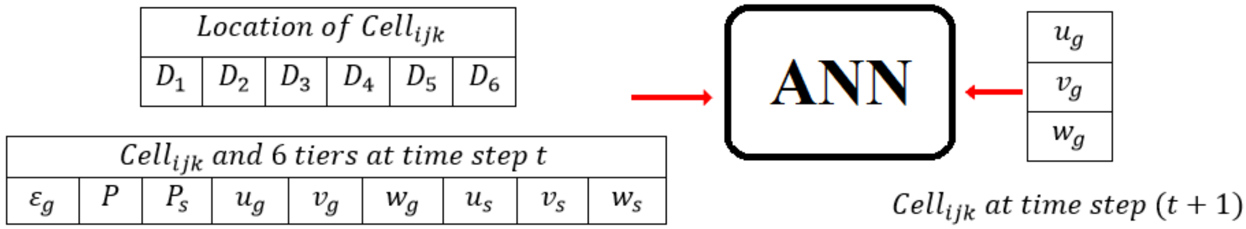


Figure 3-18 Traning with multiple outputs (3 components of gas velocity)

As another advantage of this approach, it should be stated that having multiple outputs at the same time would reduce the number of neural networks. As mentioned in the last section, there are total nine different ANN needed for cascading deployment, this number could be reduced to three if each network has three outputs at the same time. The result of this approach is also available in upcoming sections of this report.

3.4.4 Explicit versus Implicit

Regardless of the training scenario, the training process needs a pair of data; input and output (time-step *t* and time-step *t+1*). An explicit method is when all the input data come from time-step *t* and the output data come from time-step *t+1*, as shown in Figure 3-18. It is also possible to have the combination of data from time-step *t* and *t+1* as input and have time-step *t+1* as the output.

Implicit training is when the parameters that have been used as input from time-step *t+1* will not be used for output. Figure 3-19 shows one of examples of implicit training. The input consists of gas volume fraction, pressures, and gas velocity vector from time-step *t* in addition to solid velocity vector from time-step *t+1*. The output is gas velocity vector from time-step *t+1*.

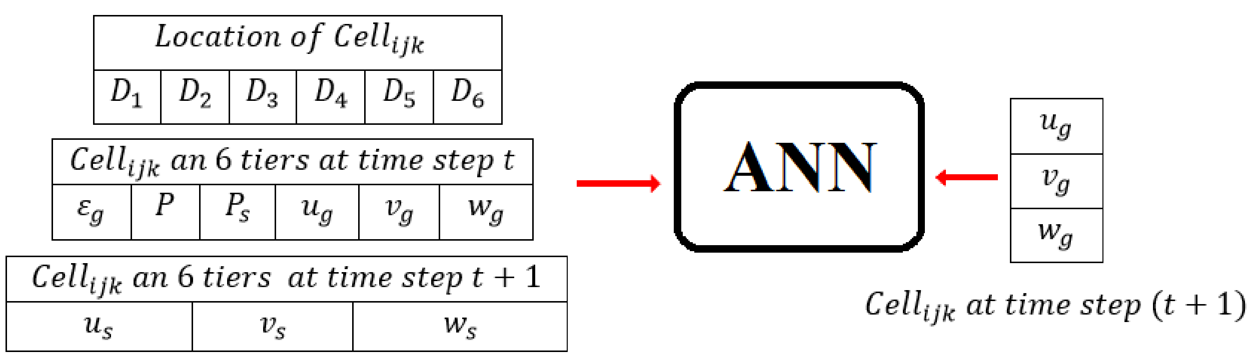


Figure 3-19 Traning implicitly with multiple outputs

This approach is very common in the numerical solution of PDE's, which increases the converging speed. Implicit approach is expected to have a lower error.

3.4.5 Training with Multiple Time-steps

The techniques outline so far uses a single time step for input and output of the ANN. This is illustrated by Figure 3-20. However, as discussed in section 3.4.1, the gas-solid flow undergoes a regime change, from a slugging flow at the beginning to a fluidized regime as time goes by. The ANN trained with the data from when the flow field is slugging does not have the predictive capability of capturing the flow dynamics, when the bed is fully fluidized. In order to train an ANN, which has a wider range of applicability, the input and output of the ANN must be trained on data from multiple time step, capturing many changes taking place in the flow.

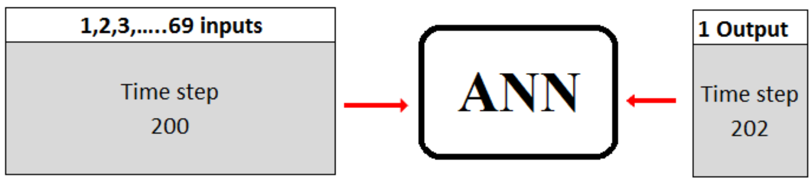


Figure 3-20 Input and output pair for the training with single time-step

Figure 3-21 shows the input and output pair for the training with three different time-steps, when the flow is slugging at first (time step of 200), then transitioning (time step 1000) and finaly fluidizing stage (time step 4000). Figure 3-22 shows the voidage contours in the CFD simulations at time steps 200 (0.2 sec elapsed time), 1000 (1 sec elapsed time) and 4000 (4 sec elapsed time).

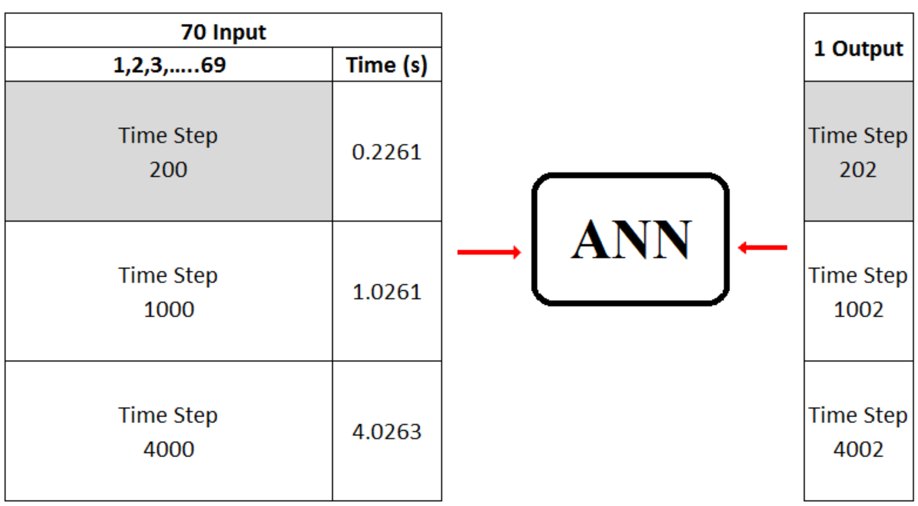


Figure 3-21 Input and output pair for the training

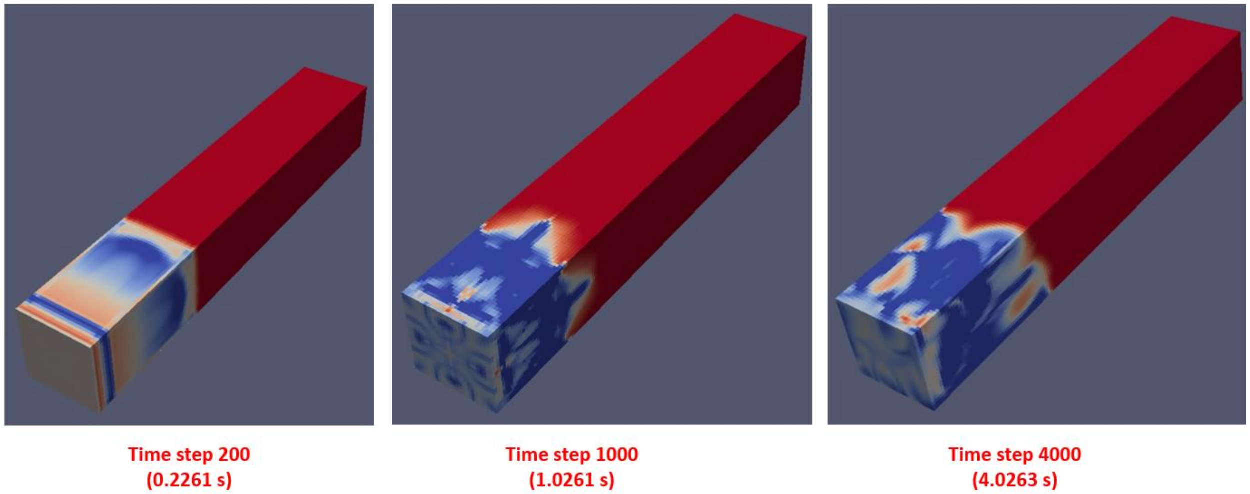


Figure 3-22 Three different time-steps with different flow characteristics

The quality of the ANN is characterized by the square root of mean square error, as defined by equ. 3-1.

$$\boldsymbol{RMSE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{n}}\left(\boldsymbol{y_{actual}} - \boldsymbol{y_{predicted}}\right)^{\mathbf{2}}} \qquad \mathbf{3\text{-}1}$$

3.4.6 Reducing the Size of the System

In this work, the input training data consists of 70 parameters (data columns) and 118,098 records (data rows) that produce a total number of 8,266,860 data points for a given time step. Reducing the data size will not only reduces the storage and data handling overhead, it also could reduce the computational cost.

3.4.7 Reducing Number of Records

One can remove cells (records) from the training data set randomly, or remove cells from the regions of fluidized bed, where the information is not as crucial to the main goal of the constructing an ANN for the bed. For example, the gas flow in the freeboard region of the fluidized bed does not affect the gas-solid behavior in the bed greatly, and therefore can be dropped from consideration. Figure 3-23 shows the voidage contours at time-step 4000 (4 seconds of elapsed time). It can be seen in Figure 3-23 that there are no solids present above the blue line in the freeboard and therefore, all the cells above the blue line can be removed from the training data set. This will reduce the size of the data set by more than one half, as seen in Figure 3-24.

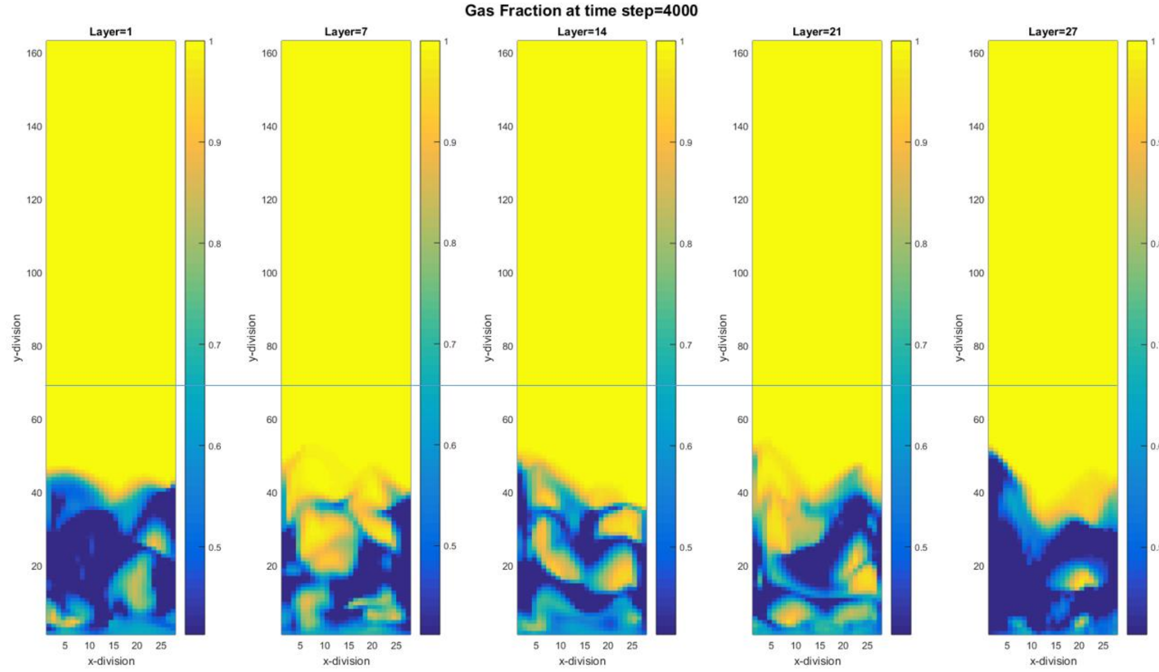


Figure 3-23 Gas volume fraction contours at time-step 4000

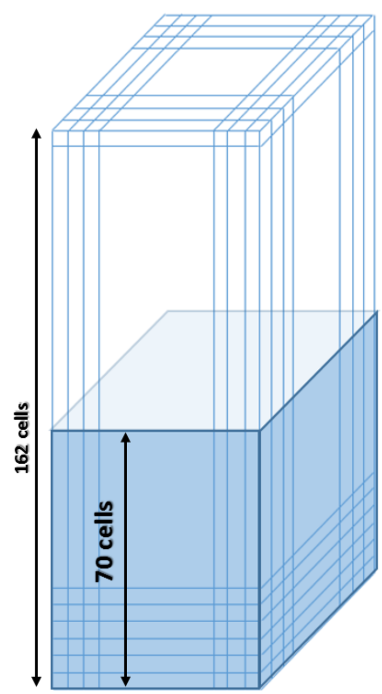


Figure 3-24 The key section of the fluidized bed used for training ANN

3.4.7.1 Reducing Number of Parameters (KPI)

In order to reduce the number of parameters, sensitivity analysis should be performed to quantify how sensitive the output is to various input parameters. This process is called identification of the Key Performance Indicator (KPI) and our effort is concentrated on finding the most influential parameters that impact the output more than the other parameters.

To perform KPI, all the weights associated with parameters should be obtained. Every parameter has several weights assigned to it to communicate with the hidden layer, as it is depicted in Figure 3-25. If all the weights assigned to one parameter ($w_{11}, w_{12}, \ldots$) are integrated to one value (w_1), that value will represent the total weight and show the priority of that particular parameter when it compares to all the other total weights. After obtaining all the total weights of the parameters, the tornado chart of each ANN could be plotted, and the key parameters could be determined.

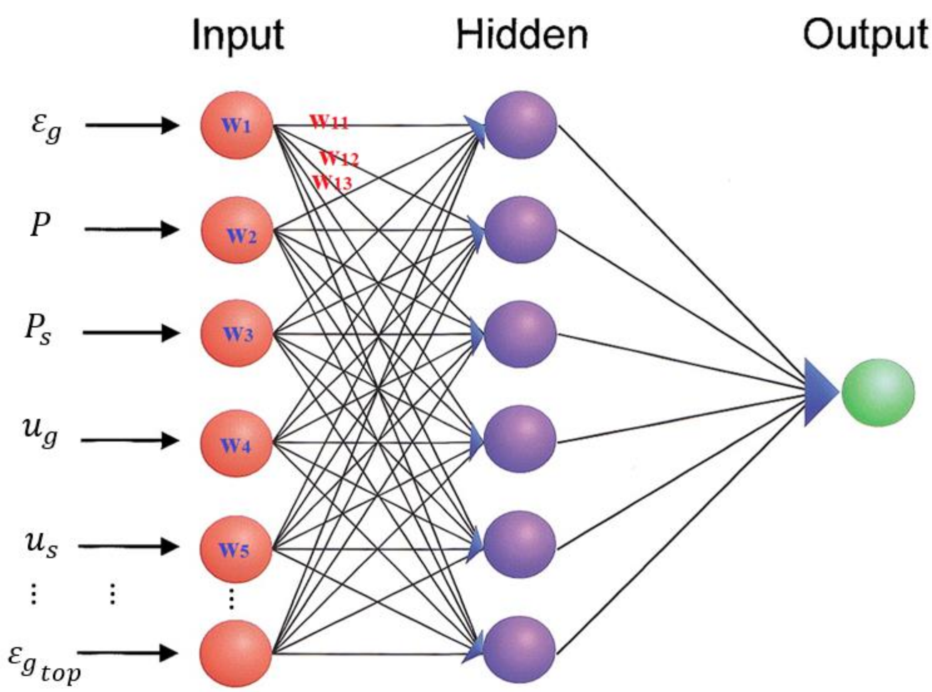


Figure 3-25 Network schematic with its weights

3.4.7.2 Changing the Data Partitioning

According to Table 3-5, 70% of the data is used for training, 15% of data is used for the calibration, and the other 15% of data is used for the validation. Since training of the network is the most computationally intensive part of the entire process, reducing the amount of training will accelerate the process of constructing an ANN.

3.4.7.3 Reducing Number of Records Using Smart Sampling

Smart sampling is a process through which the data used for training and calibration of the artificial neural network is not selected through a complete random process. During the smart sampling process all the available data (both input parameters and the output parameter) are scrutinized in detail to make sure that proper data is used for training and calibration purposes. For example, highly pronounced skewness of the distribution of the output parameter can bias the training process toward the portion of the data with much higher frequency. In such cases, while the skewness is preserved to a certain degree, the frequency of the data used for training and calibration is selected such that the shape of the distribution is communicated with the neural network while giving the learning algorithm a real chance of learning all aspects of the distributed data.

A wide spectrum of flow regime is encountered in a fluidized bed. At a given time, portion of the flow can be in a dense phase, while part of the flow is in the dilute phase and yet other part of the flow is in the emulation phase. Figure 3-26 shows the distribution of gas volume fraction at time step 4000. The distribution is bi-modal. Cells with voidage value of 0.4 represents cells that are at the solid packing limit (solid volume fraction of 0.4) and cells with voidage of 1.0 represent cells with no solid phase present (all air). Due to the strong bi-modal distribution of voidage, the learning will be biased towards the regions of the distribution, where the peeks reside. This is not desirable, since the region in between voidage of 0.4 and 1.0 represents the interface between bubbles and solid phase and the more dilute regions of the fluidized bed. To avoid this bias in training, the concept of smart sampling is introduced.

In the training process, the ANN is taught by the data. The data should provide even information in different ranges. If the data is not uniform and it emphasizes more on one part, the ANN will learn that parts very well and will not learn the other parts. Thus, the best practice is to have a uniform distribution rather than the bi-modal distribution. The simplest way to make the distribution uniform is to sample the data randomly. Figure 3-27 shows the distribution after smart sampling. This reduces the number of records from 118,098 to 25,827.

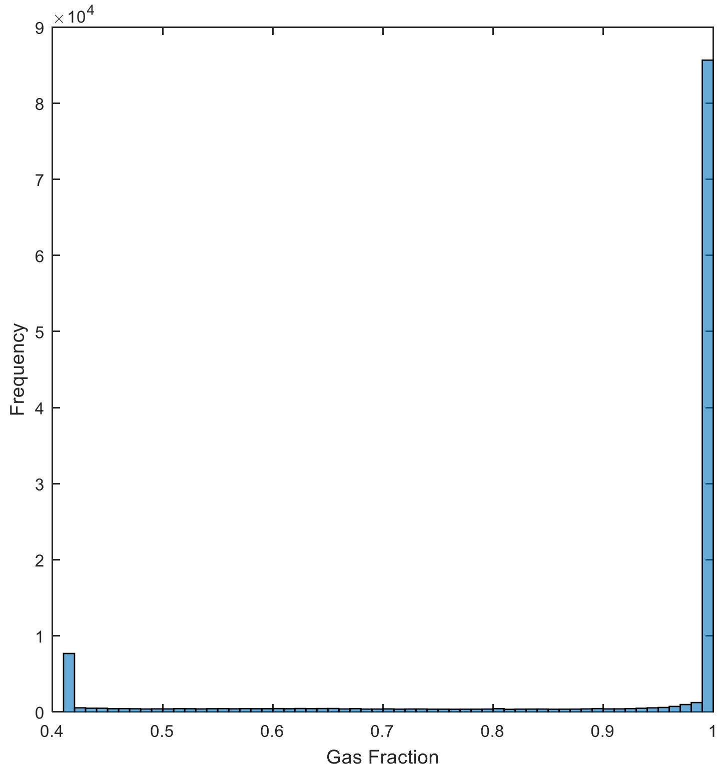


Figure 3-26 Distibution of gas volume fraction at time step 4000

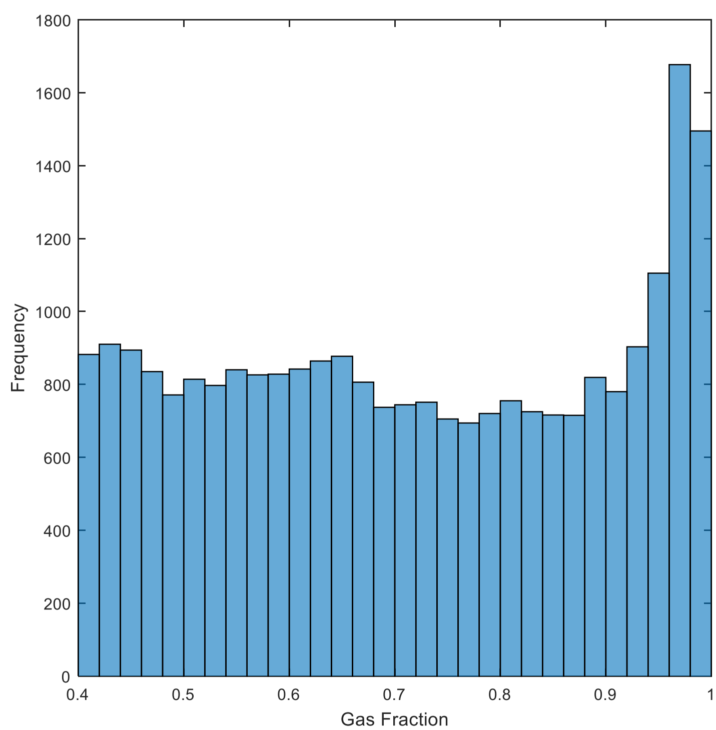


Figure 3-27 Distibution of gas volume fraction at time step 4000 after smart sampling

4. RESULTS AND DISCUSSIONS

Different scenarios were introduced in the previous section. In this section of the report, the results of all the scenarios are presented and discussed in detail. These results are coming from different approaches; early time or late time, single time-step or multiple time-steps for training, cascading or non-cascading deployment, single output or multiple outputs, and explicit or implicit method. Before proceeding with the results, there next section provides a short description of how the results are going to be presented.

4.1 PRESENTATION OF RESULT

To compare the CFD results with the smart proxy results, 5 vertical cross-sectional planes, 3 cm apart, are selected, where the contour plots for the various output parameters are presented. The locations of these vertical planes are shown in Figure 4-1. Each figure has three subplots, the left plot is the result of CFD simulation model which is coming directly from MFiX, the middle plot is the result of the smart proxy which is the output of ANN, and the right plot is the error distribution which is the difference between CFD and the smart proxy.

4.2 EARLY TIME-STEP, NON-CASCADING, SINGLE OUTPUT, EXPLICIT

The simplest case to consider is when one-time step selected as the input, along with one output. 9 separate ANN have been trained for all the 9 parameters. Time-steps 100 and 101 were used to train the system. After the training was completed, to deploy the model, all the time-steps from 101 all the way to 120 were used as input to the ANN and acceptable results were obtained. In the next sections, the results of gas volume fraction and gas pressure are provided for time-step 102 and remaining time-steps could be found in Appendix I. The results for the rest of the parameters are in Appendix II.

4.2.1 Gas Volume Fraction

Comparison between the CFD results and smart proxy for voidage at time step of 102 (0.102 sec elapsed time) are shown in Figure 4-2 through Figure 4-6. It can be seen that the smart proxy is able to replicate the MFIX simulation results very well, with less than 3% error especially around the inlet region.

