Eleventh International Conference on CFD in the Minerals and Process Industries
CSIRO, Melbourne, Australia
7-9 December 2015

ON PRAGMATISM IN INDUSTRIAL MODELING
PART II: WORKFLOWS AND ASSOCIATED DATA AND METADATA

Josip ZORIC[1,*], Alexander BUSCH[2], Ernst A. MEESE[1], Milad KHATIBI[3], Rune W. TIME[3], Stein T. JOHANSEN[1,2], H. A. RABENJAFIMANANTSOA[3]

[1] SINTEF Materials and Chemistry, Trondheim, NORWAY
[2] Norwegian University of Science and Technology (NTNU), Trondheim, NORWAY
[3] University of Stavanger (UiS), Stavanger, NORWAY
[*] Corresponding author, E-mail address: josip.zoric@sintef.no

ABSTRACT

At CFD2014 in Trondheim, Norway a framework (FW) for pragmatic industrial modeling was suggested and demonstrated on industrial use cases with a process centric approach by Zoric et al. (2014). The present paper elaborates further on these concepts with focus on modeling and experimental data and metadata, their organization, syntax and semantics. This is exemplified on a new use case, related to drilling of oil & gas wells. The analytical problem is to produce a model which can predict the motion of a spherical particle embedded in laminar non-Newtonian flow. An overview of tasks, procedures, organization, structure and flow of data (to/from various modeling and experimental phases), required metadata, and technical and quality requirements is given. The uncertainty in the characterization of the fluid will impact the accuracy of the model predictions. The necessary uncertainty assessments and the corresponding metadata needed to support the entire modeling process are also discussed. Quantification of uncertainty is demonstrated using polynomial chaos to assess the effect of uncertainties in the parameters for the non-Newtonian viscosity model. Finally, we summarize the findings and discuss how the "pragmatism in industrial modeling" concept can help building more consistent industrial models, answering to customer needs for actual accuracy and delivery speed.

1. INTRODUCTION

Position and Motivation

Zoric et al. (2014) identified and discussed the basic concepts, processes and frameworks needed for pragmatic industrial modeling incl. terms and abbreviations. The main purpose of so-called "pragmatic modeling" is to adjust the research models to the realism/world of industrial processes, their scope, perspective and challenges. An important key element in the pragmatic modeling approach is exploitation of modeling and experimental data, organized in the optimal manner to support the prediction quality of the model. The organization of all types of data is based on "bridging" (using modeling middleware, various programming interfaces, protocols, standards etc.) between complex scientific (aspect/phenomenon oriented) physical models, simplified industrial models, process data and experimental data. In this paper we elaborate on creation, organization, flow and use of analytical information, i.e. case descriptions, modeling and experimental metadata, raw data, quality requirements and parameters, and other analytical information. By example simulations, a use case related to drilling of oil & gas wells – prediction of the motion of a spherical particle embedded in laminar non-Newtonian pipe flow, we illustrate the individual steps of the analytical process, and related flow of the data and metadata. We demonstrate examples of the necessary uncertainty assessments and the corresponding metadata needed to support the entire modeling process.

Existing research body and praxis

This work complements the references of Zoric et al. (2014) by data-centric literature, and some newer contributions on standardization of modeling FWs. Basic concepts, terms and abbreviations related to modeling frameworks are also given there.

Slotnick et al. (2014) stress the challenge of managing the vast amounts of data generated by current and future large-scale simulations and motivate to develop and maintain integrated simulation and software development infrastructures for integrated experimental testing and computational validation campaigns.

We agree with Novere et al. (2005) that most of the published quantitative models are difficult to reuse because they are either not made available or they are insufficiently characterized e.g. lack of a standard description format, lack of stringent reviewing and incomplete model descriptions. References in Lombard and Yesilbas (2006) found that almost 75% of an engineer's design work consists of seeking, organizing, modifying and translating information, often unrelated to his own personal discipline. Knowledge-based related activities consume most of a project's resources in terms of time and money Freiberg et al. (2012). Lombard and Yesilbas (2006) suggest defining a complete repository for information and knowledge management in collaborative design. When undertaking a new study, the reuse of existing models considerably reduces engineering time Malleron et al. (2011), and the possible lack of information in an existing model can be compensated for by coupling it to another model, thereby giving access to the missing information.

Varnell-Sarjeant et al. (2015) stress the importance of reuse in modeling and give an overview of reuse strategies, reusable artifacts and development constructs. Donn (2001) identifies commonality in the types of information that simulation and design tools produce.