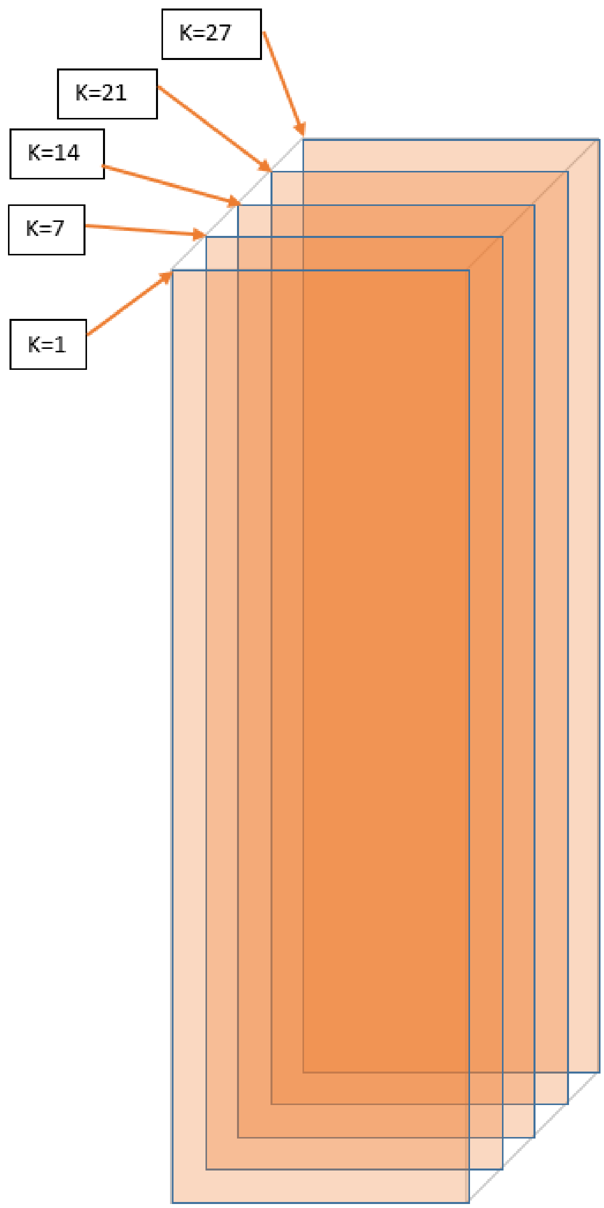


Figure 4-1 Cross-sectional planes, 3 cm apart, where results are presented

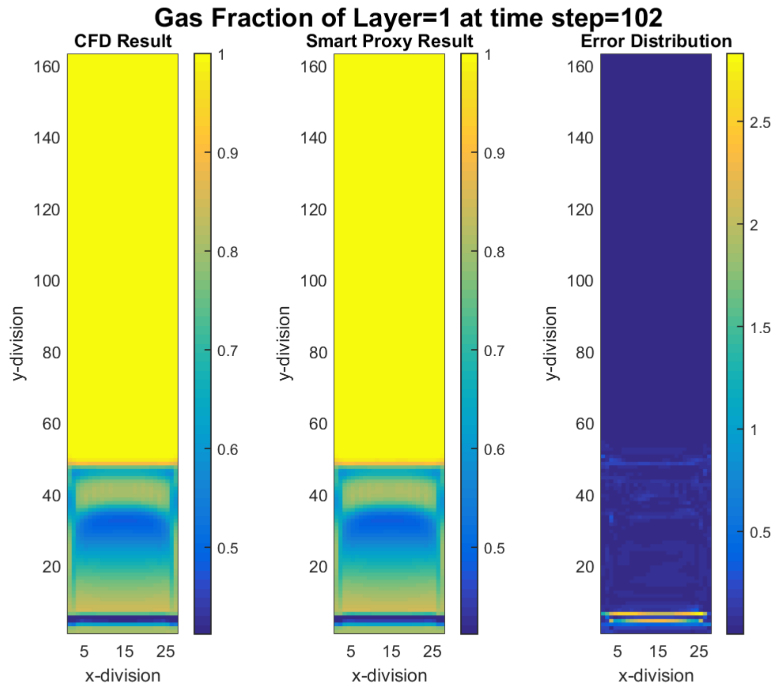


Figure 4-2 CFD and smart proxy results for gas volume fraction at K = 1 cross-sectional plane

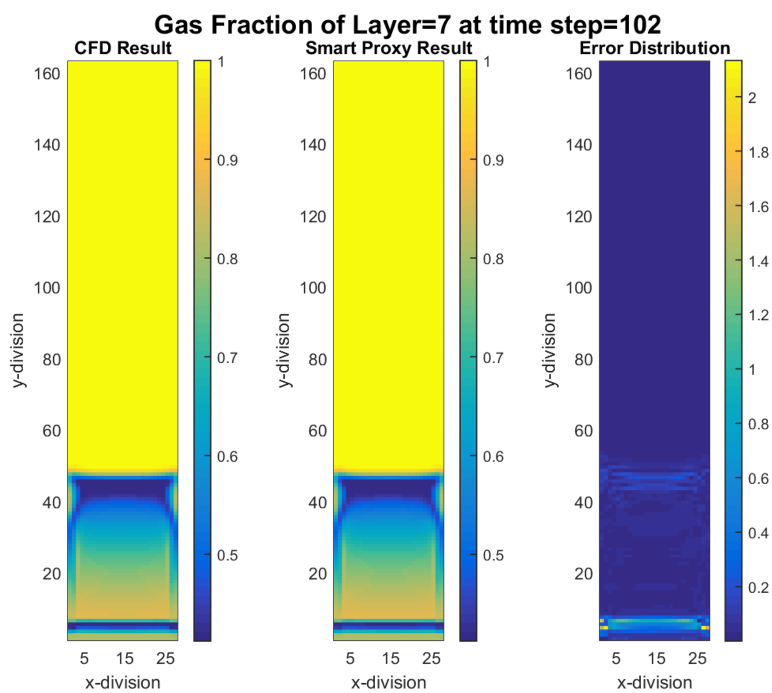


Figure 4-3 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane

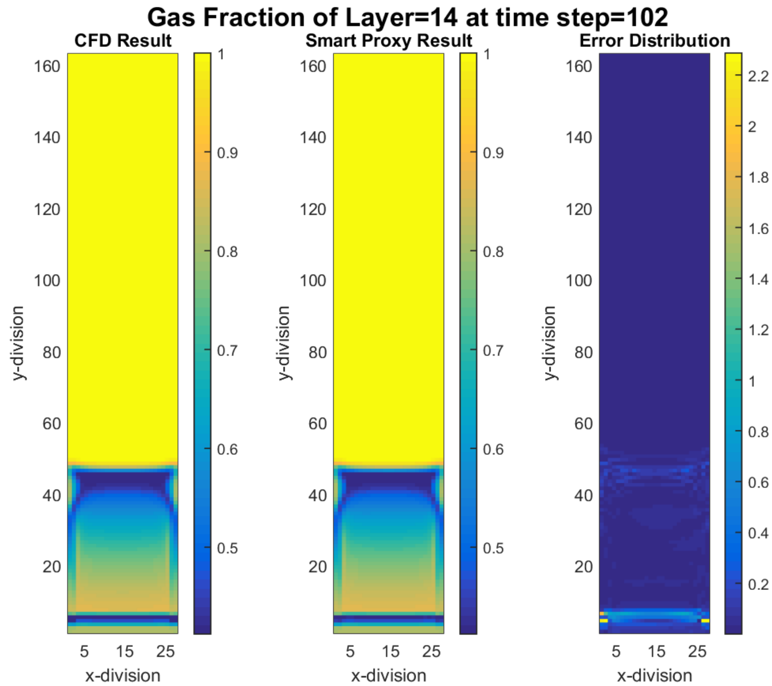


Figure 4-4 CFD and smart proxy results for gas volume fraction at K = 14 cross-sectional plane

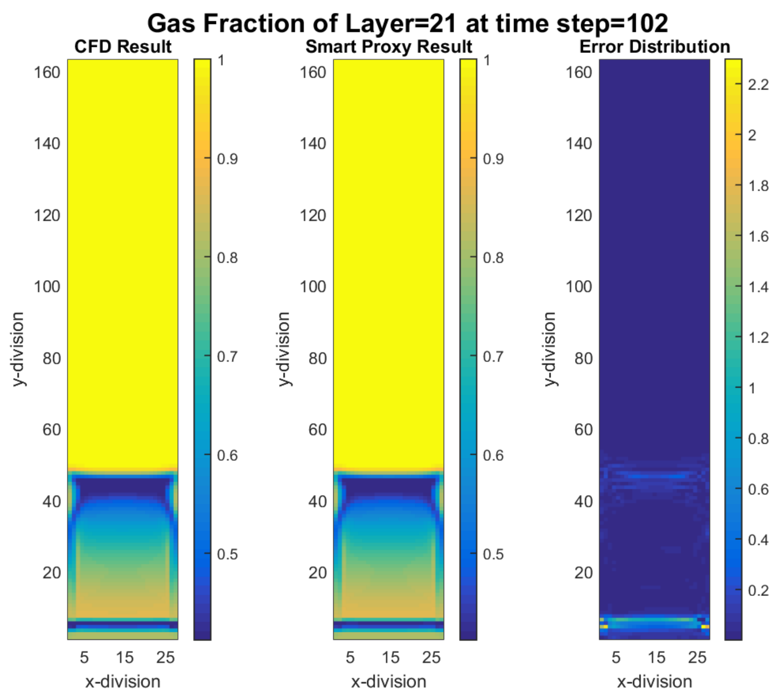


Figure 4-5 CFD and smart proxy results for gas volume fraction at K = 21 cross-sectional plane

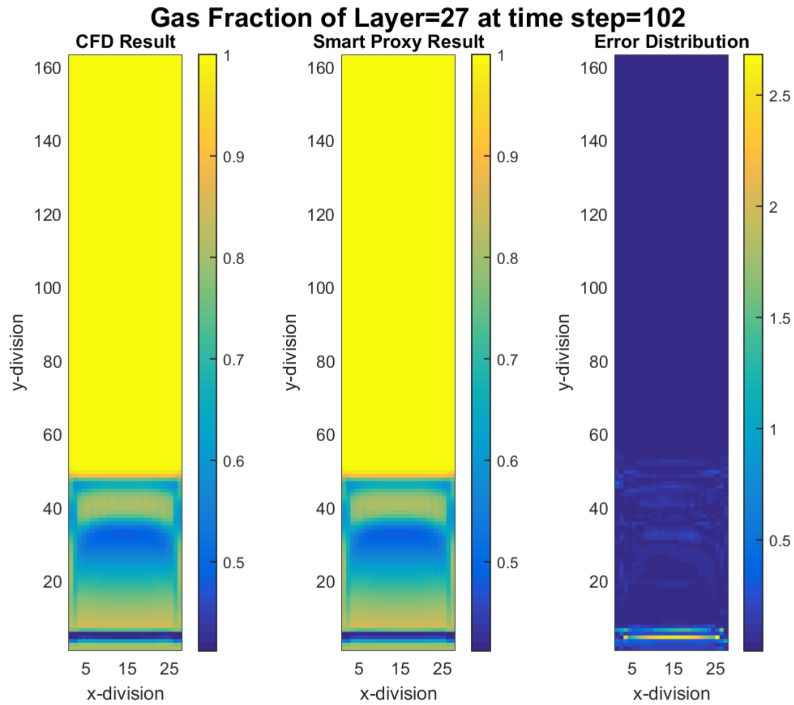


Figure 4-6 CFD and smart proxy results for gas volume fraction at K = 27 cross-sectional plane

4.2.2 Gas Pressure

The results of the smart proxy versus CFD for the gas pressure are shown in Figure 4-7 through Figure 4-11. As it is shown in Figure 4-7, Smart proxy is able to replicate the pressure distribution well, when compared to CFD simulation results, at time-step 102. The error is less than 2% everywhere except in the vicinity of inlet, where the maximum error is less than 20%.

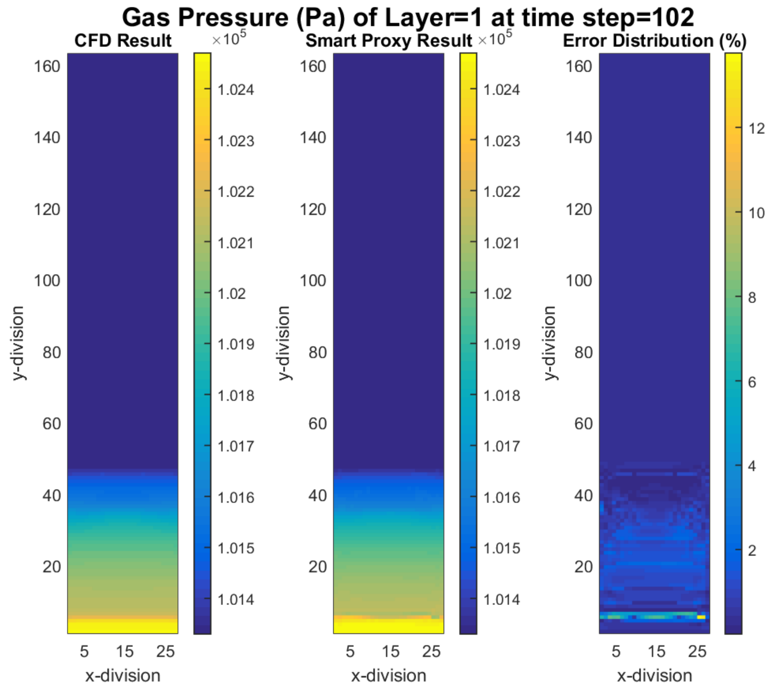


Figure 4-7 CFD and smart proxy results for gas pressure at K = 7 cross-sectional plane

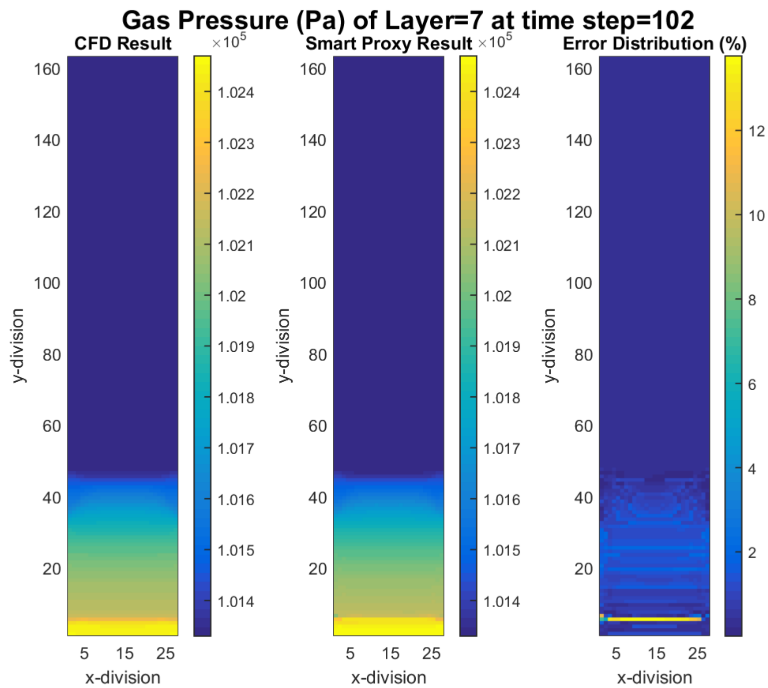


Figure 4-8 CFD and smart proxy results for gas pressure at K = 7 cross-sectional plane

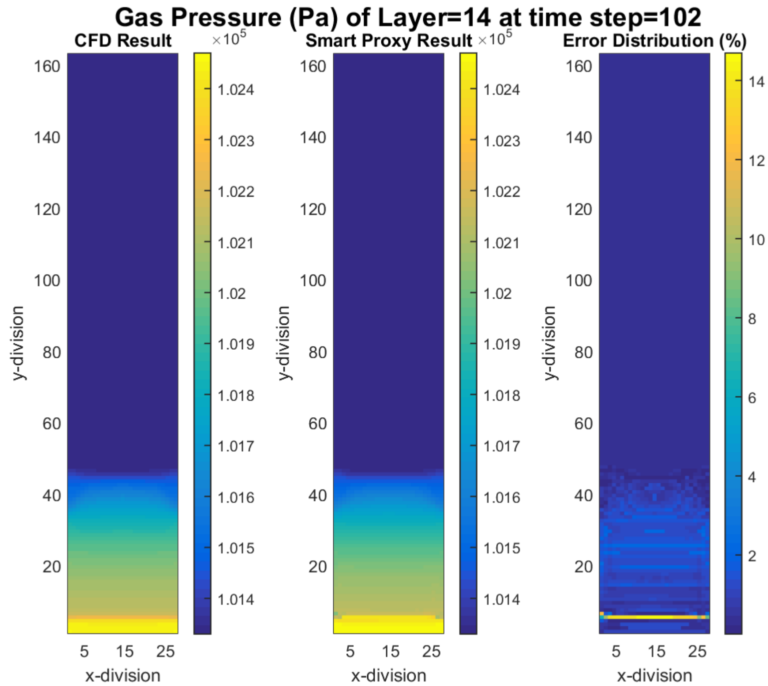


Figure 4-9 CFD and smart proxy results for gas pressure at K = 14 cross-sectional plane

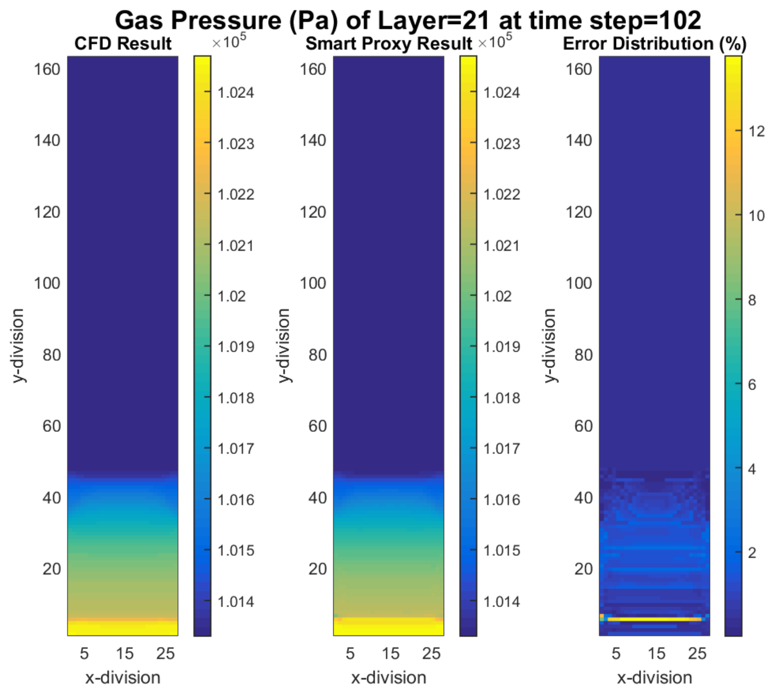


Figure 4-10 CFD and smart proxy results for gas pressure at K = 21 cross-sectional plane

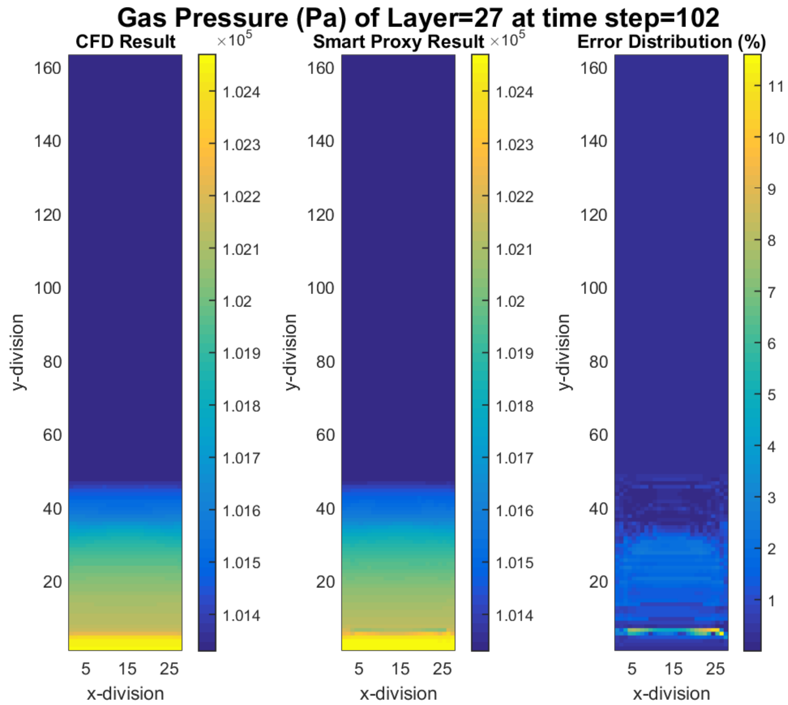


Figure 4-11 CFD and smart proxy results for gas pressure at K = 27 cross-sectional plane

4.3 LATE TIME-STEP, NON-CASCADING, SINGLE OUTPUT, EXPLICIT

In this section, the results of the training and deployment are presented for late time-steps when the bed is fully fluidized. The time steps 4000 and 4002 are selected to train the network. Similar to the previous scenario, the ANN had only one output, so 9 different ANN are trained for all 9 parameters. For the deployment, all the time-steps from 4002 to 4040 were input to the ANN and acceptable results were obtained. In the next sections, the results of the gas volume fraction are provided for time-step 4004, additional figures for next time-steps are provided in Appendix III.

4.3.1 Gas Volume Fraction

The results of the smart proxy versus CFD for gas volume fraction when the bed is fluidized are shown by Figure 4-12 through Figure 4-16. The smart proxy is able to capture the bubbles in the fluidized bed, with maximum error of around 4%. The error is mostly around the interface of gas and solid.

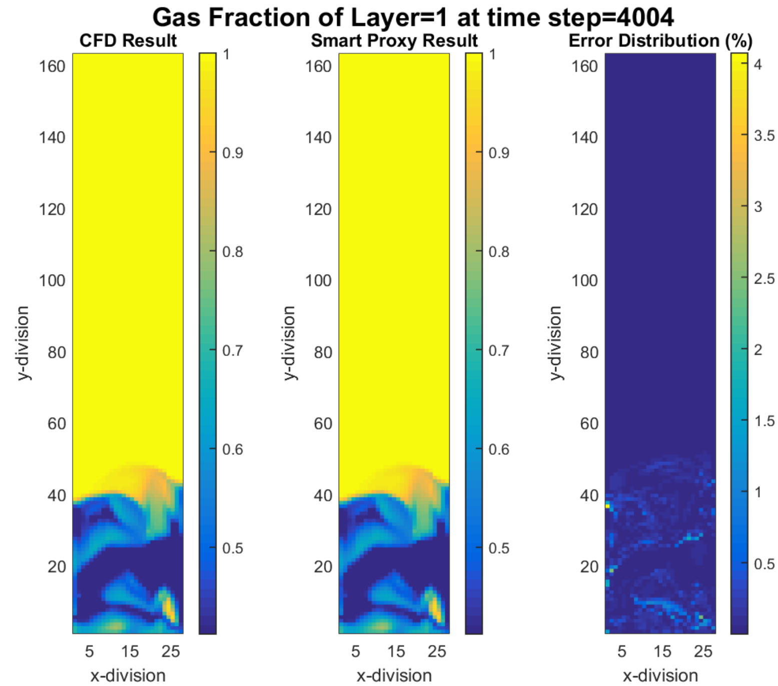


Figure 4-12 CFD and smart proxy results for gas volume fraction at K = 1 cross-sectional plane

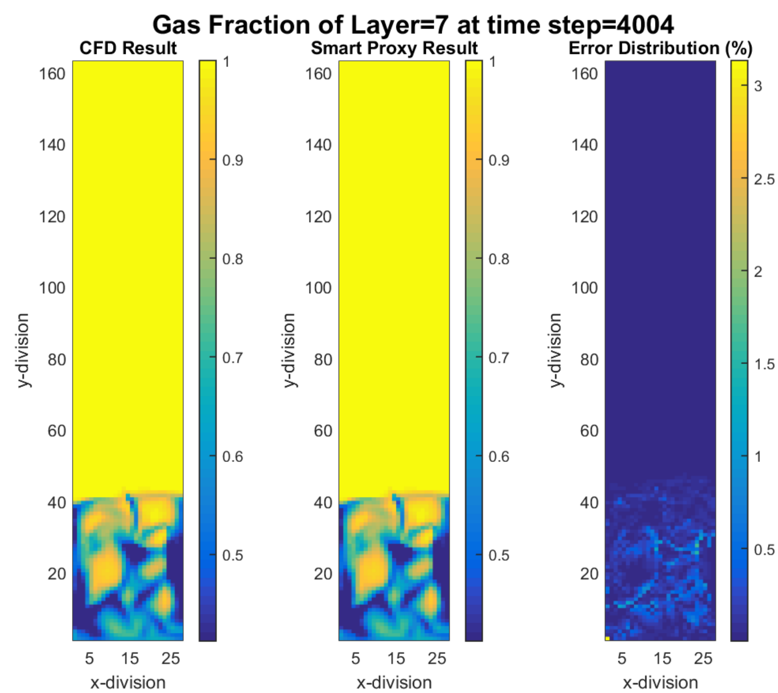


Figure 4-13 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane

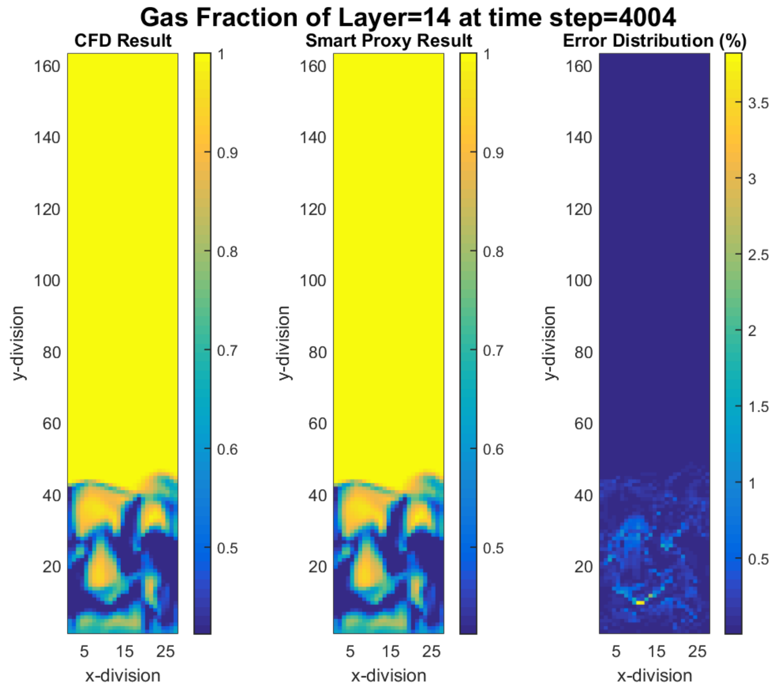


Figure 4-14 CFD and smart proxy results for gas volume fraction at K = 14 cross-sectional plane

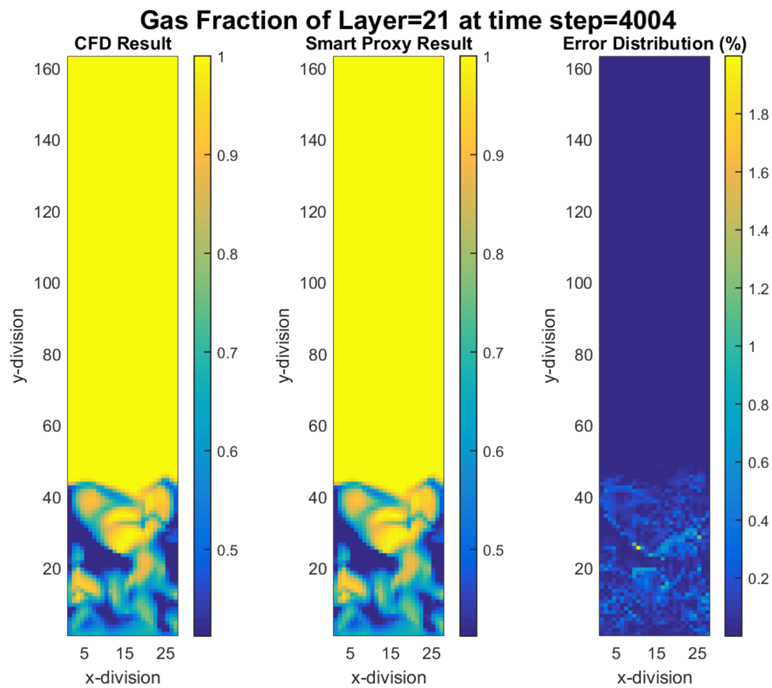


Figure 4-15 CFD and smart proxy results for gas volume fraction at K = 21 cross-sectional plane

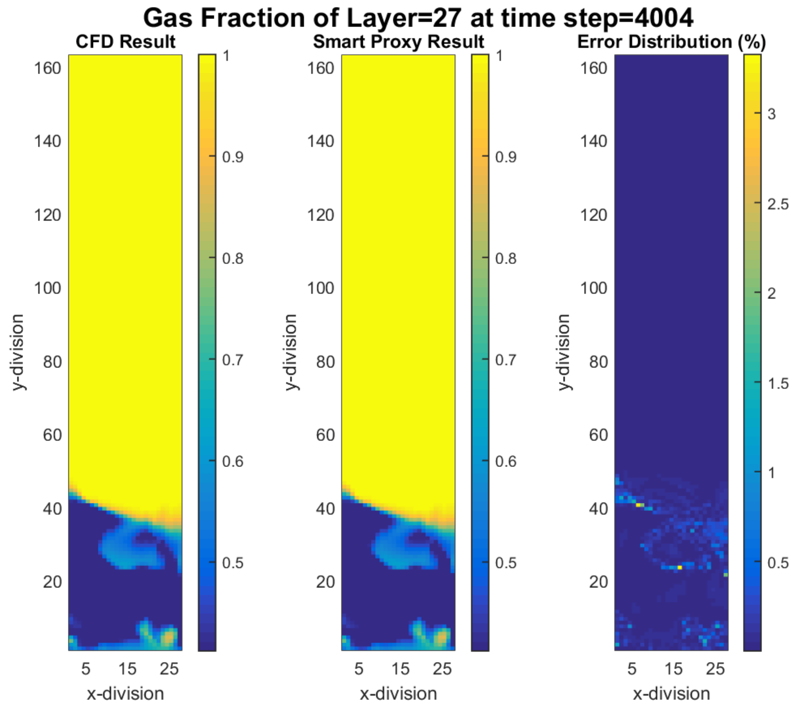


Figure 4-16   CFD and smart proxy results for gas volume fraction at K = 27 cross-sectional plane

4.4 CASCADING, SINGLE OUTPUT, EXPLICIT

In the previous section, it is shown that ANN is able to mimic the CFD results both when there is moderate change in the dynamics of the multi-phase flow and when the bed is fully fluidized, when non-cascading approach is selected. In this section, the results of the cascading approach are presented.

4.4.1 Gas Volume Fraction for Early Time

The network is trained for gas volume fraction by introducing time-step 100 as input and time-step 101 as output. Then for the deployment process, time-step 100 was used as the input to ANN, with the output of each ANN being used as input for the next time step. Figure 4-17 through Figure 4-26 show the voidage contours based on CFD and ANN, along with the percent error between time step 101 (0.101 sec elapsed time) to time step 110 (0.11 sec elapsed time). These figures show that percent error increases in time. The error propagation from time step to time step can eventually go beyond the user defined tolerance level that forces the cascading deployment process to terminate. In order to overcome the error propagation, more time-steps should be used for training. This will be discussed in later sections.

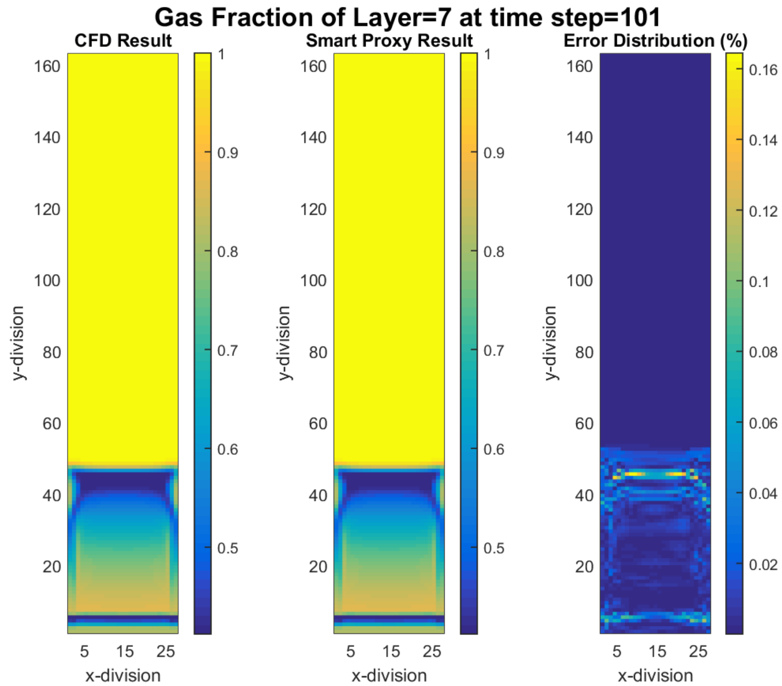


Figure 4-17 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 101

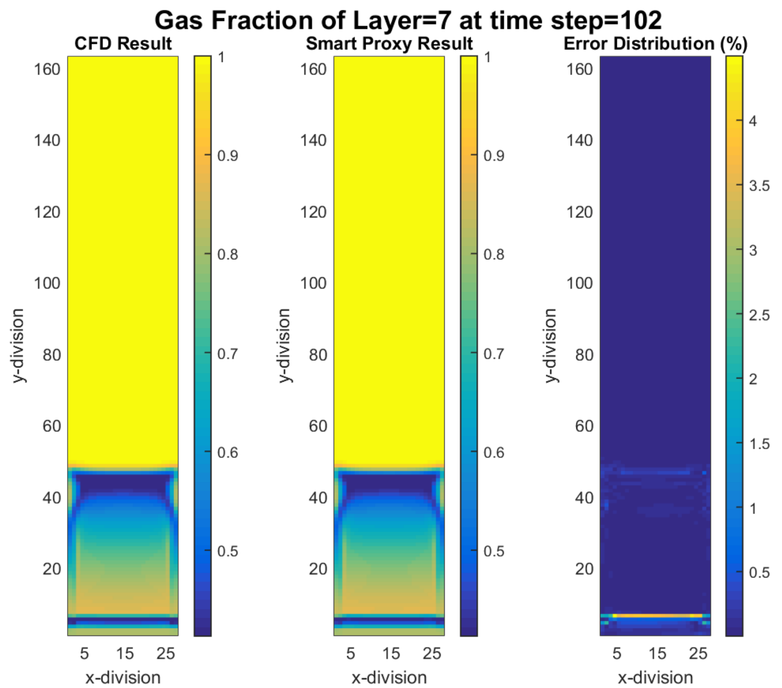


Figure 4-18 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 102

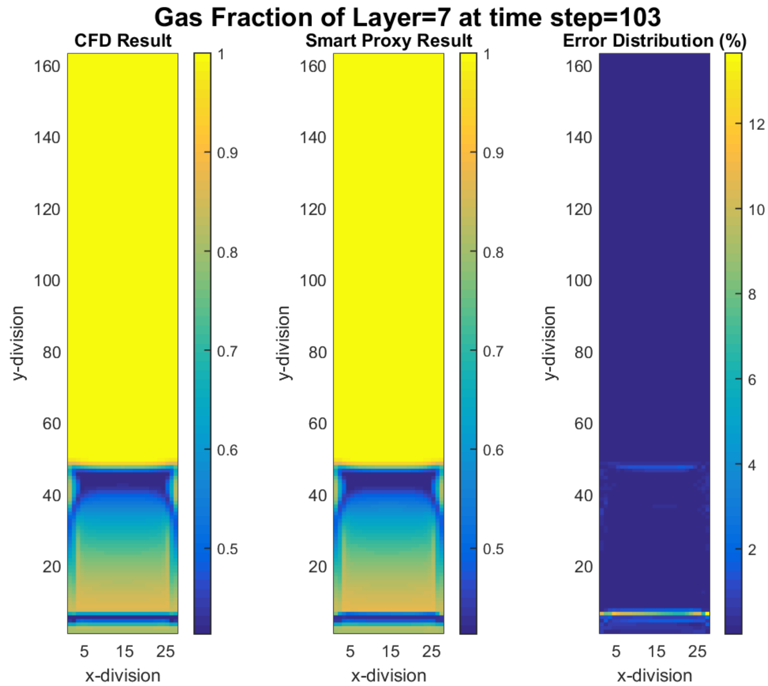


Figure 4-19 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 103

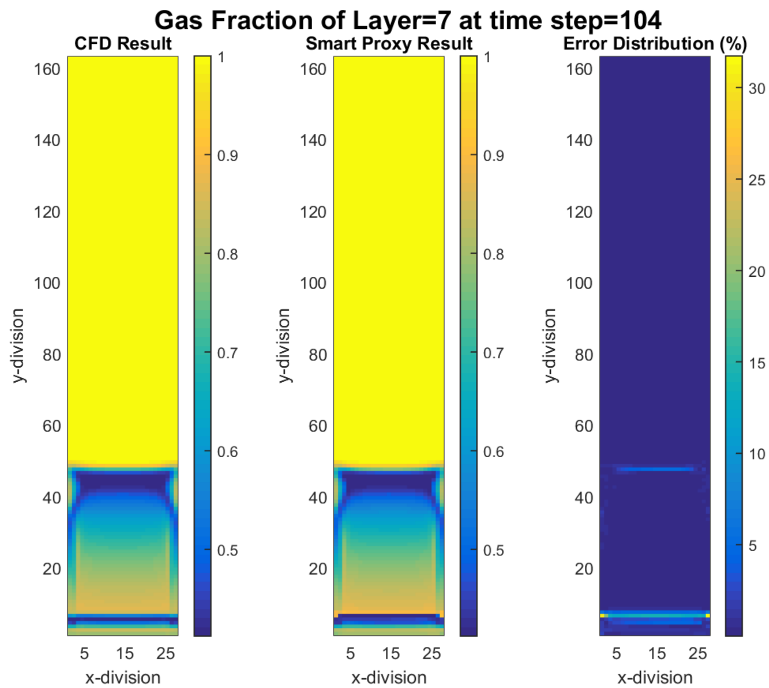


Figure 4-20 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 104

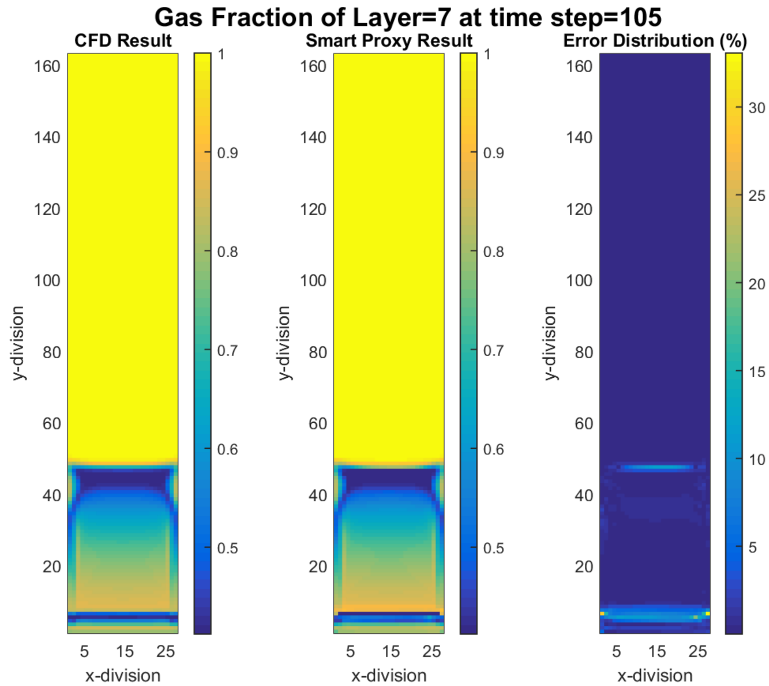


Figure 4-21 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 105

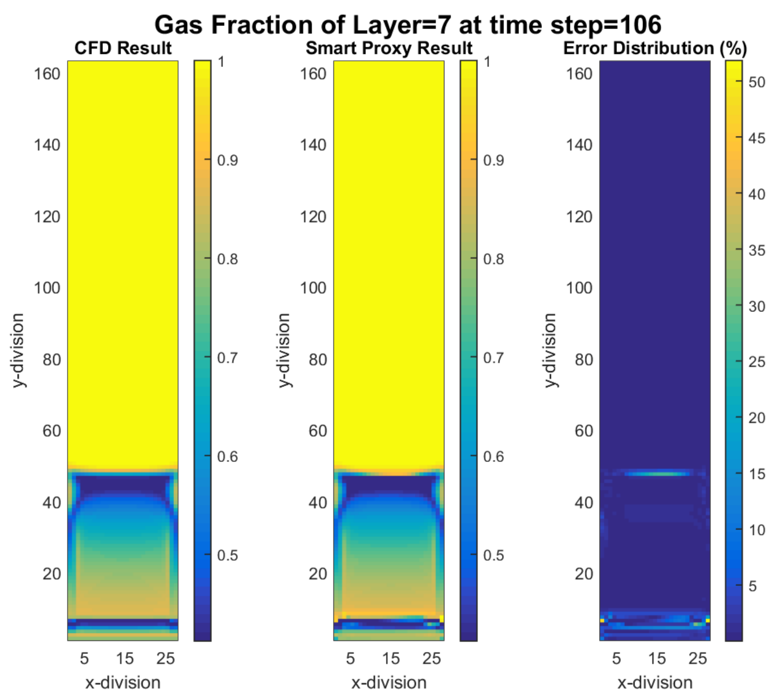


Figure 4-22 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 106

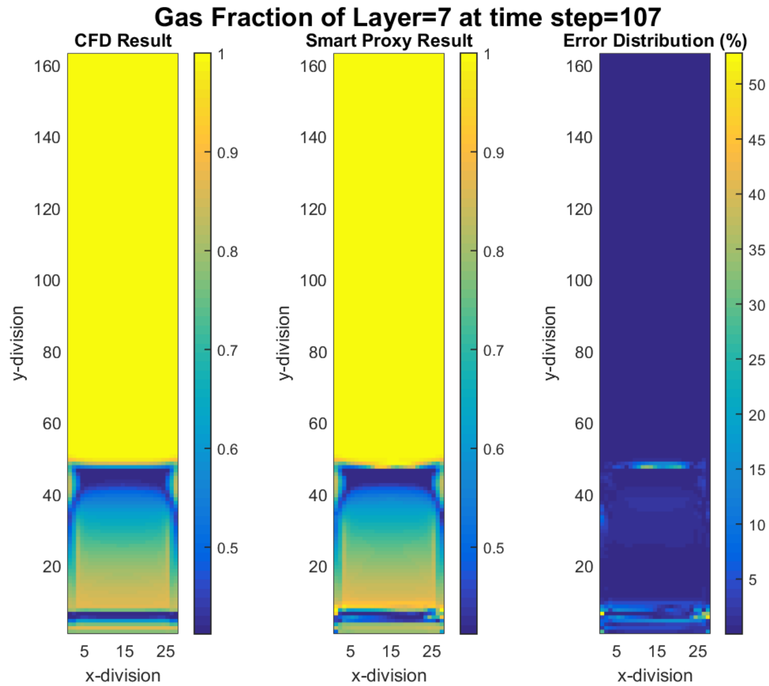


Figure 4-23 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 107

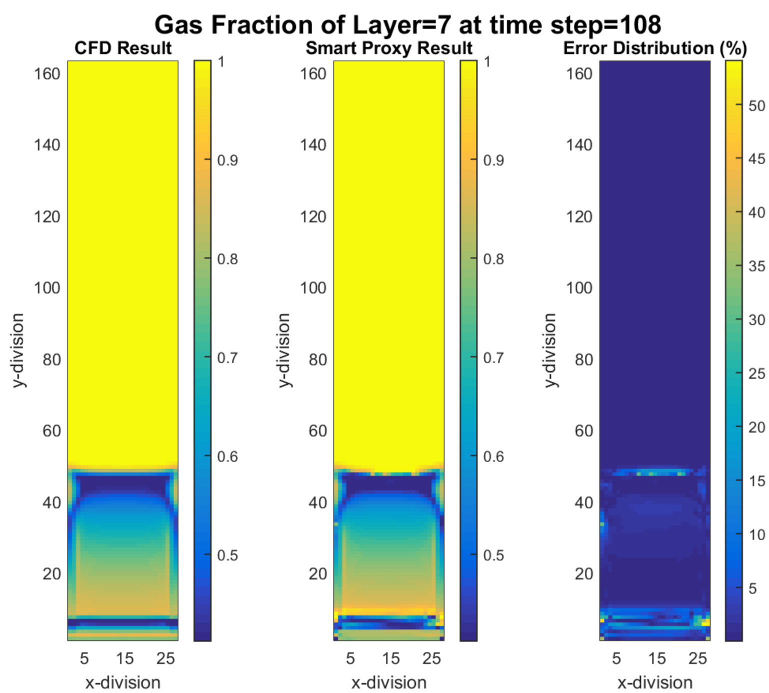


Figure 4-24 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 108

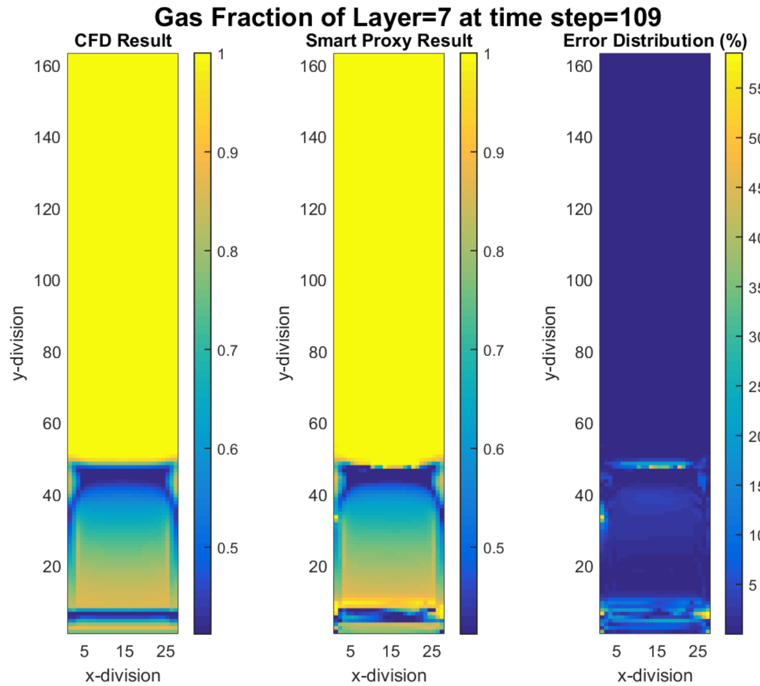


Figure 4-25 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 109

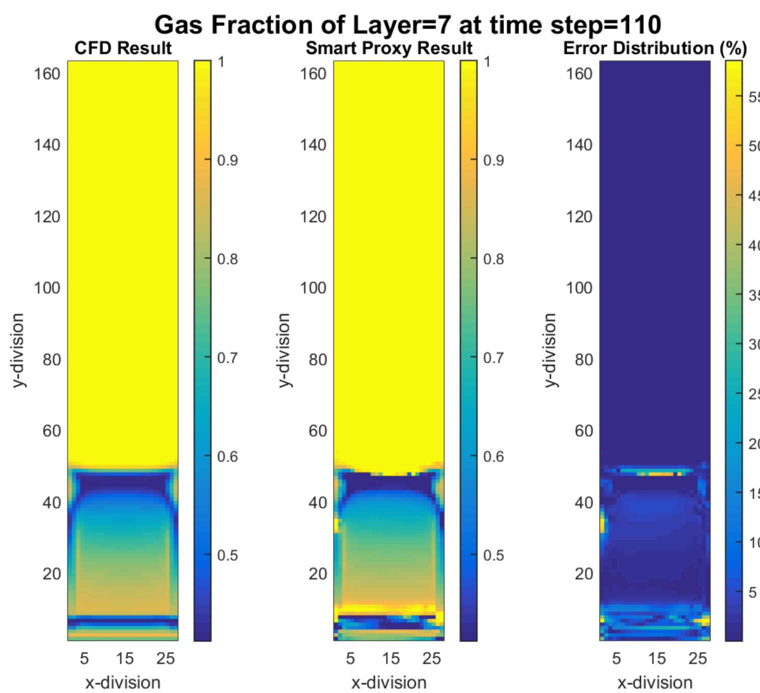


Figure 4-26 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 110

4.4.2 Gas Volume Fraction for Late Time

Figure 4-27 through Figure 4-30 show the contours of voidage, for time step 4002 (4.002 sec elapsed time) through 4020 (4.02 sec of elapsed time), when flow is fully fluidized. As in section 4.4.1, error propagation in ANN from time frame to time frame increases in time. Techniques for minimizing the accrued error will be discussed in the following sections.

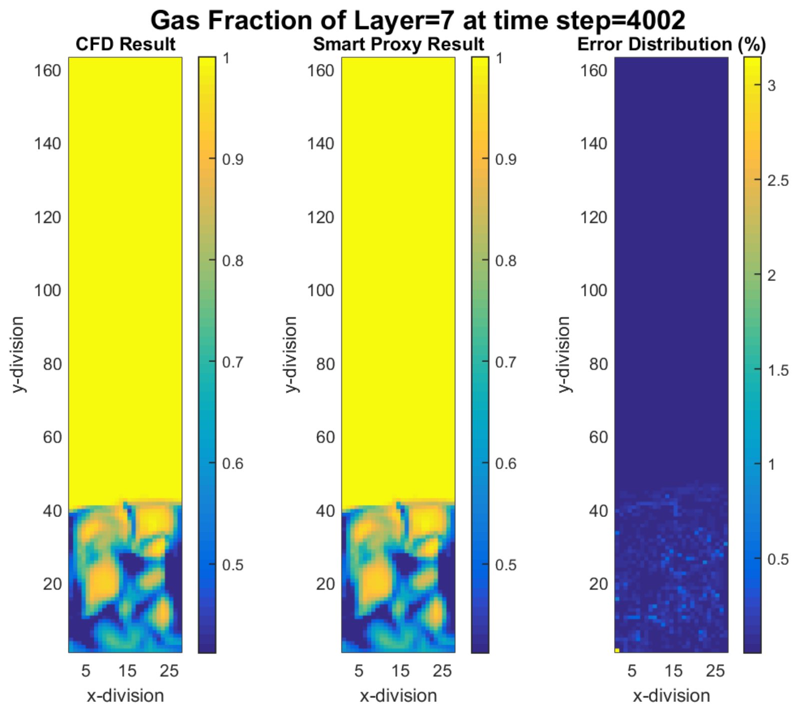


Figure 4-27 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4002

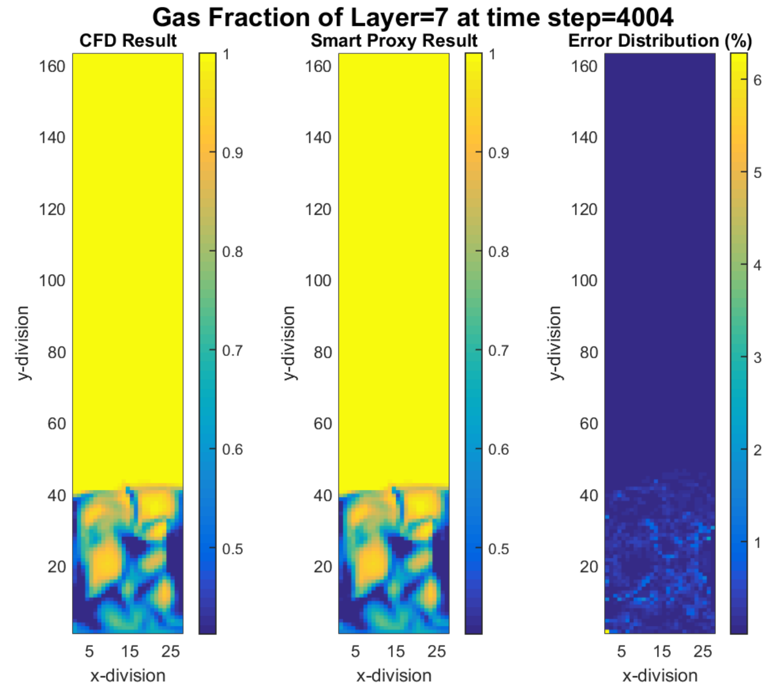


Figure 4-28 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4004

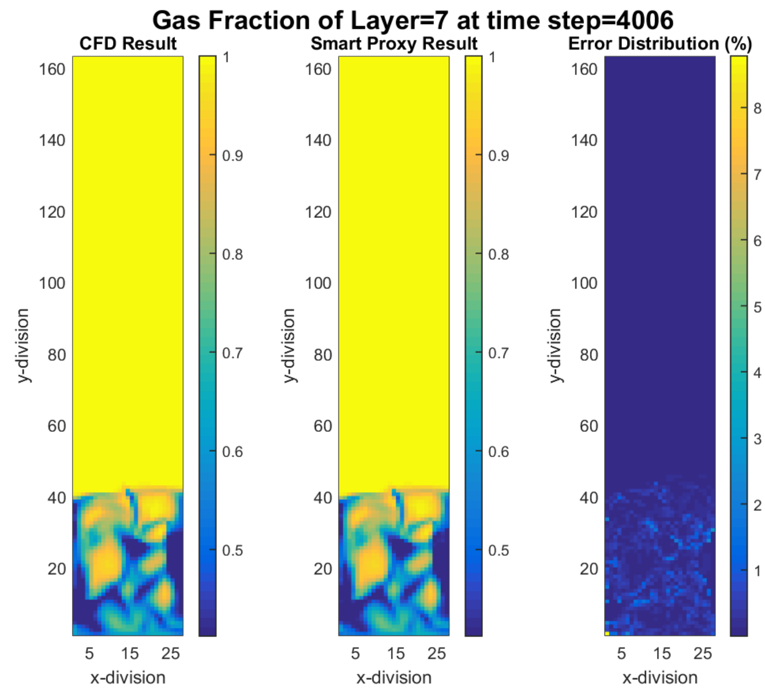


Figure 4-29 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4006

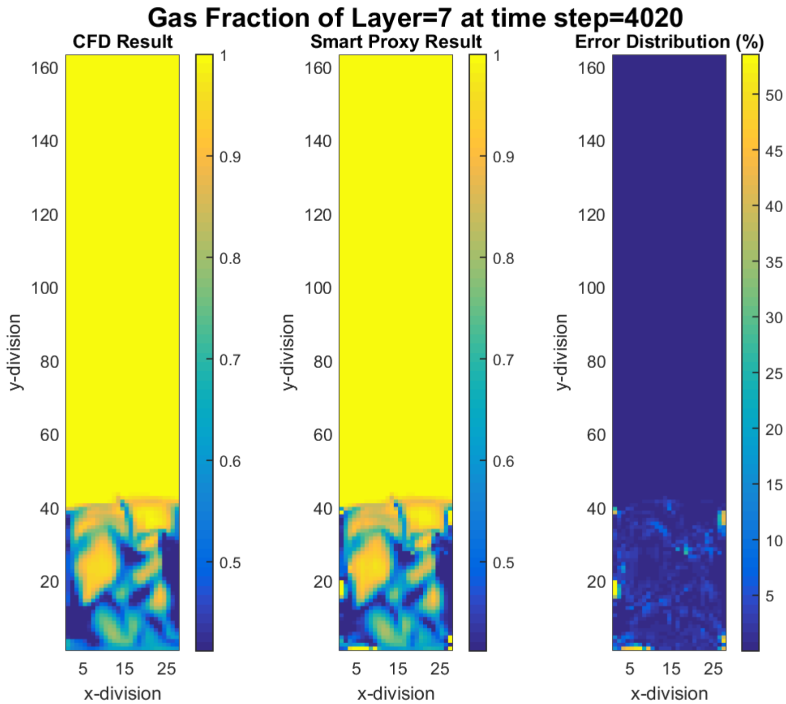


Figure 4-30 CFD and cascading smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4020

4.5 EARLY TIME-STEP, NON-CASCADING, MULTIPLE OUTPUT, EXPLICIT

Based on the discussion in the previous sections, it is sometimes beneficial to have multiple outputs rather than only a single output. Three components of gas velocity were selected to be the output of the ANN; the input data is exactly the same as previous scenarios. The inputs come from time-step 100 and the outputs come from time-step 101. The ANN is trained successfully and the time-step 102 is predicted.

Figure 4-31 to Figure 4-33 show the results of smart proxy and comparison with CFD simulation results. Smart proxy result compares favorably to CFD simulation result. The highest error is about %20 in the vertical component of velocity, Figure 4-32.

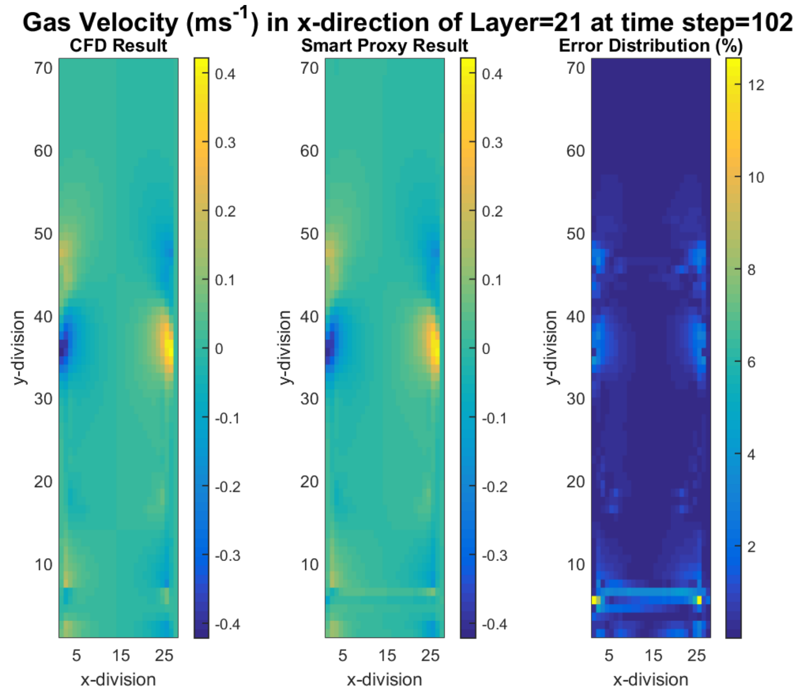


Figure 4-31 CFD and smart proxy results for x-component of gas velocity at K = 21 cross-sectional plane, based on expilictly constructed ANN

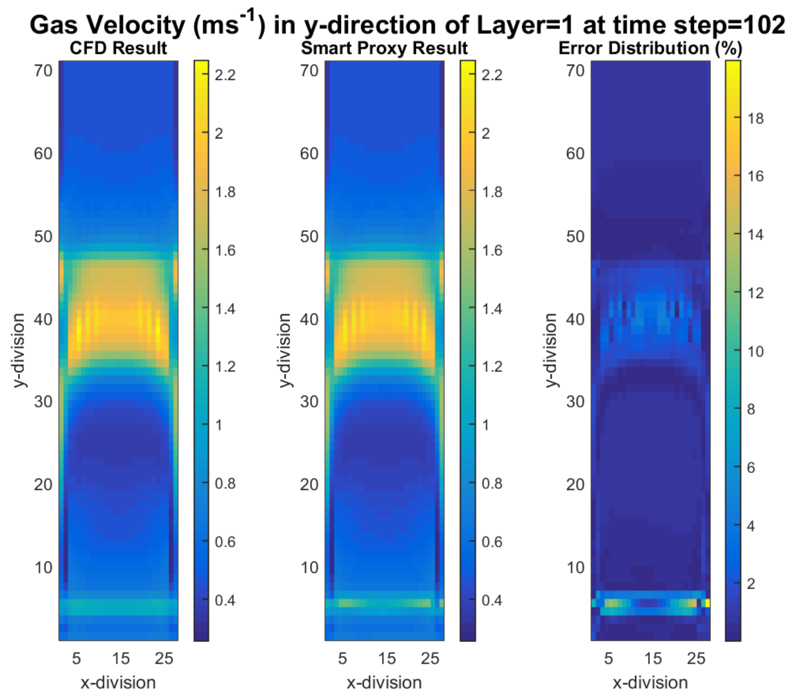


Figure 4-32 CFD and smart proxy results for y-component of gas velocity at K = 1 cross-sectional plane, based on expilictly constructed ANN

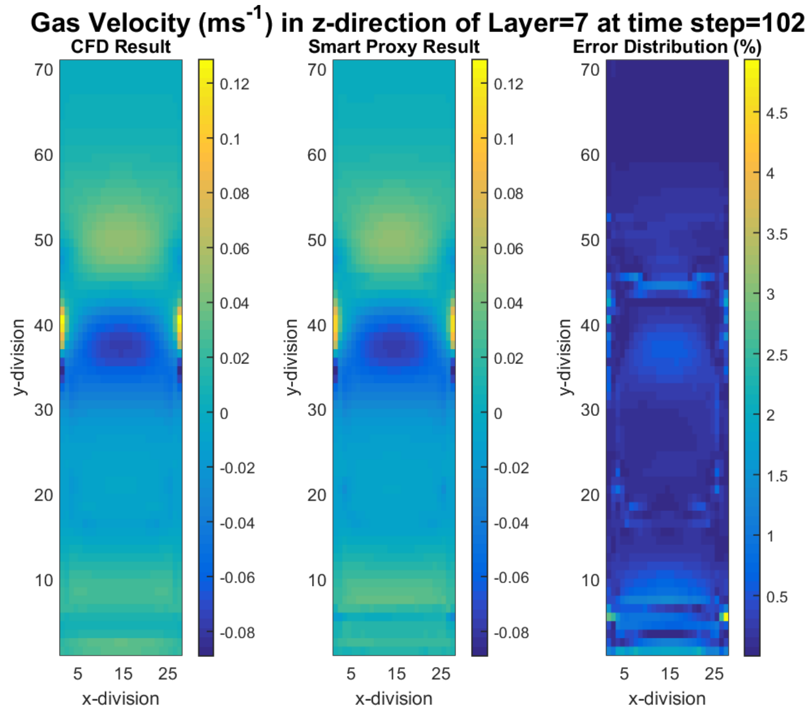


Figure 4-33 CFD and smart proxy results for z-component of gas velocity at K = 7 cross-sectional plane, based on expilictly constructed ANN

4.6 EARLY TIME-STEP, NON-CASCADING, MULTIPLE OUTPUT, IMPLICIT

The process in section 4.5 is repeated, this time with an implicit scheme. The input of the ANN consists of all parameters at time step 100, except solid velocity that comes from time step 101. The output is the three components of gas velocity from time-step 101. Figure 4-34 to Figure 4-36 show the results of this scenario. A comparison of Figure 4-31 and Figure 4-34 shows that maximum percent error in the x-component of velocity decreases from 12% to 9%, when an implicit scheme is used instead of explicit. Maximum error in the y-component of velocity deceases from 19% to 12%, when using an implicit scheme, as shown in Figure 4-32 and Figure 4-35. A comparison of Figure 4-33 and Figure 4-36 show that the maximum error in z-component of velocity reduces from 5% to 3% when the implicit scheme is utilized.

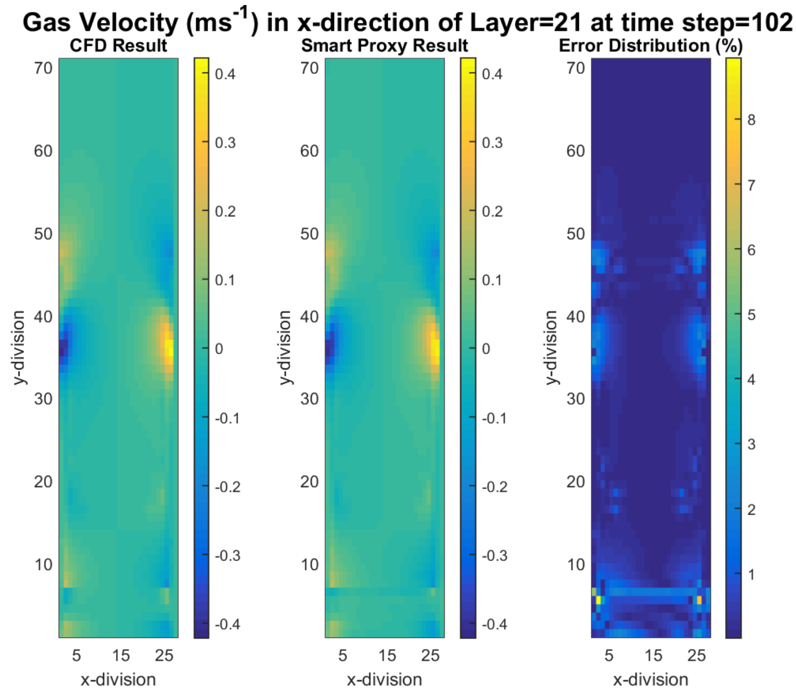


Figure 4-34 CFD and smart proxy results for x-component of gas velocity at K = 21 cross-sectional plane, based on implicitly constructed ANN

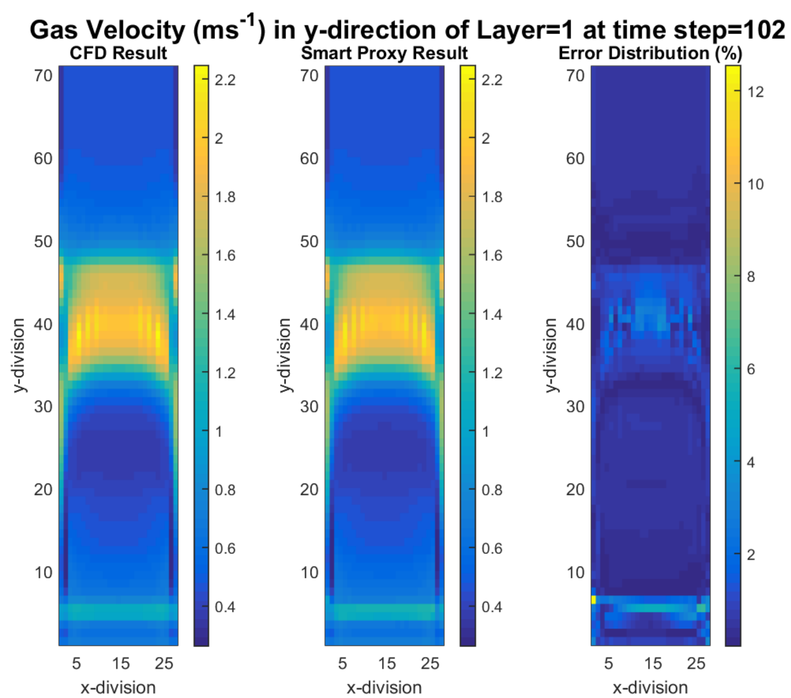


Figure 4-35 CFD and smart proxy results for y-component of gas velocity at K = 1 cross-sectional plane, based on implicitly constructed ANN

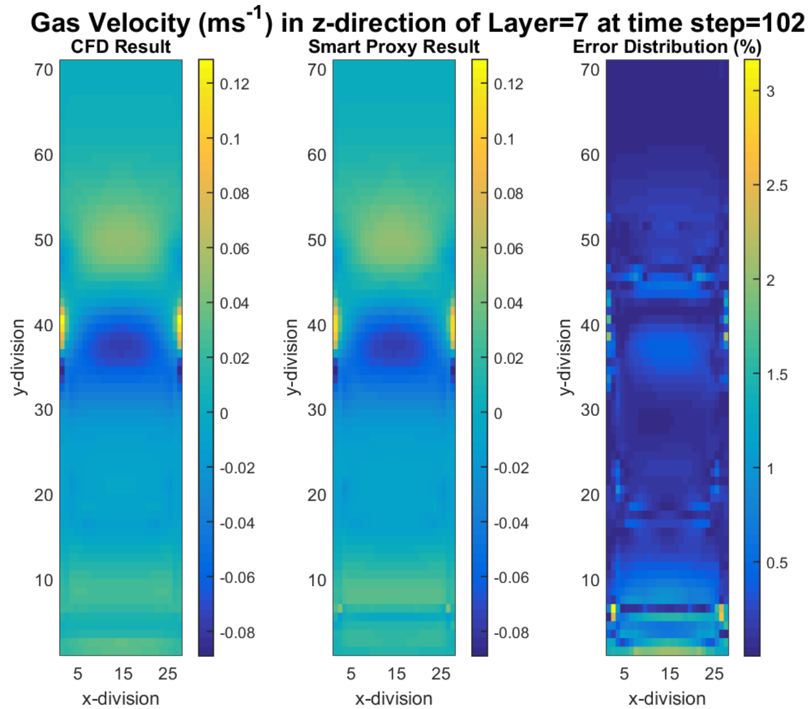


Figure 4-36 CFD and smart proxy results for z-component of gas velocity at K = 7 cross-sectional plane, based on implicitly constructed ANN

4.7 USING MULTIPLE TIME-STEPS FOR TRAINING, NON-CASCADING, SINGLE OUTPUT, EXPLICIT

As discussed in the previous sections, the ANN for early time steps (startup period) is valid only for early period and the ANN for later time steps (fully fluidized state) only is valid for the later time periods. A more general approach has to be able to cover the entire time span (different flow regimes). The ANN is therefore constructed using additional time steps that account for the transition from the startup condition to fully fluidized condition. Time steps 200-202 and 1000-1002 have been added to the training data set (4000-4002). The deployment process is then conducted with the trained ANN by inputting time step 200 all the way to time step 4000. The results are presented in terms of RMSE of gas volume fraction in Figure 4-37. Figure 4-37 shows the RMSE distribution versus time-steps. It is clear that in the time-steps that we had training data, the amount of error is minimum but in the other time-steps the RMSE increased. The time steps with larger peaks in Figure 4-37 point to the need for additional ANN training at those time steps. Or potentially other dynamic behaviors are taking place in the bed that additional training is needed. The contour plot of voidage at time step 500, where the RMSE is high (see Figure 4-37) is shown in Figure 4-38. It is clear that at around this time step, flow is transitioning from the initial plug flow behavior to a more bubbling flow. This change in the flow regime is not properly captured, since the input data to ANN at the training stage did include any data from the time steps, when transitioning is taking place. Figure 4-39 shows a decrease in RMSE, when data from time steps where RMSE peaks in Figure 4-37 are added to the training data set.

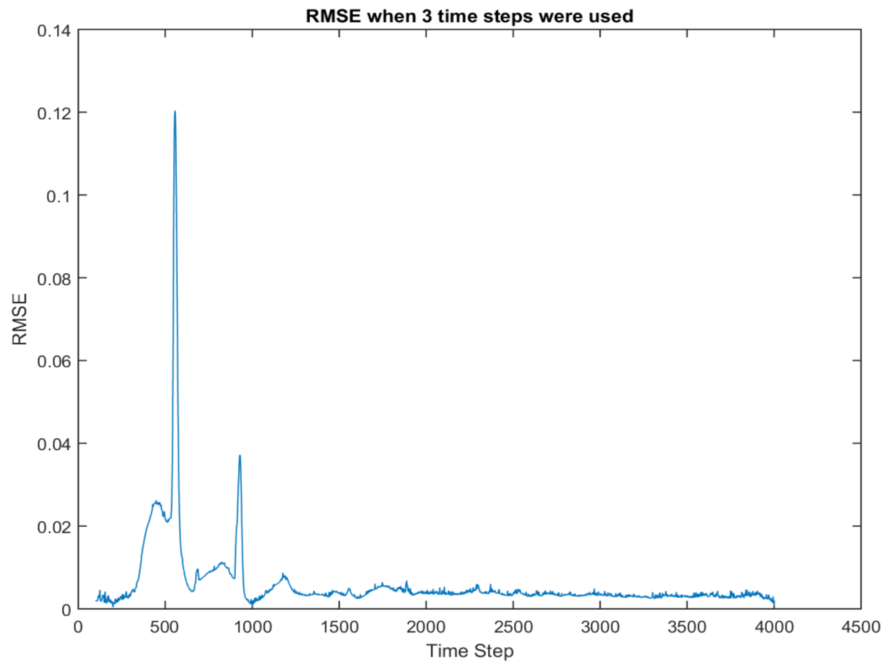


Figure 4-37 RMSE distribution over time, when three pairs of data are used for training

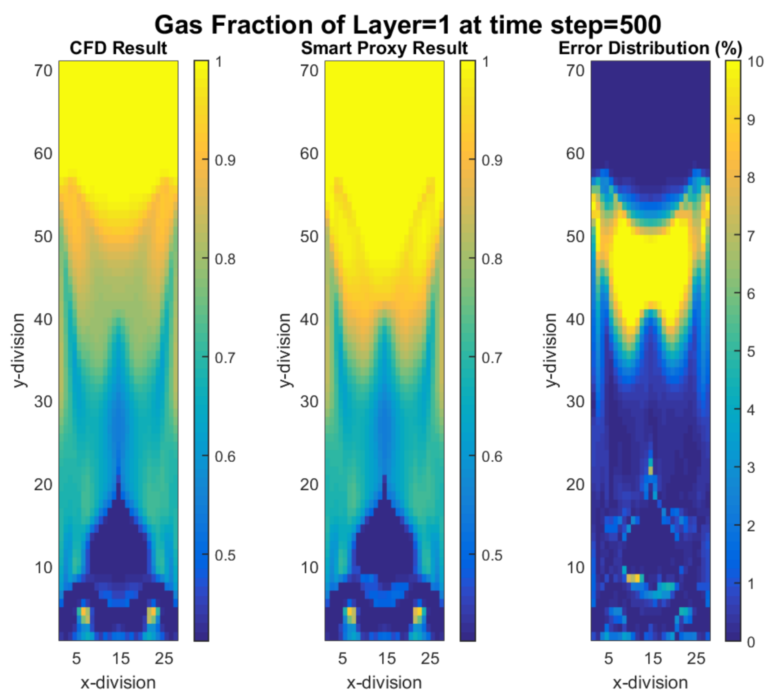


Figure 4-38 CFD and smart proxy results for gas volume fraction at K = 1 cross-sectional plane

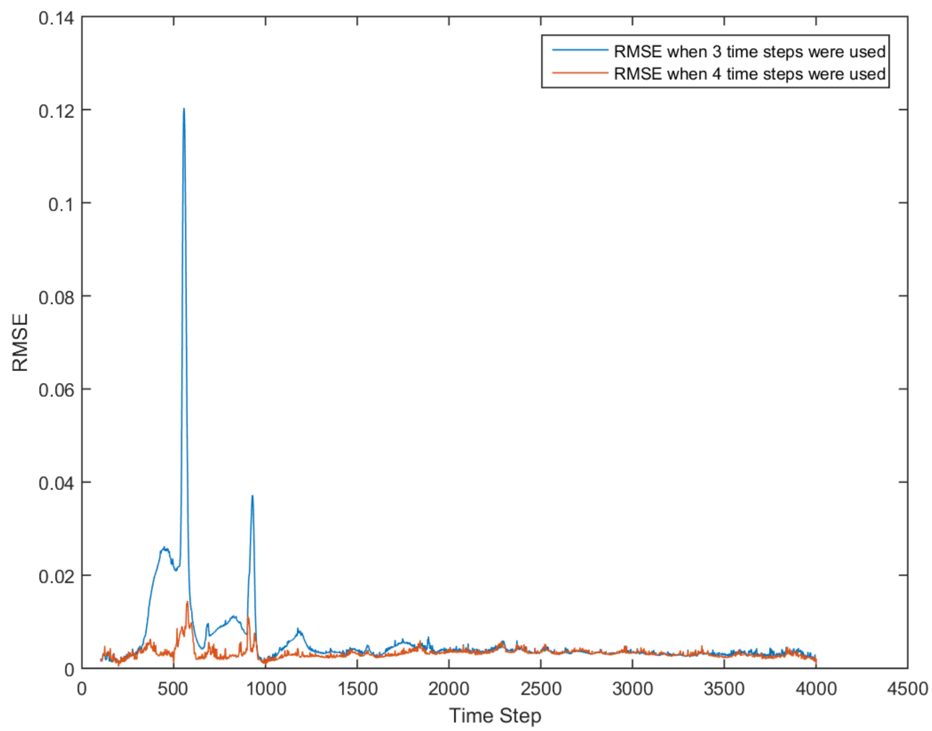


Figure 4-39 RMSE distribution over time, when four pairs of data are used for training

4.8 REDUCING THE NUMBER OF PARAMETERS (KPI)

It was shown earlier that eliminating the free board section from the network reduces the size of the input data. By eliminating some of the parameters from the training process, the size of the input data can be reduced even further. This can be done through a ranking analysis of total weights of the parameters used in training of ANN. Since in back propagation method, there is a weighted summation between all the parameters from each layer to the next layer, the total weight could be obtained by averaging all the weights corresponding to a specific parameter. There are two different ways to find the total weights; averaging all the weights by considering their signs or averaging all the weights by considering the absolute value of each weight. Table 4-1 and Table 4-2 show the 14 least important parameters, using the two-different averaging scheme, described above. The remaining 56 parameters are used to train the network. Figure 4-40 and Figure 4-41 show the Tornado chart for some of the remaining 56 parameters that is used for training the ANN. Figure 4-42 shows the comparison of RMSE distribution from ANN results based on the technique described in section 4.7 and the two different averaging techniques, with reduced number of parameters. Clearly prioritizing the parameters, based on the absolute value of their total weight leads to less error in the ANN results. The fact that reducing the number of input parameters, which are used for training the ANN, leads to reduction in RMSE shows the important role optimization of the spatio-temporal database has in the training process. In effect, eliminating the less important parameters from ANN training has improved the training process, by eliminating irrelevant connections between the parameters and increasing the potential for detecting stronger correlations among the key parameters.

Table 4-1 The least important parameters based on averaging of the weights

By parameter	By location
ug-south	us-top
ug-east	
ug-north	ug-north
	us-north
wg-east	ws-north
wg-west	
	ug-south
us-top	
us-north	wg-west
us-east	ws-west
ws-north	ug-east
ws-west	wg-east
	us-east
distance to south	distance to south
distance to east	distance to east
distance to north	distance to north
distance to west	distance to west

Table 4-2 The least important parameters based on averaging of absolute value of the weights

By parameter	By location
ug-west	vg-self
ug-north	
ug-south	vs-bottom
ug-east	
	ug-north
vg-self	wg-north
vg-west	us-north
wg-north	ug-south
us-north	ws-west
	vs-west
ws-west	ug-west
	vg-west
vs-west	
vs-bottom	ug-east
distance to south	distance to south
distance to east	distance to east
distance to west	distance to west

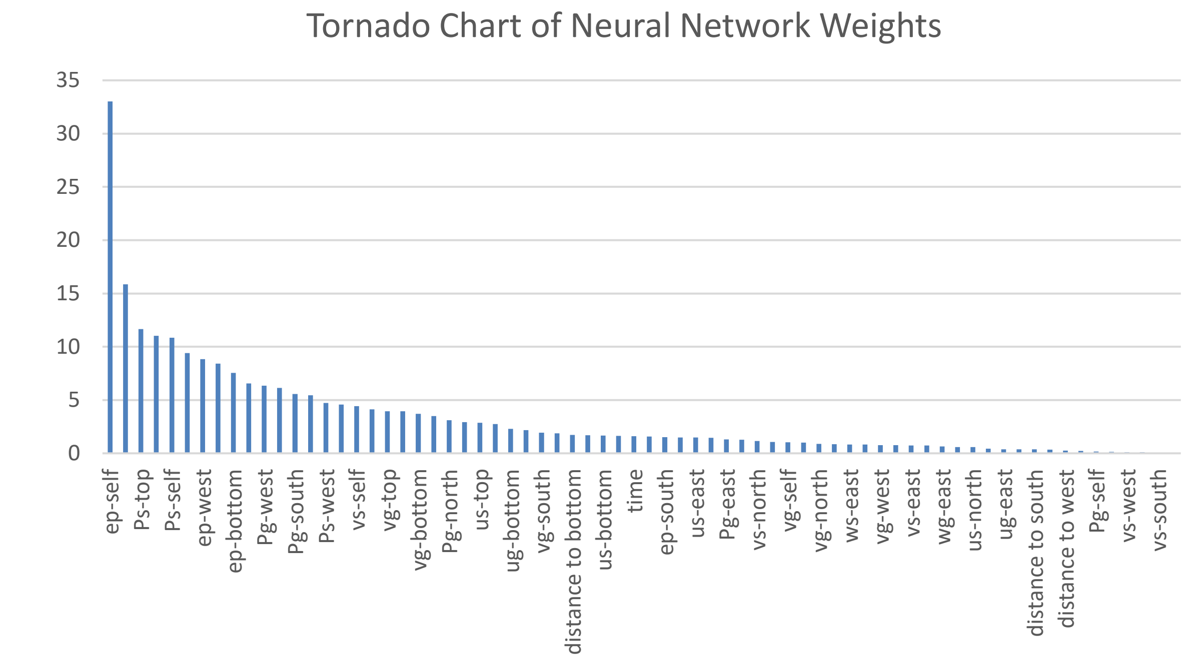


Figure 4-40 Parameters ranking for construction of gas volume fraction ANN, based on averaging all weights

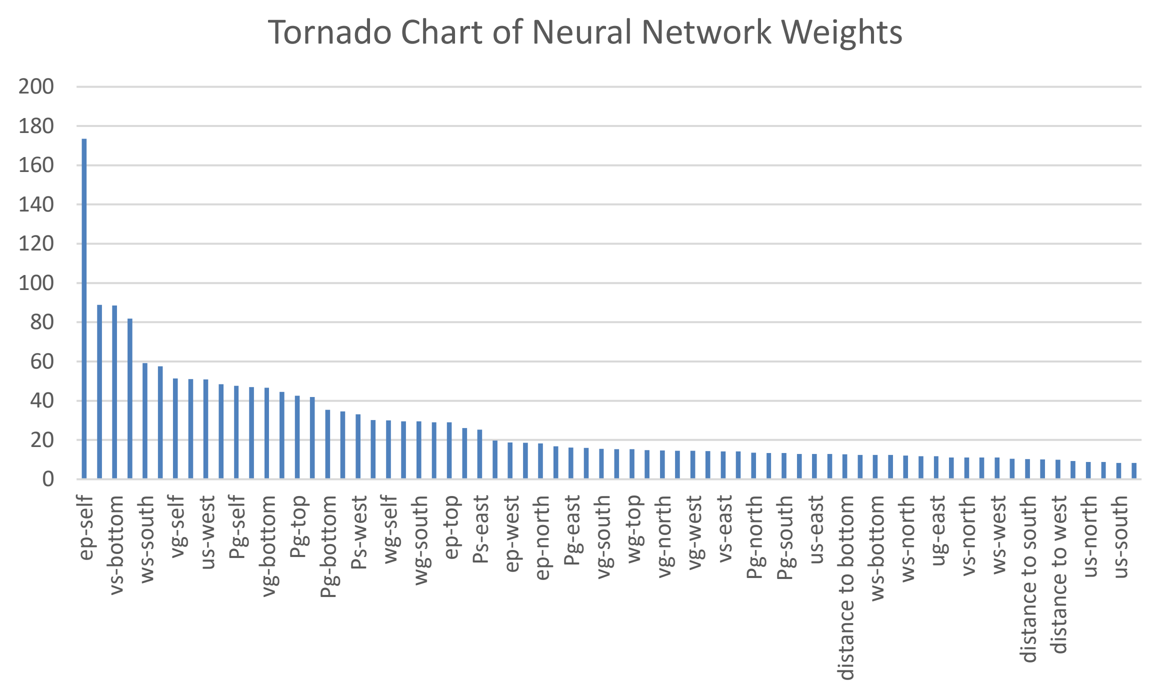


Figure 4-41 Parameters ranking for construction of gas volume fraction ANN, averaging of absolute value of all weights

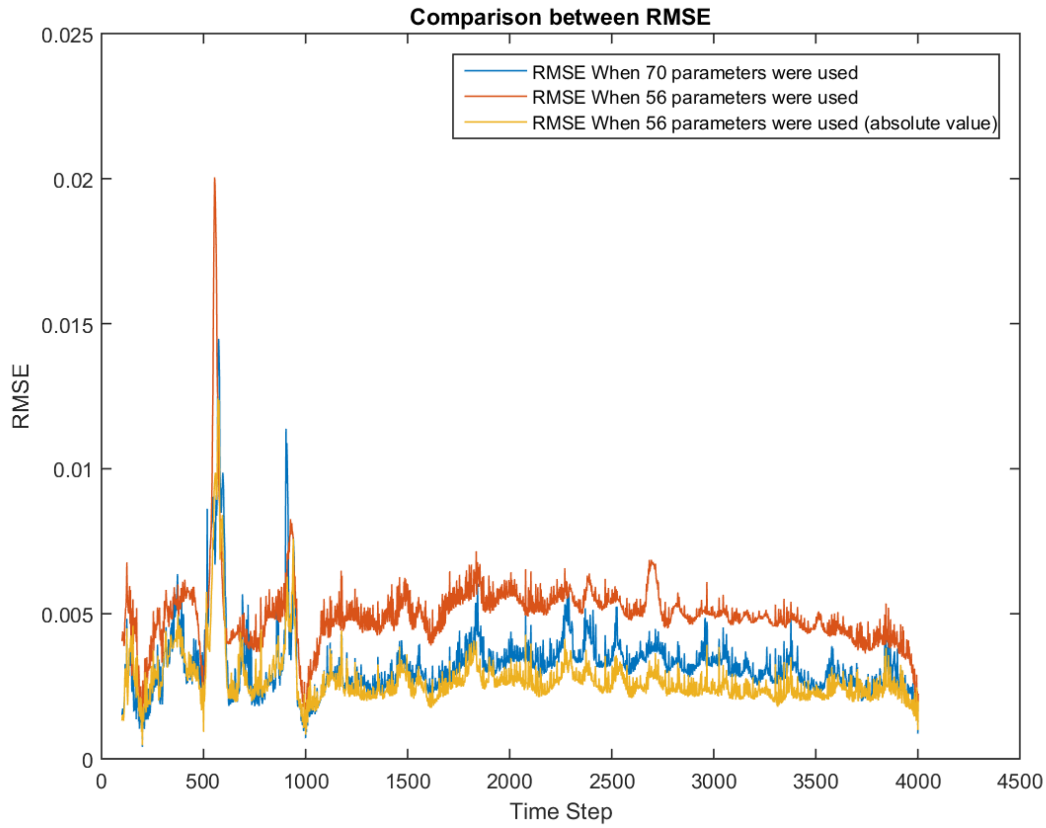


Figure 4-42 RMSE distribution over time with two different averaging approach

Since reducing the number of input parameters for training led to a reduction in RSME, additional sensitivity analysis is performed in order to reduce the sampling size even further. Figure 4-43 shows the RMSE, with 42 and 35 input parameters, along with the original 70 parameters and 56 parameters.

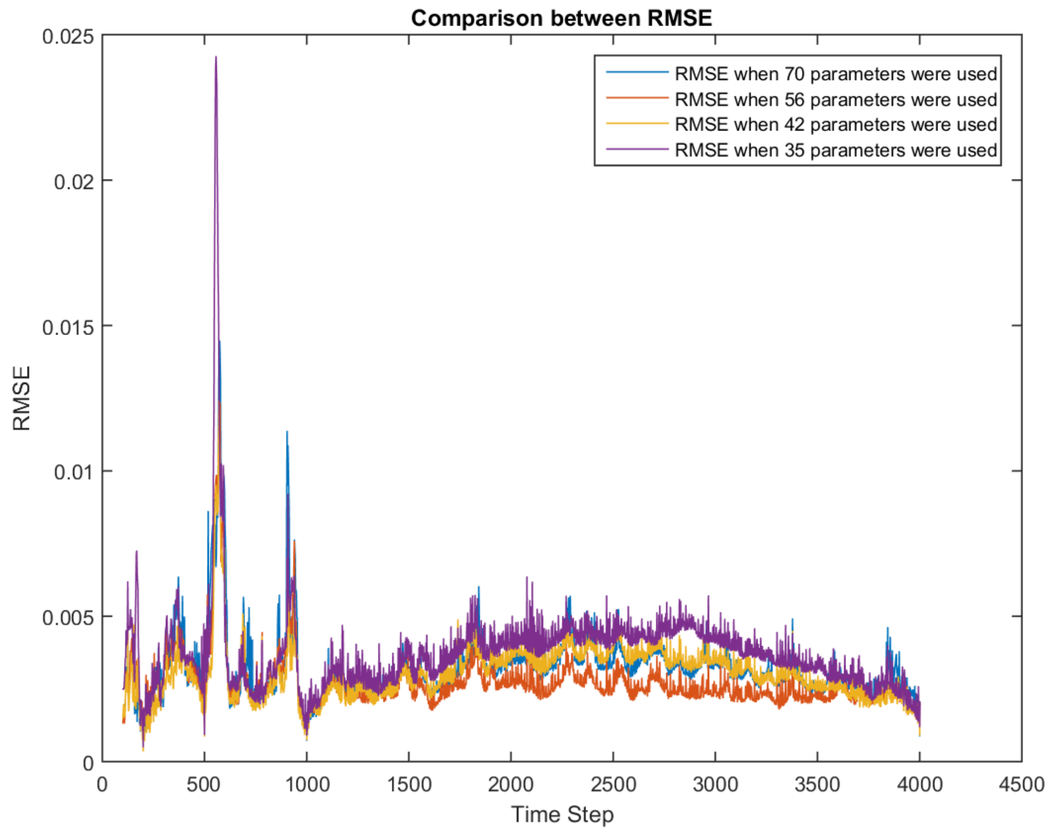


Figure 4-43 RMSE distribution with different parameters used for training

The same analysis is performed for all other parameters, such as pressure and velocities, in order to reduce the training data size. Although the results are not shown here, the number of parameters for training of ANN is set at 42, for all parameters shown in Table 3-2. It is noteworthy that the 42 parameters are not necessarily the same for all the parameters in Table 3-2.

4.9 USING SEVEN TIME-STEPS FOR TRAINING, CASCADING, SINGLE OUTPUT, EXPLICIT

In the previous section, it was shown that the RMSE can be reduced by eliminating non-essential parameters from training dataset. This reduction in sample size allows for using more time steps for training without hitting the hardware memory limitation.

In the previous section, it was mentioned that in order to be able to perform the cascading deployment, more time-steps should be used in the training process. Because of the memory limitations, it was not possible to add more time-steps in the training but now, when the size of input has been decreased; more time-steps could be used in the training, which is the discussion of the next section. Table 4-3 shows the size of data before and after size reduction, the data size is reduced by a factor of 5, if 35 parameters are used for training the ANN.

Table 4-3 Database size before and after optimization

Model	*Size of input*	*Total Data Point*
Original Model	*118,098 by 70*	*8,148,762*
Latest Model	*51,030 by 35*	*1,786,050*

So far, 4 time-steps were used to build the model (time-steps: 200, 500, 1000, and 4000). In this section, the cascading approach is repeated by using 3 more time-steps. These three time-step were chosen based on the highest RMSE in Figure 4-44. As it is shown in the figure, time step 574, 904, and 1842 were selected. Once the ANN is trained using data from the above 7-time steps, the ANN is deployed in a cascading manner, with data from time step of 4000. Prediction is terminated at time step 4020, after the error surpasses the user defined termination threshold, as seen in Figure 4-45.

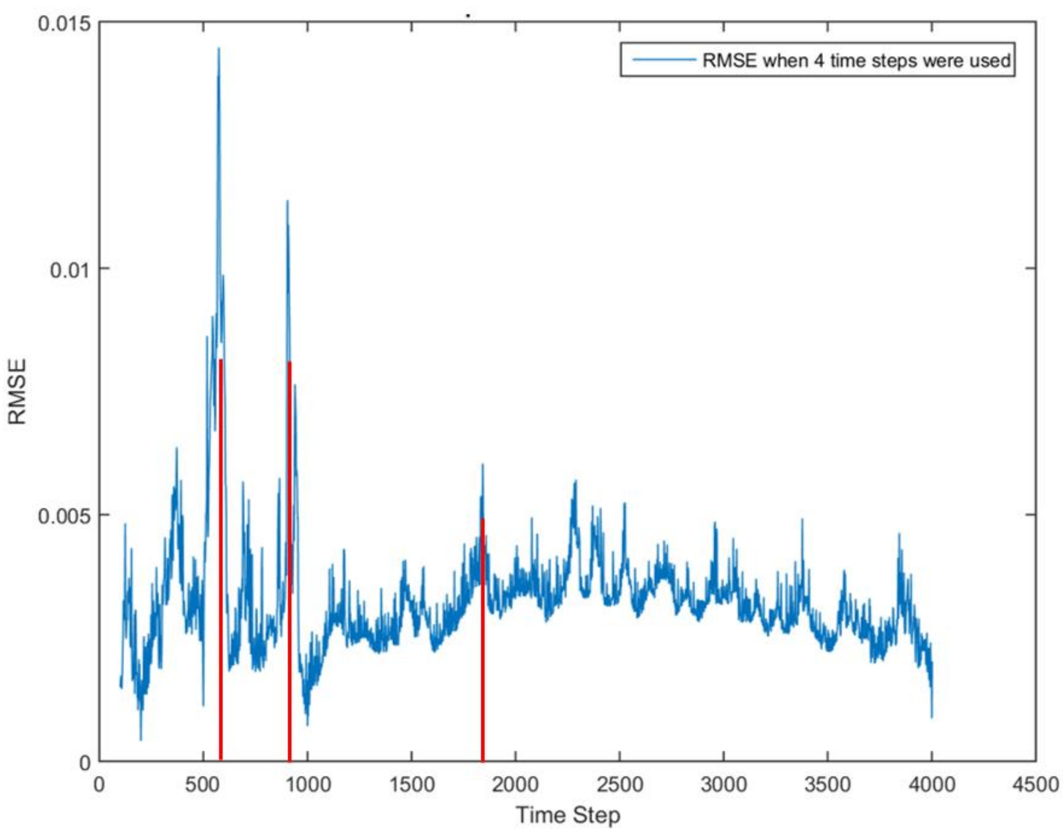


Figure 4-44 RMSE distribution in time, with three of the high RMSE values identified

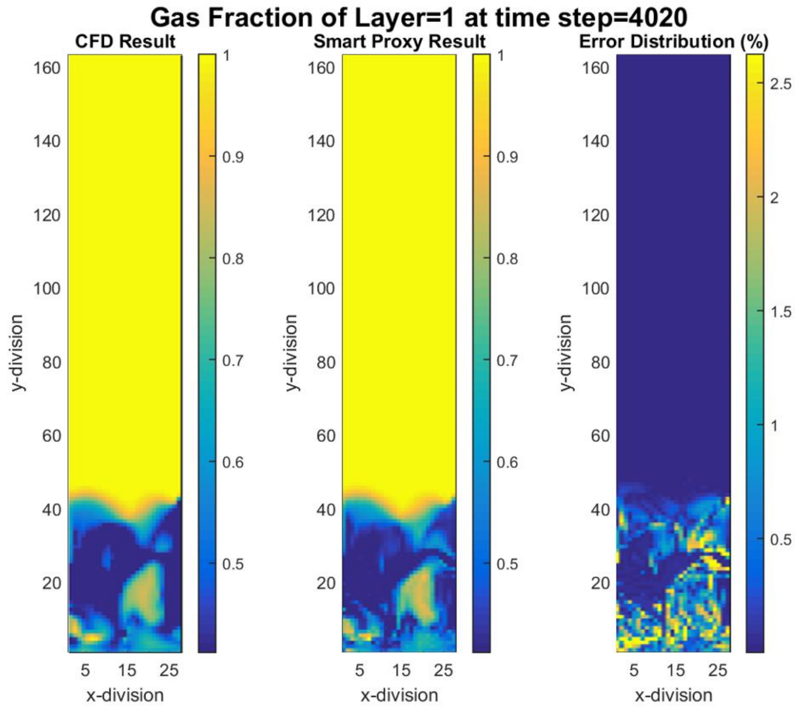


Figure 4-45 CFD and cascading smart proxy results for gas volume fraction at K = 1 cross-sectional plane

4.10 CHANGING THE DATA PARTITIONING

Another approach for reducing the input data size is partitioning of the CFD data used for training, validation and testing of the ANN. Table 4-4 shows the 4-different data partitioning method tested in this work. The results are shown in Figure 4-46 to Figure 4-48, where the error increases from about 2.2% to about 6% when the percentage of data used for training was reduced from 60% to 30%.

Table 4-4 Data partitioning in different scenarios

Data	*Training (%)*	*Calibration (%)*	*Validation (%)*
Original simulation	*70*	*15*	*15*
First attempt	*60*	*20*	*20*
Second attempt	*40*	*30*	*30*
Third attempt	*30*	*35*	*35*

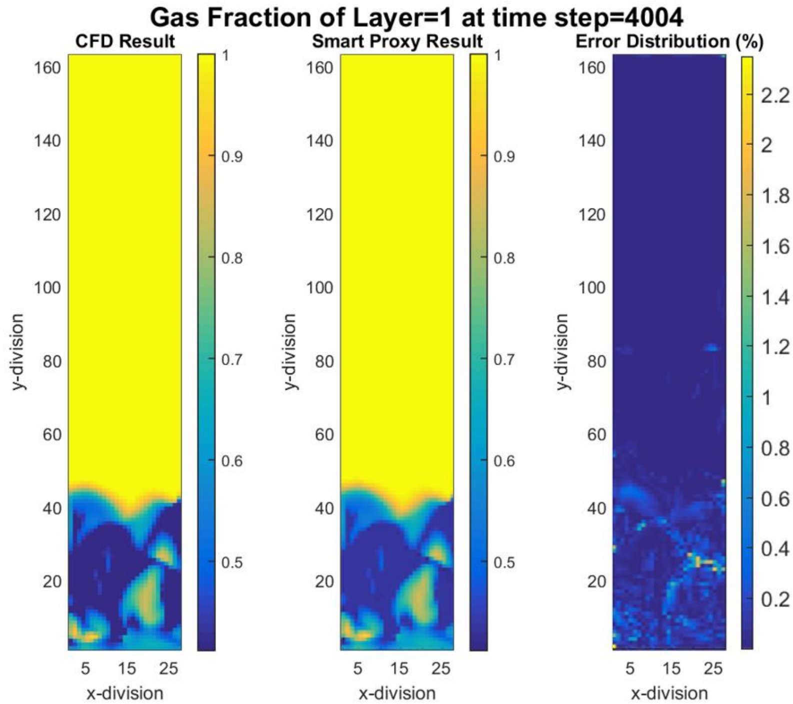


Figure 4-46 CFD and non-cascading smart proxy results for gas volume fraction at K = 1 cross-sectional plane with 60% of data used for training

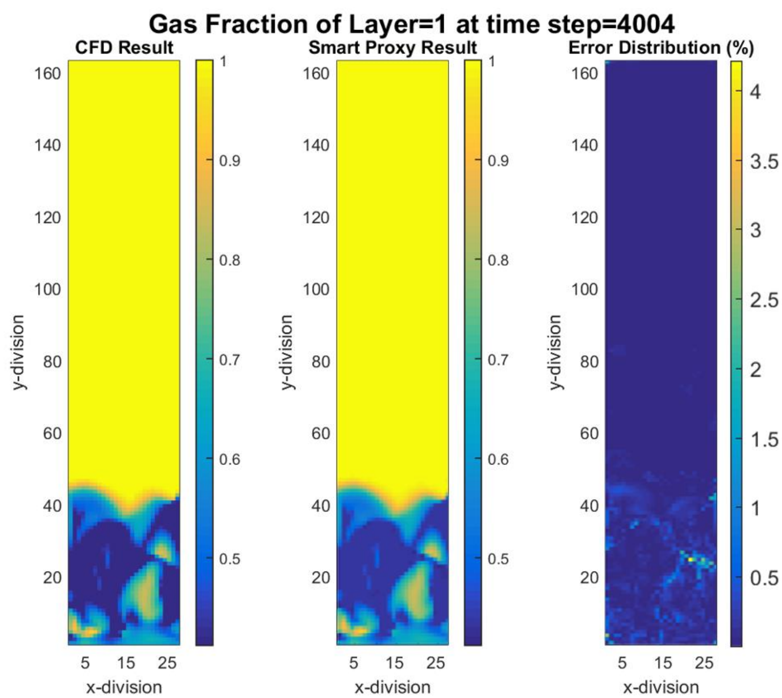


Figure 4-47 CFD and non-cascading smart proxy results for gas volume fraction at K = 1 cross-sectional plane with 40% of data used for training

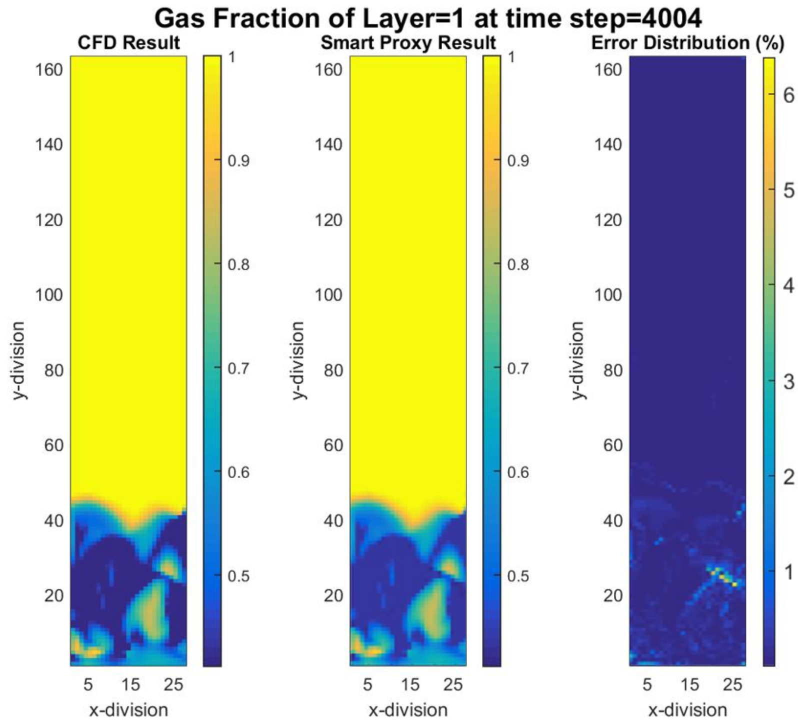


Figure 4-48 CFD and non-cascading smart proxy results for gas volume fraction at K = 1 cross-sectional plane with 30% of data used for training

4.11 SMART SAMPLING

The data from time step 4000 was used for training the ANN, using the smart sampling procedure that is outlined in section 3.4.7.3. After training is completed, the deployment starts from time step 3950 to time step 4150. Figure 4-49 shows RMSE during the deployment, with and without the smart sampling technique. It can be seen that the RMSE is comparable between the two methods. It is noteworthy that smart sampling only uses 20% of the records compare to when smart sampling is not used.

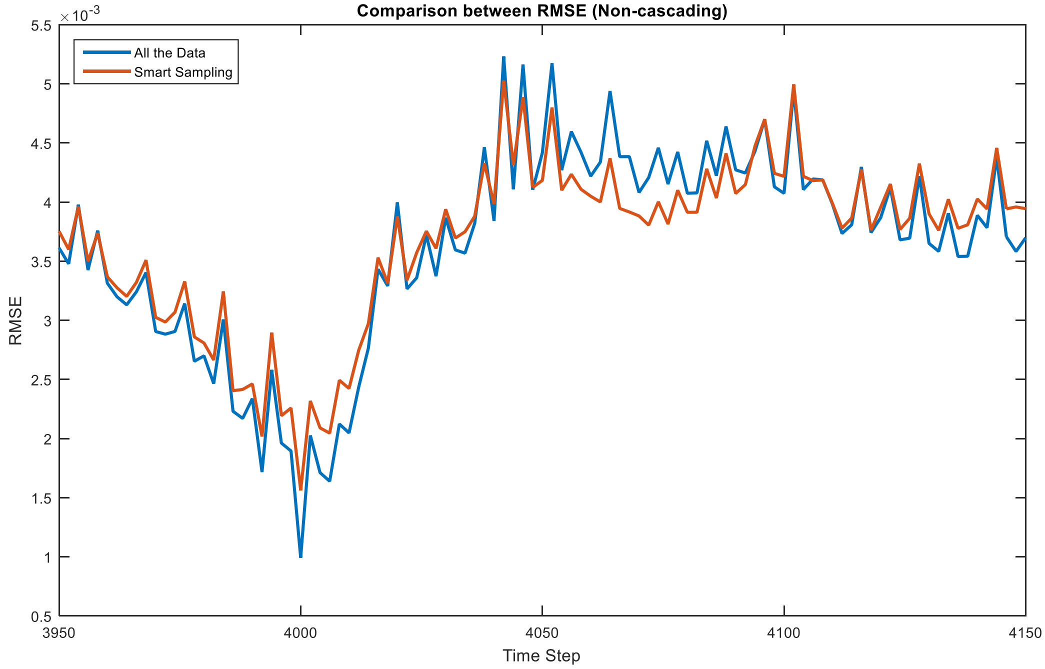


Figure 4-49 RMSE distribution in time with and without smart sampling

5. CONCLUSIONS

A data-driven smart proxy was developed to mimic the CFD simulation results of a three-dimensional fluidized bed, with a good accuracy and faster speed. Table 5-1 shows the comparison of run time of these two approaches. On average, training of ANN takes between 24 hours to 36 hours, depending on the scenario under consideration. The training time of an ANN is also strongly affected by the computer hardware. Once the smart proxy is trained, 4 seconds of simulation can be achieved in 180 seconds. This is considerably shorter compare to CFD execution time (a wall time speedup of 1440%). This study shows that machine learning and artificial intelligence can be an important tool in multiphase flow modeling and warrants further investigation.

Table 5-1 Comparison between speed of run for CFD and Smart proxy

Method	*Execution Time*
CFD	*4 seconds simulation = 3 days on 4 CPUs*
Smart Proxy	*4 seconds simulation = 180 s on 1 CPU = 3 min*

The original database (Spatio-Temporal database) included 70 parameters and 118,098 records. This database was then sent through an optimization process to get an optimized database with the smaller size but the same efficiency. In the optimization process, the size of Spatio-Temporal database was reduced more than 25 times by. This optimization was done by:

1. Reducing the number of parameters (using Key Performance Indicator)
2. Reducing the number of records (focusing on the more important cells using smart sampling)
3. Reducing the percentage of training (by Intelligent Partitioning).

The size of the optimized training data set is shown in Table 5-2

Table 5-2 Comparison between Spatio-Temporal database and optimized database

	No. Records	*No. of Parameters*	*Training Percentage*	*Total data*
Spatio-Temporal Database	*118,098*	*70*	*100*	*8,266,860*
Optimized Database	*25,827*	*43*	*30*	*333,168*

Data from time step t is used as the input to the model, while data from time step *(t+1)* is used as the output to the model (since a supervised learning is used). A model has been trained by using the original database. Blind validation of ANN during the deployment stage (use of data that was not used during the training of ANN) demonstrated that the smart proxy is able to predict the entire flow regime.

5.1 RECOMMENDATIONS AND FUTURE WORKS

This study showed that the smart proxy could be a viable tool for predicting gas-solid flow behavior is a fluidized bed. More research is needed for establishing a more efficient methodology for including more data from more time steps for training the ANN. Additionally, to make the trained ANN more general, the cascading deployment needs further research in order to minimize the error propagation over time.

6. REFERENCES

[1] Gel, A., Garg, R., Tong, C., Shahnam, M., and Guenther, C., "Applying uncertainty quantification to multiphase flow computational fluid dynamics," *Powder Technology,* vol. 242, pp. 27-39, 2013.

[2] Gel, A., Shahnam, M. and Subramaniyan, A.K., "Quantifying uncertainty of a reacting multiphase flow in a bench-scale fluidized bed gasifier. A Bayesian approach," *Powder Technology,* vol. 311, pp. 484-495, 2017.

[3] Gel, A., Shahnam, M., Musser, J, Subramaniyan, A.K. and Dietiker, J.F., "Non-intrusive Uncertainty Quantification of Computational Fluid Dynamics Simulations of a Bench-scale Fluidized Bed Gasifier," *Industrial & Engineering Chemistry Research,* no. DOI: 10.1021/acs.iecr.6b02506, 2016.

[4] Shahnam, M., Gel, A., Dietiker, J.-F., Subramaniyan, A. K. and Musser, J., "The Effect of Grid Resolution and Reaction Models in Simulation of a Fluidized Bed Gasifier through Non-intrusive Uncertainty Quantification Techniques," *ASME Journal of Verification, Validation and Uncertainty Quantification*, 2016.

[5] Fullmer, W. and Hrenya, C., "Quantitative assessment of fine-grid kinetic theory based predictions of mean-slip in unbounded fluidization," *AIChE Journal*, Vols. DOI 10.1002/aic, 62, no. 2016, pp. 11-17.

[6] Amini, S., Mohaghegh, S.D., Gaskari, R. and Bromhal, G.S., "Pattern Recognition and Data Driven Analytics for Fast and Accurate Replication of Complex Numerical Reservoir Models at Grid Block Level," in *SPE Intelligent Energy Conference and Exhibition*, Utrecht, The Netherlands, 2014.

[7] Mohaghegh, S. D., "Converting Detail Reservoir Simulation Model into Effective Reservoir Management Tools Using SRMs: Case Study - Three Green Fields in Saudi Arabia," *International Journal of Oil, Gas and Coal Technology,* pp. 115-131, 2014.

[8] Mohaghegh, S.D., Abdulla, F., Abdou, M., Gaskari, R. and Maysami, M., "Smart Proxy: An Innovative Reservoir Management Tool, Case Study of a Giant Mature Oilfield in UAE," in *Abu Dhabi International Petroleum Exhibition and Conference*, Abu Dubai, 2015.

[9] Shahkarami, A., Mohaghegh, S.D., Gholami, V., Haghighat, A. and Moreno, D., "Modeling Pressure and Saturation Distribution in a CO2 Storage Project Using a Surrogate Reservoir Model (SRM)," *Greenhouse Gas Science and Technology, Society of Chemical Industry, John Wiley & Sons, LTD,* pp. 289-315, doi:10.1002/ghg, 2013.

[10] NETL Multiphase Flow Science, "MFiX software suite," National Energy Technology Laboratory, Morgantown, WV, 2016. [Online]. Available: http://mfix.netl.doe.gov.

[11] Van der Hoef, M. A., Van Sint Annaland, M., Andrews, A. T., Sundaresan, S. and Kuipers, J. A. M., "Numerical Simulation of Dense Gas-Solid Fluidized Beds: A Multiscale Modeling Strategy," *Annual Review of Fluid Mechanics,* vol. 40, no. doi:101146/annurev.fluid.40111406.102130, pp. 47-70, 2008.

[12] Van der Hoef, M. A., Ye, M., Van Sint Annaland, M., Andrews, A. T., Sundaresan, S. and Kuipers, J. A. M., "Multiscale Modeling of Gas-Fluidized Beds," *Advances in Chemical Engineering,* vol. 31, no. 6, pp. 65-149, 2006, doi:10.1016/S0065-2377(06)31002-2..

[13] Deen, N. G., Van Sint Annaland, M., Van der Hoef, M. A. and Kuipers, J. A., M., "Review of Discrete Particle Modeling of Fluidized Beds," *Chemical Engineering Science, doi:10.1016/j.ces.2006.08.014,* vol. 62, no. 1-2, pp. 28-44, 2007.

[14] Syamlal, M., Rogers, W., O'Brien, T. J., "MFiX Documentation Theory Guide," [Online]. Available: https://mfix.netl.doe.gov/download/mfix/mfix_legacy_manual/Theory.pdf.

[15] Benyahia, S., Syamlal, M., "Summary of MFiX Equations," 2012. [Online]. Available: https://mfix.netl.doe.gov/download/mfix/mfix_current_documentation/MFIXEquations2012-1.pdf.

[16] A. Samuel, "Some Studies in Machine Learning using the Game Checkers," *IBM Journal,* vol. 3, no. 3, pp. 535-554, 1959.

[17] Alpaydin, E., Introduction to Machine Learning, The MIT Press, second edition, 2010.

[18] Haykin, S., Neural Networks - A Comprehensive Foundation, IEEE Press, 1994.

[19] Mohaghegh, S., Ameri, S., Artificial Neural Network as a Valuable Tool for Petroleum Engineers, Network, 1995.

[20] Mohaghegh, S., Arefi, R., Ameri, S., Rose, Deanna, "Design and Development of an Artificial Neural Network for Estimation of Formation Permeability," *SPE Computer Applications, doi:10.2118/28237-PA,* vol. 7, no. 6, pp. 151 - 155, 1995.

[21] Alizadehdakhel, A., Rahimi, M., Sanjari, J., Abdulaziz, A., "CFD and Artificial Neural Network Modeling of Two-phase Flow Pressure Drop," *International Communication in Heat and Mass Transfer,* vol. 36, no. doi.org/10.1016/j.icheatmasstransfer.2009.05.005, pp. 850 - 856, 2009.

[22] Esmaili, S., Mohaghegh, S., "Full Field Reservoir Modeling of Shale Assets using Advance Data-driven Analytics," *Geoscience Frontiers, Elsevier Ltd,* no. doi:10.1016/j.gsf.2014.12.006, 2015.

[23] S. D. Mohaghegh, Data Driven Reservoir Modeling, Richardson, Texas: Society of Petroleum Engineers (SPE), 2017.

[24] Kalantari-Dahaghi, A., Mohaghegh, S., Esmaili, S., "Coupling Numerical Simulation and Machine Learning to Model Shale Gas Production at Different Time Resolutions," *Journal of Natural Gas Science and Engineering,* vol. 25, pp. 380 - 392, 2015.

[25] Kalantari-Dahaghi, A., Mohaghegh, S., Esmaili, S., "Data Driven Proxy at Hydraulic Fracture Cluster Level: A Technique for Efficient CO2 Enhanced Gas Recovery and Storage Assessment in Shale Reservoir," *Journal of Natural Gas Science and Engineering,* vol. 27, no. doi:10.1016/j.jngse.2015.06.039, pp. 515 - 530, 2015.

7. APPENDIX I: GAS VOLUME FRACTION IN EARLY TIME-STEPS, NON-CASCADING, SINGLE OUTPUT, EXPLICIT

Figure 7-1 to Figure 7-12 show the contour plots for gas volume fraction as a result of non-cascading deployment at early time steps for a single output. The location of each cross-sectional plane is given in Figure 4-1.

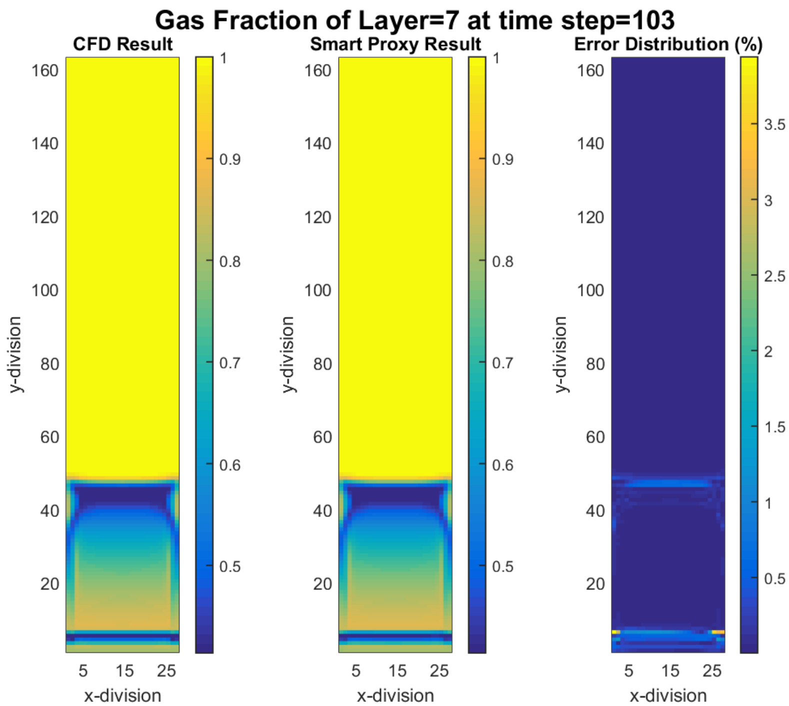


Figure 7-1 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 103

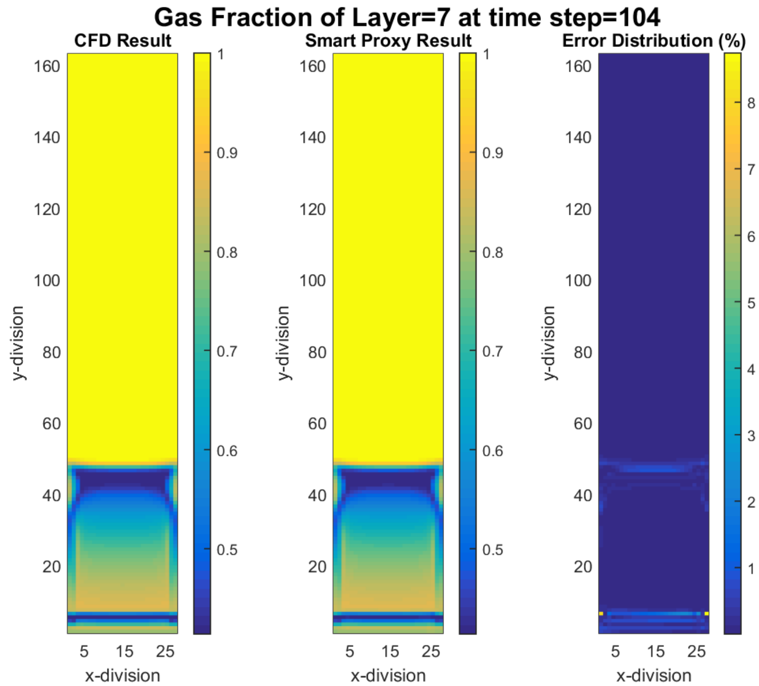


Figure 7-2 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 104

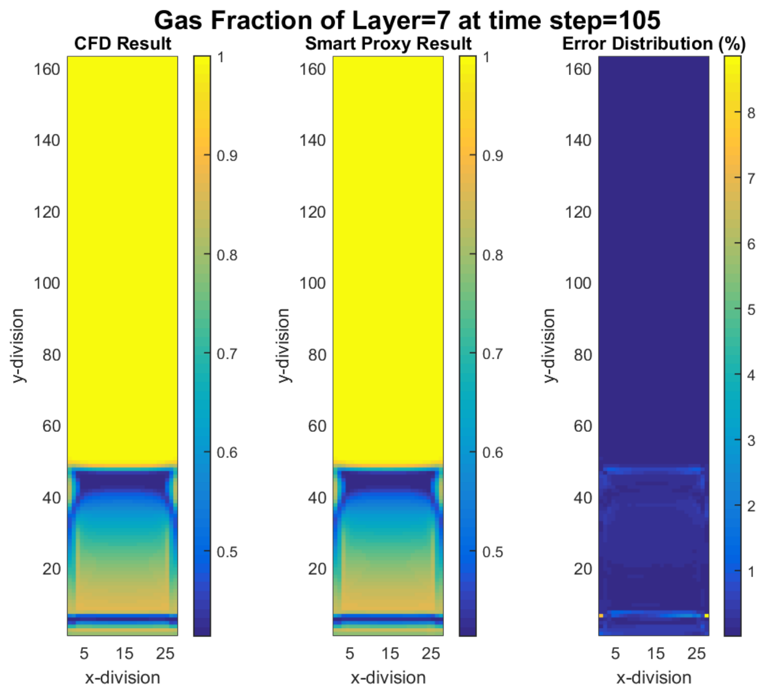


Figure 7-3 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 105

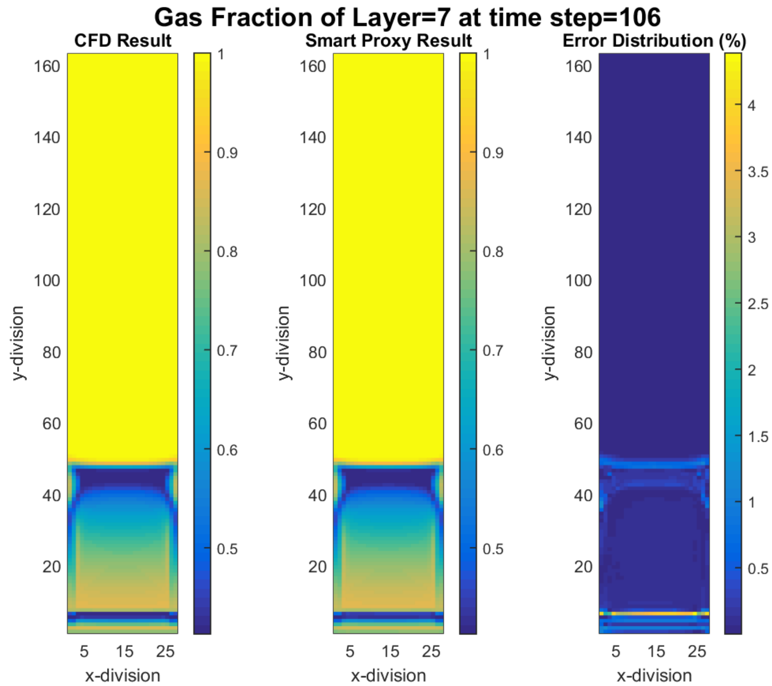


Figure 7-4 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 106

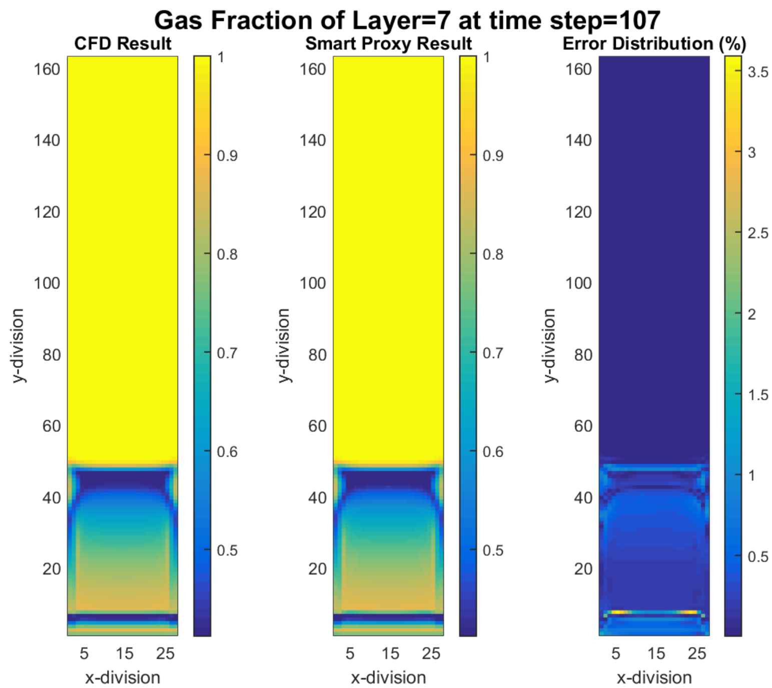


Figure 7-5 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 107

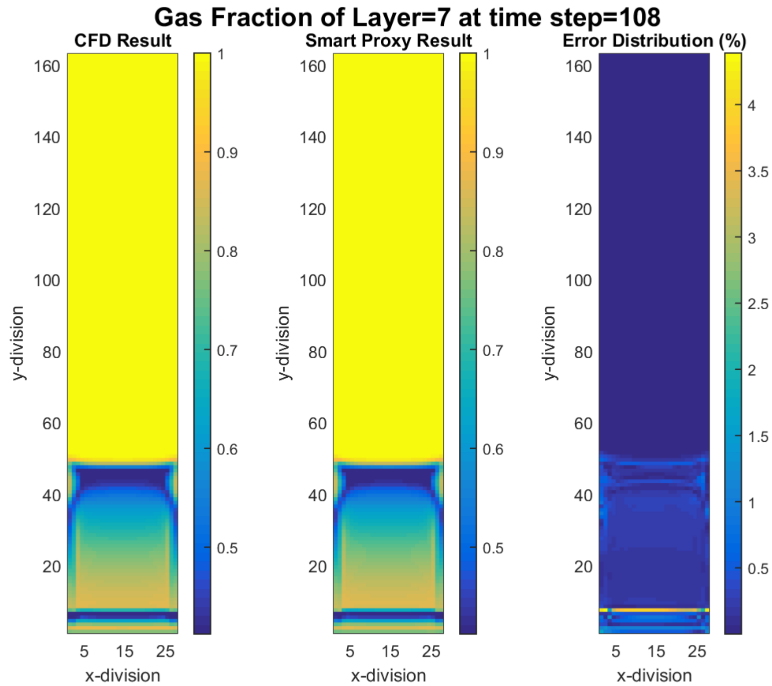


Figure 7-6 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 108

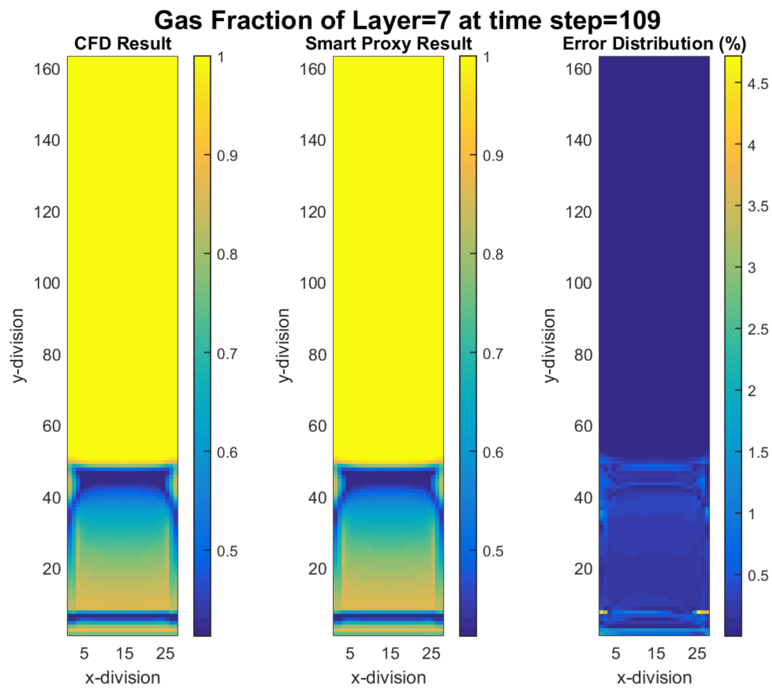


Figure 7-7 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 109

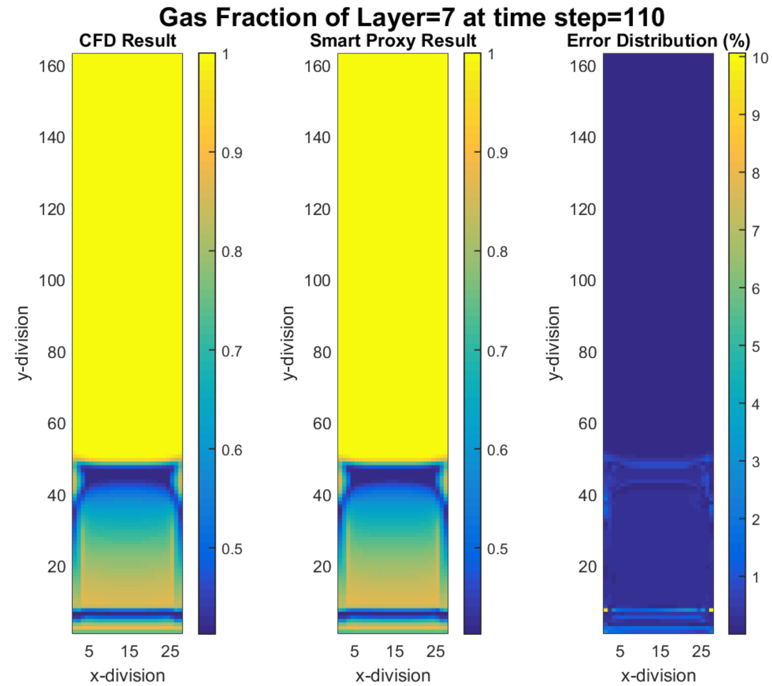


Figure 7-8 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 110

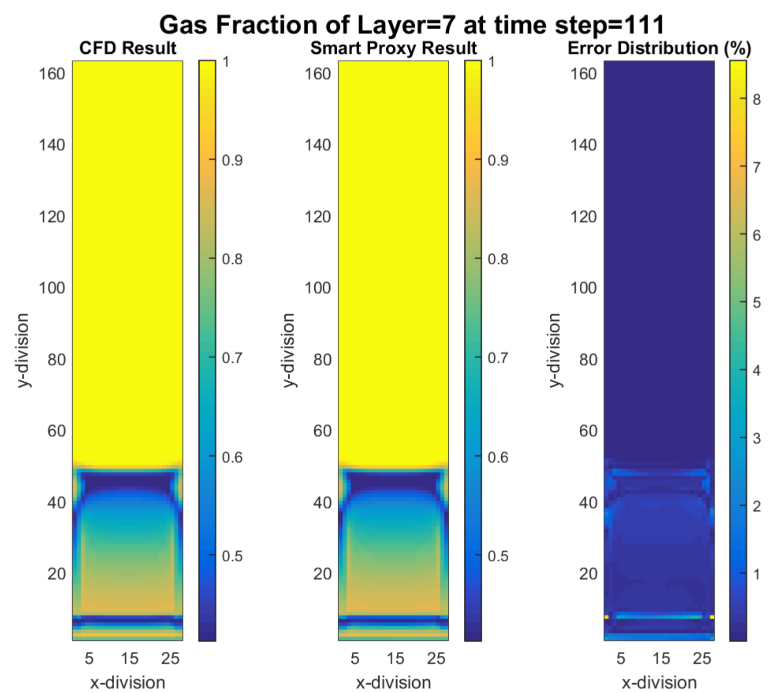


Figure 7-9 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 111

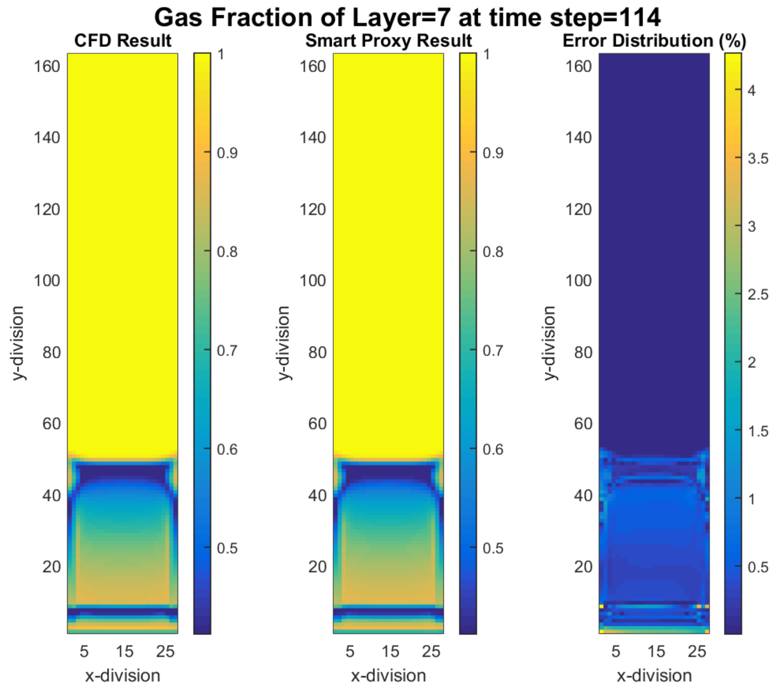


Figure 7-10 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 114

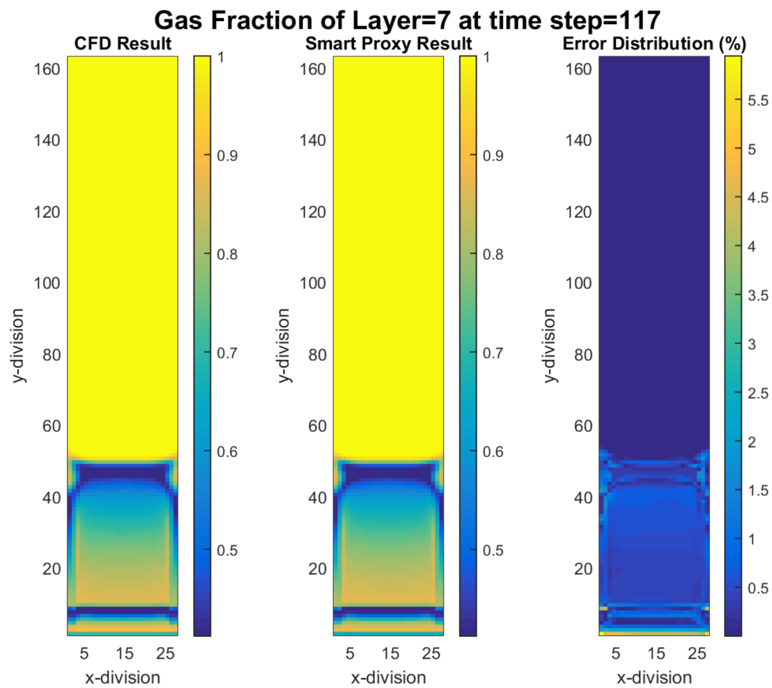


Figure 7-11 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 117

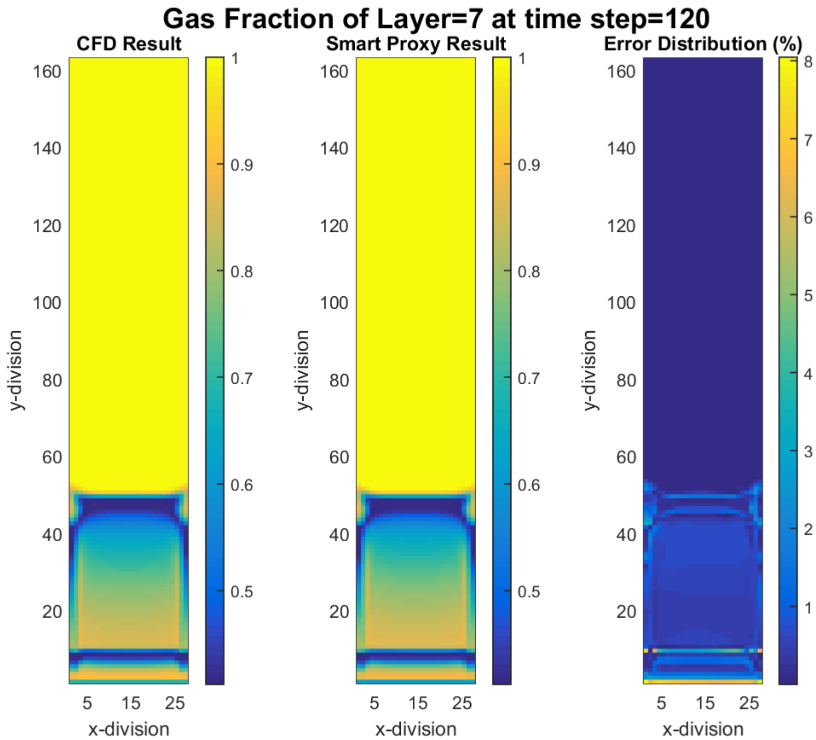


Figure 7-12 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 120

8. APPENDIX II: OTHER PARAMETERS IN EARLY TIME-STEP, NON-CASCADING, SINGLE OUTPUT, EXPLICIT

Figure 8-1 to Figure 8-11 show the contour plots for various flow variables as a result of non-cascading deployment at early time steps. The location of each cross-sectional plane is given in Figure 4-1.

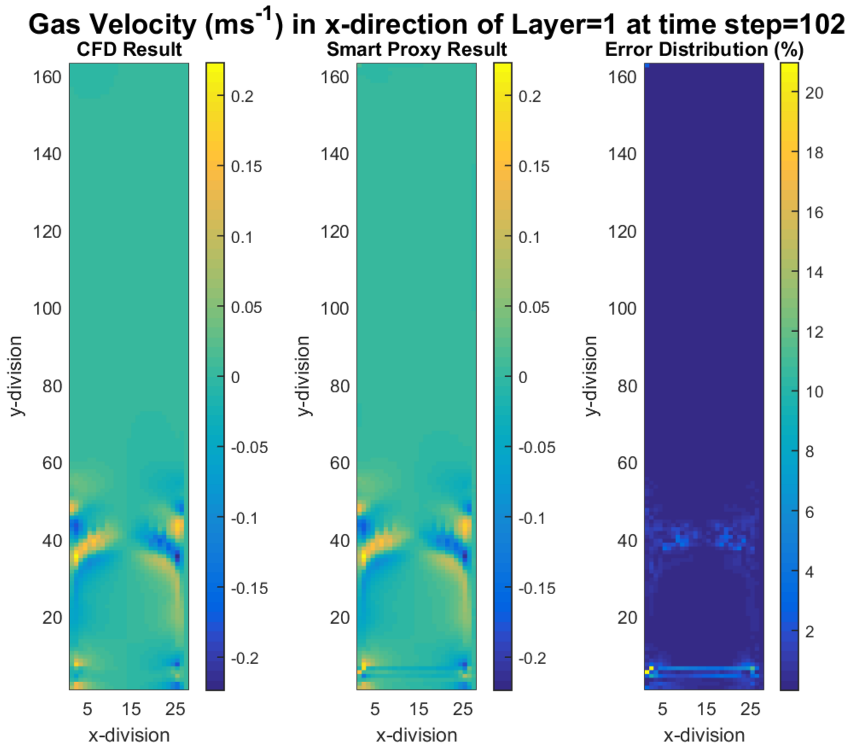


Figure 8-1 CFD and smart proxy results for x-component of gas velocity at K = 1 cross-sectional plane

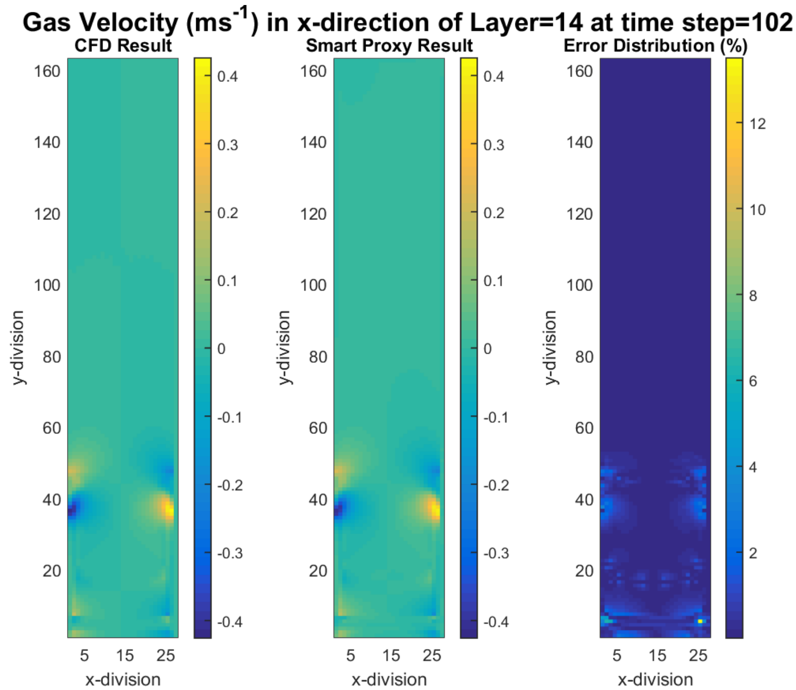


Figure 8-2 CFD and smart proxy results for x-component of gas velocity at K = 14 cross-sectional plane

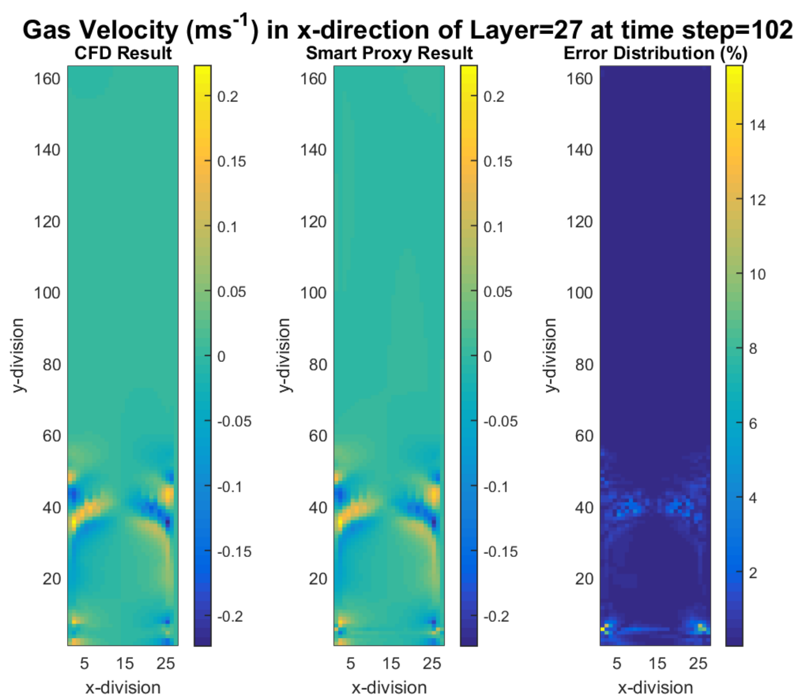


Figure 8-3 CFD and smart proxy results for x-component of gas velocity at K = 27 cross-sectional plane

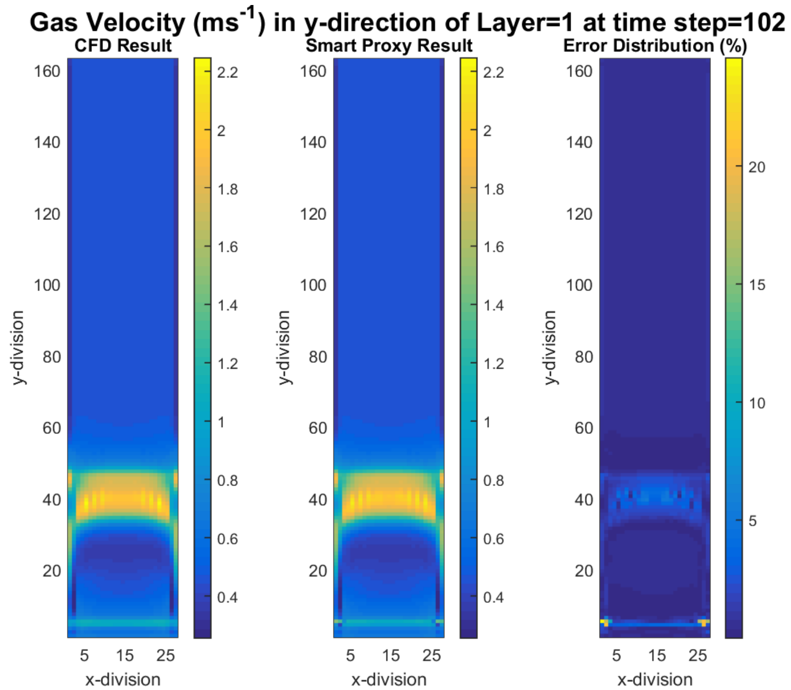


Figure 8-4 CFD and smart proxy results for y-component of gas velocity at K = 1 cross-sectional plane

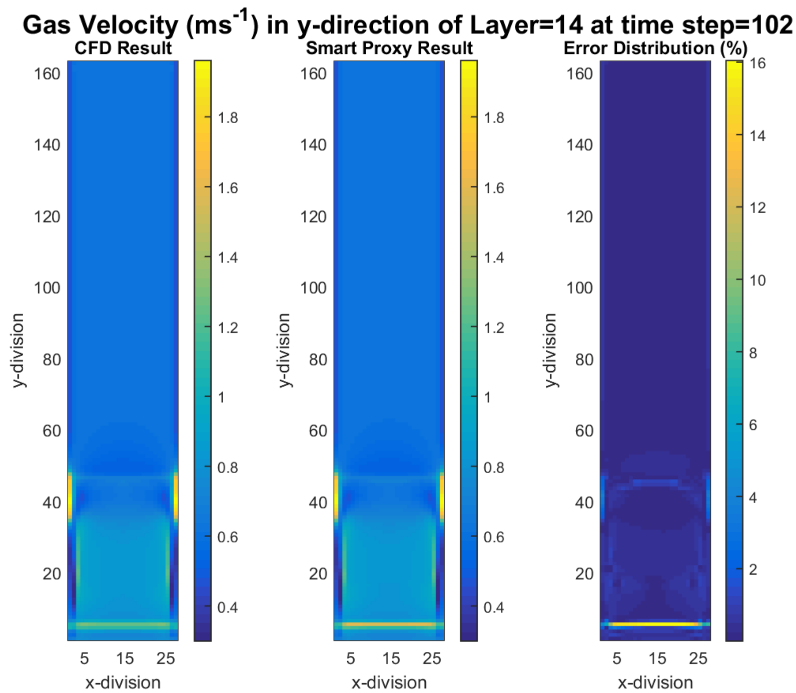


Figure 8-5 CFD and smart proxy results for y-component of gas velocity at K = 14 cross-sectional plane

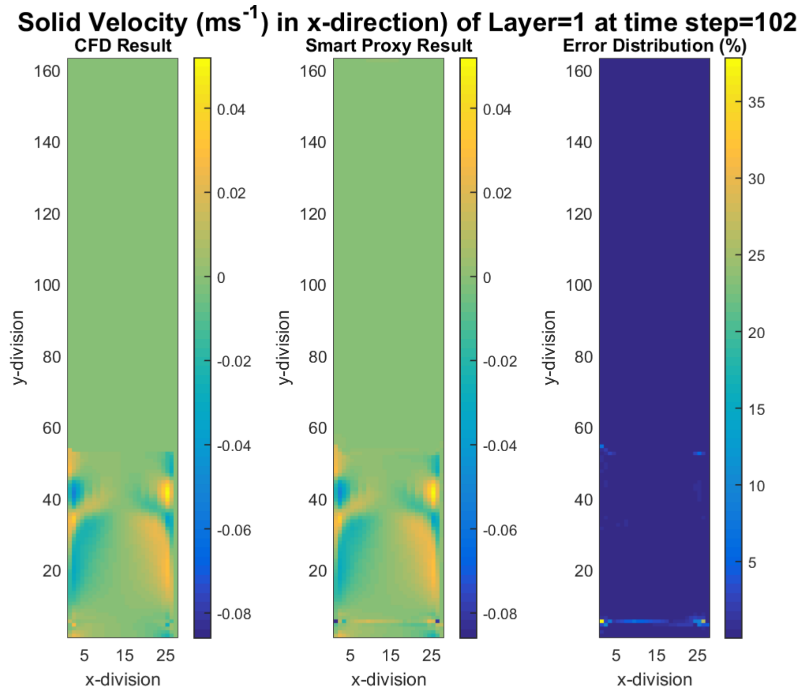


Figure 8-6 CFD and smart proxy results for x-component of solid velocity at K = 1 cross-sectional plane

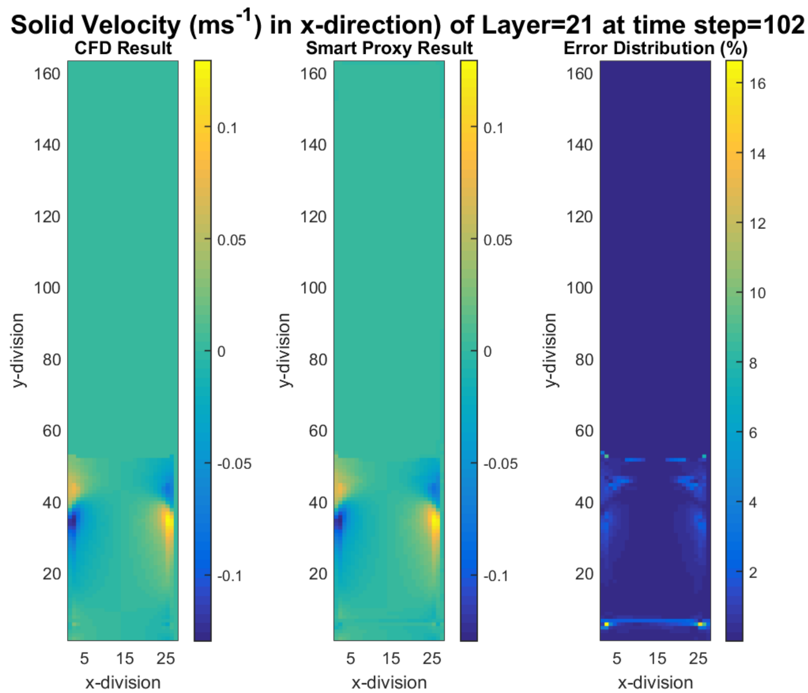


Figure 8-7 CFD and smart proxy results for x-component of solid velocity at K = 21 cross-sectional plane

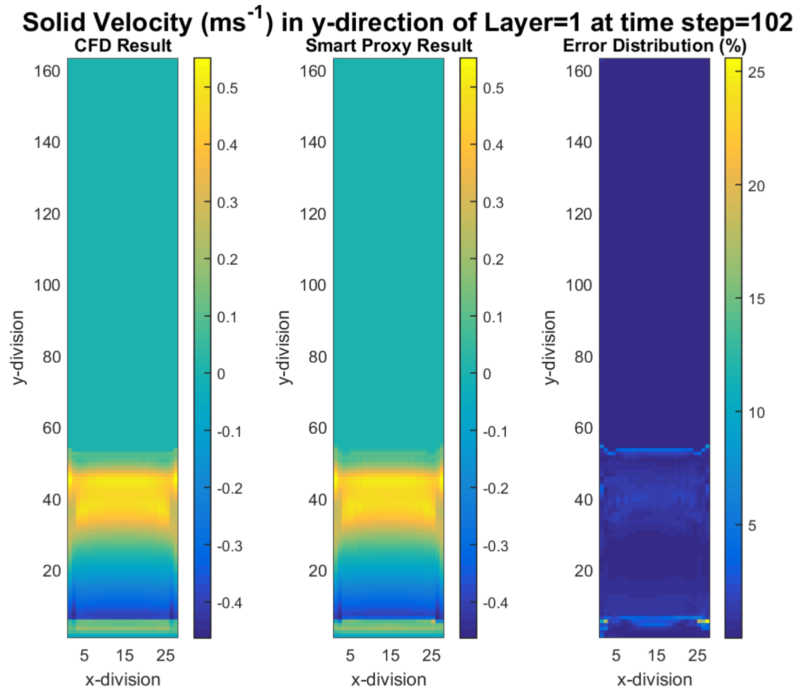


Figure 8-8 CFD and smart proxy results for y-component of solid velocity at K = 1 cross-sectional plane

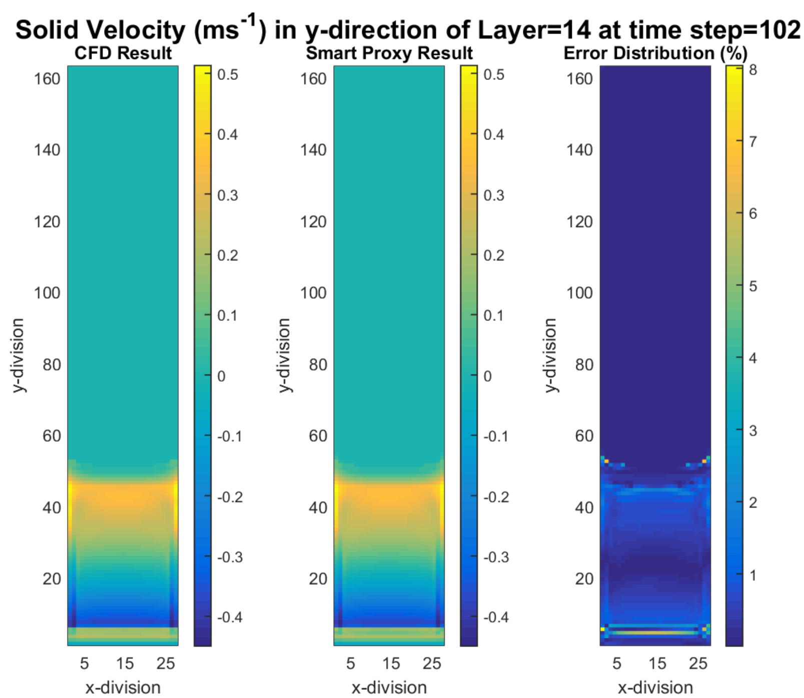


Figure 8-9 CFD and smart proxy results for x-component of solid velocity at K = 14 cross-sectional plane

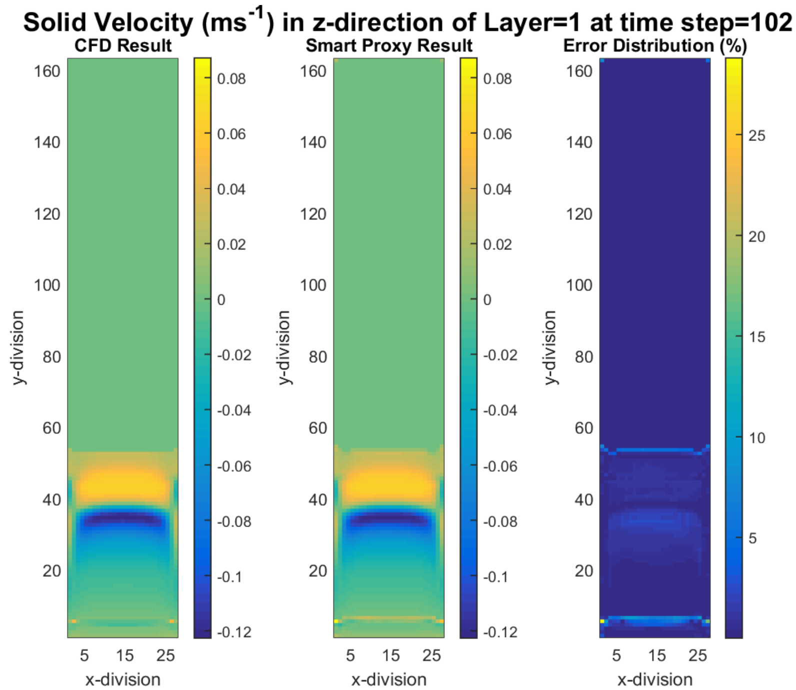


Figure 8-10 CFD and smart proxy results for z-component of solid velocity at K = 1 cross-sectional plane

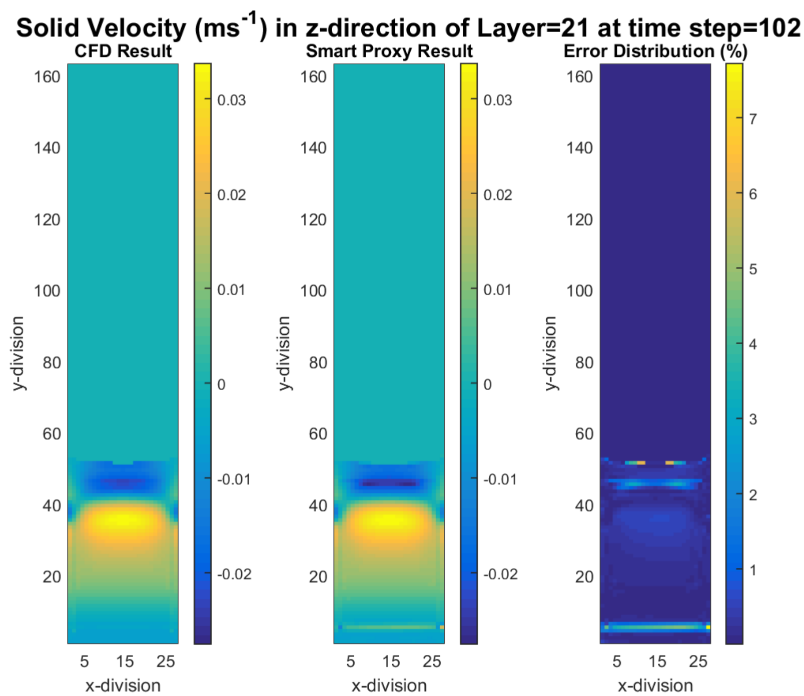


Figure 8-11 CFD and smart proxy results for z-component of solid velocity at K = 21 cross-sectional plane

9. APPENDIX III: GAS VOLUME FRACTION IN LATE TIME-STEPS, NON-CASCADING, SINGLE OUTPUT, EXPLICIT

Figure 9-1 to Figure 9-10 show the contour plots for various flow variables as a result of non-cascading deployment at late time steps for a single output. The location of each cross-sectional plane is given in Figure 4-1.

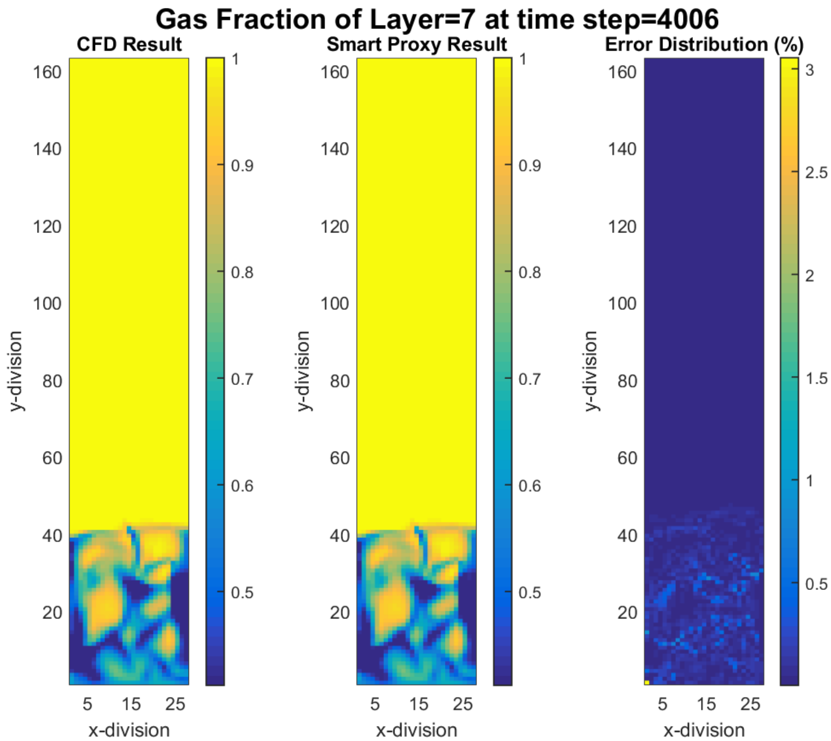


Figure 9-1 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4006

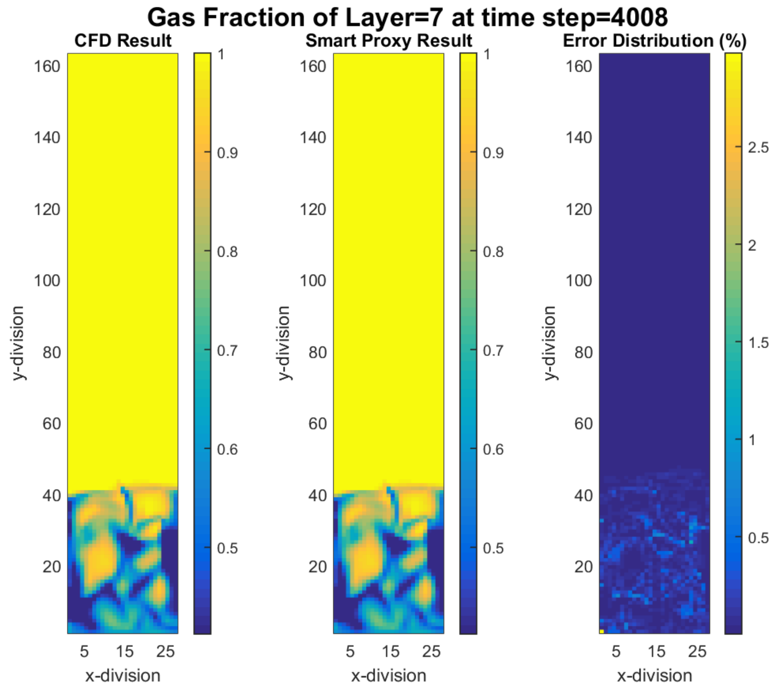


Figure 9-2 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4008

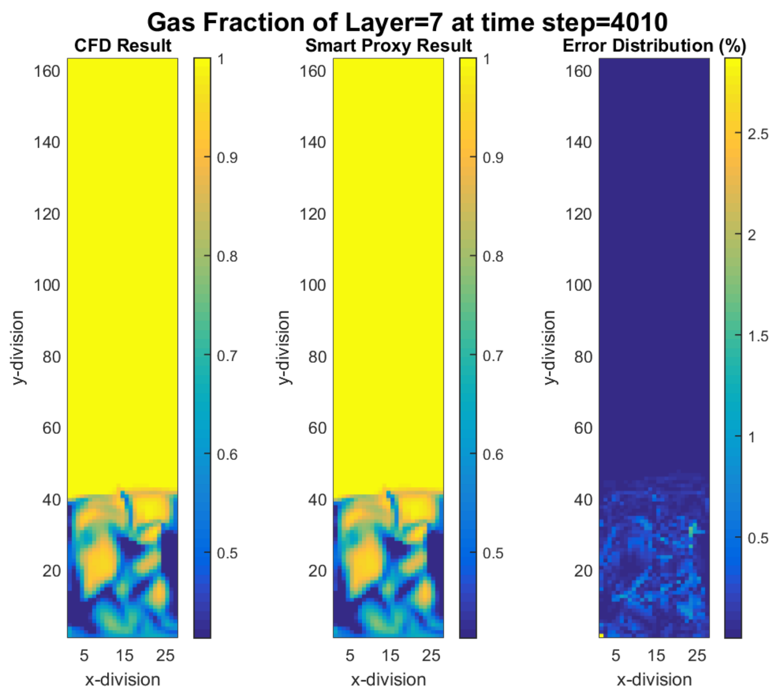


Figure 9-3 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4010

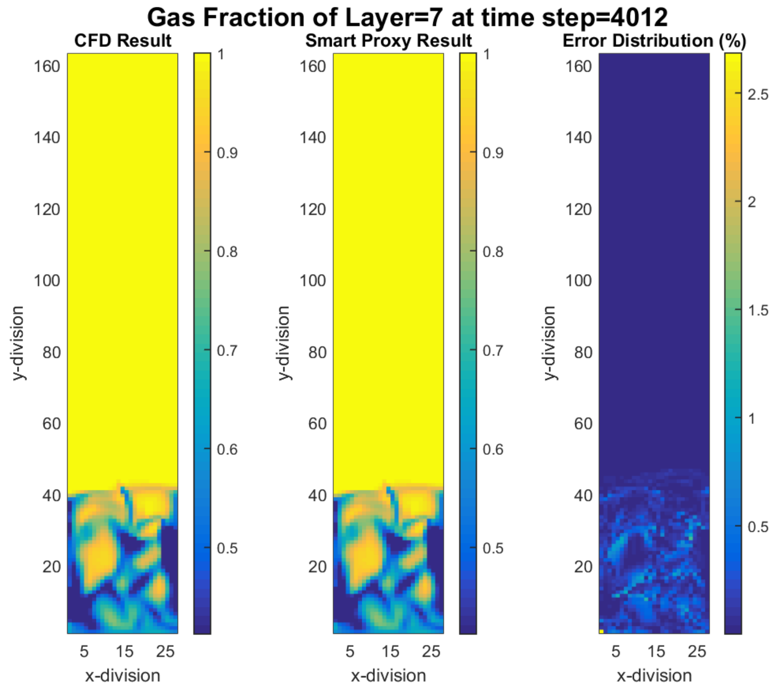


Figure 9-4 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4012

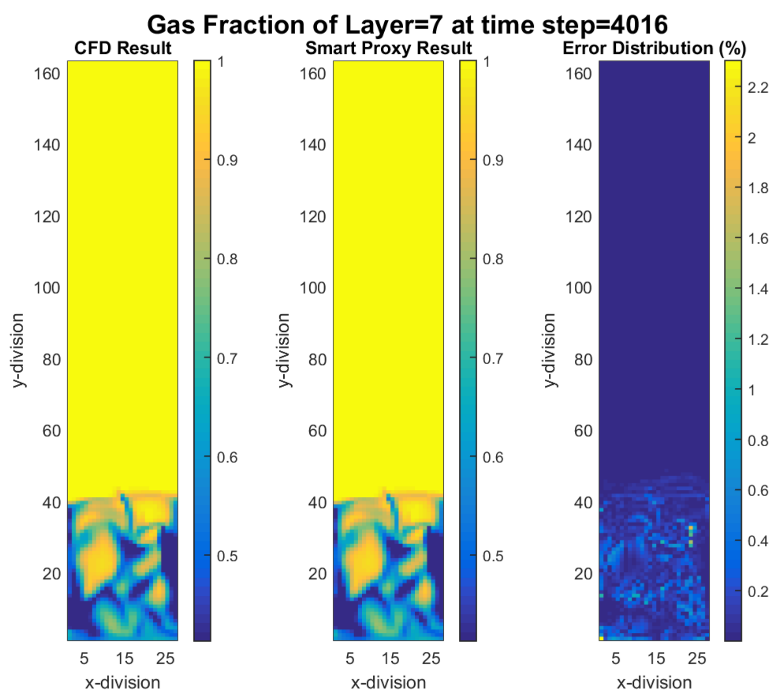


Figure 9-5 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4016

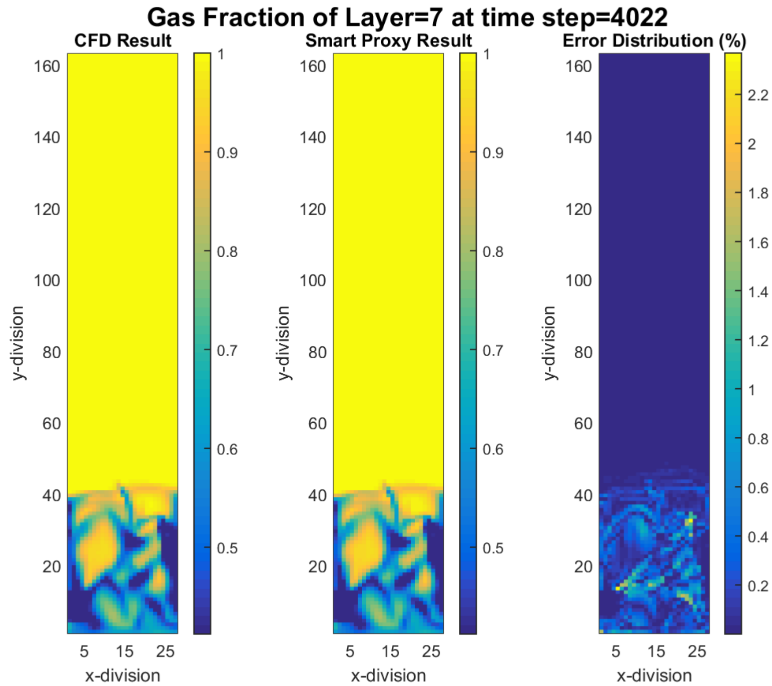


Figure 9-6 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4022

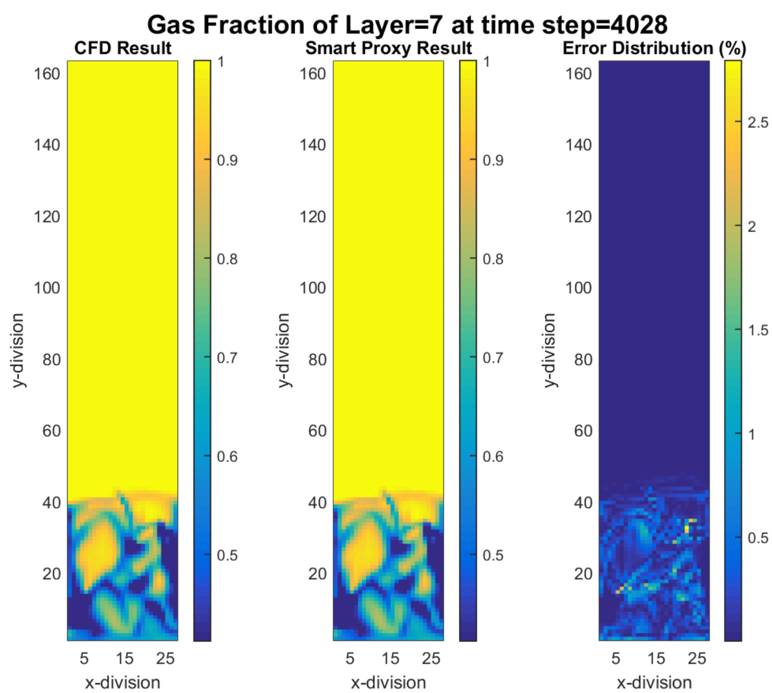


Figure 9-7 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4028

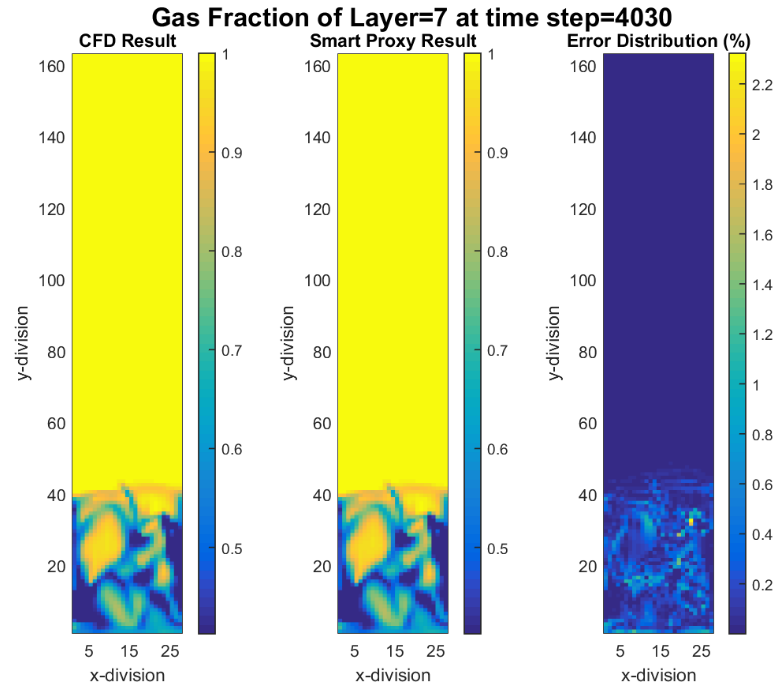


Figure 9-8 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4030

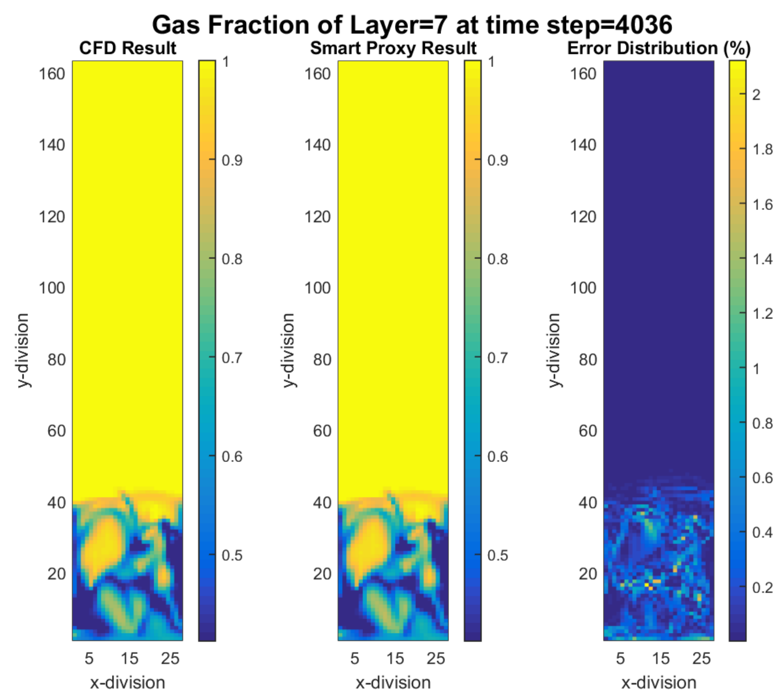


Figure 9-9 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4036

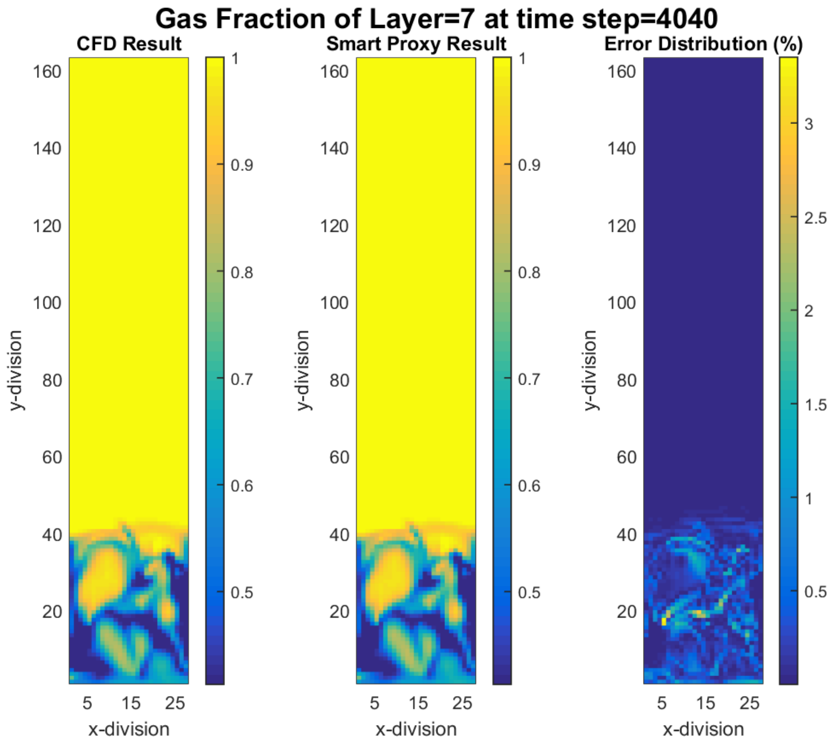


Figure 9-10 CFD and smart proxy results for gas volume fraction at K = 7 cross-sectional plane and time step = 4040

Sean Plasynski
Executive Director
Technology Development & Integration Center
National Energy Technology Laboratory
U.S. Department of Energy

John Wimer
Associate Director
Strategic Planning
Science & Technology Strategic Plans & Programs
National Energy Technology Laboratory
U.S. Department of Energy

David Alman
Executive Director
Research & Innovation Center
National Energy Technology Laboratory
U.S. Department of Energy

NETL Technical Report Series