Best Practice Guidelines (BPG) based approaches Zigh and Solis (2013); ASME (2009) have been developed to ensure as accurate CFD results as possible given the time and budget constraints. BPG complement well the modeling platforms and technologies by offering practical guidelines, procedures, rules, criteria, requirements and other knowledge artifacts. AIAA (1998) was pioneering the formalization of BPG through concise definitions of the terms verification, validation, and calibration, and by recommending related work processes. Various FW approaches are suggested in the literature, each proposing its own architecture, technology, interfacing and data structuring standards, e.g. Lombard and Yesilbas (2006); Freiberg et al. (2012); Surridge et al. (2014); Klement et al. (2014); Mooij et al. (2014); Gopsill et al. (2013); Lloyd et al. (2007); Novere et al. (2005); Hu et al. (2012). Unfortunately, they approach differently the analysis process and related interaction with FW entities (lack of industry standards). Many FW designers focus on standardizing FWs' components and technologies, e.g.: Bulatewicz et al. (2014) with the Open Modeling Interface (OpenMI) data component (based on Web services and a REST technology), Swayne et al. (2010) present a review of available tools for code/model coupling stressing a wide range of interfaces and frameworks and challenging their interconnection and collaboration. Hinsen (2011) gives a data and code model for reproducible research and executable papers, based on the Hierarchical Data Format 5 (HDF5) suitable for electronic publication and permitting the reproduction of computational results.

Many references focus on data exchange in experimental and modeling FWs and processes, and from their work we can take over good ideas and modeling elements e.g.: (a) syntax and semantics of data Gayer et al. (2010), (b) sources of metadata and data Tecplot, Inc. (2011), (c) organization of knowledge, information and data Greenwald (2004); Erdemir et al. (2012); Frenklach (2007), (d) requirements Novere et al. (2005), reusability, quality of data, validation, verification Roache (1998); Oberkampf et al. (2002); Oberkampf and Barone (2006); Babuska and Oden (2004), uncertainty estimates Russi et al. (2010); Oberkampf et al. (2002); Helton et al. (2004); Mendenhall et al. (2003); Waiters and Huyse (2002); Roy and Oberkampf (2011); Sankararaman and Mahadevan (2015), etc.

Proprietary environments such as the ANSYS Workbench Platform or modeFrontier allow for an out-of-the box connection and orchestration of various tools, whereas open-source solutions such as Dakota, NLopt, and Porto have to be coupled with the respective solvers/CFD software (SW) i.e. using C-based APIs and scripting. ANSYS Workbench allows Python-based scripting and uses XML for "Engineering Data files" and also integrates dedicated tools such as FLUENT/CFX ANSYS - CFD and MECHANICAL ANSYS - Structural. Tools such as Dakota lack the sophisticated GUI but allow for much broader and cross-platform tool integration (Jareteg (2014)). As alternative to integration into a higher framework, STAR-CCM+ offers the above mentioned functionalities with the "Optimate" Add-In (Maley (2012)).

Data pre-processing standards manifest in various file formats e.g. IGES, STEP or STL for the geometry of the case. Another generic container format in use is JT, an open and standardized lightweight format for exchange of 3D data ISO/TC 184/SC 4 (2012). For other types of pre-processing data e.g. material properties, XML seems to be the most generic file format. Data standards for post-processing are mainly based on HDF5 e.g. CGNS CFD General Notation System, Silo, or HDF5/XML. The CFD General Notation System (CGNS) might be considered the current CFD data standard as it has been adopted as a recommended practice by AIAA (2005). A coupled XDMF (for metadata) and HDF5 (for heavy data) approach has been used as an alternative in cases dealing with multiphase particle-loaded flows by Strandenes (2013) and de la Cruza et al. (2011).

Given the vast and heterogeneous tool landscape, one could require that FW implementation standards and data management standards, interfaces and protocols get more consistently standardized, and in such a way facilitate efficient and flexible interworking (combining existing and new analyses in solution of industrial problems).

Structure of this work

Section 1 of this paper introduces our position and motivation for pragmatism in industrial modeling. Section 2 takes a process view on pragmatic industrial modeling, and reflects on the interplay of practical, holistically organized and orchestrated activities as explained in Zoric et al. (2014). Section 3 takes a data-centric view and focuses on creation, flow, exchange and use of modeling metadata, raw data and information in various phases of pragmatic modeling. Section 4 addresses organization and structure of modeling metadata and data. We offer a system analysis and discuss the candidate implementation technologies and standards. Section 5 discusses our experiences with a practical use-case driven exercise and suggests future steps and improvements.

2. PROCESS VIEW

Pragmatic industrial modeling requires a structured approach not just with respect to models, simulations, experiments, information and data, but also regarding analytical processes, concluding with a well-structured communication of the results and their analytical and validity context. We consider these important elements being parts of the analytical framework (FW) of Zoric et al. (2014), illustrated in Figure 1.

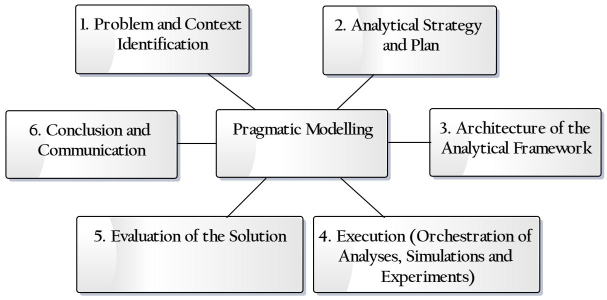

Figure 1: Important phases, processes and results in a typical pragmatic analysis acc. to Zoric et al. (2014).

Zoric et al. (2014) detail these modeling phases, and the results they produce. It is crucial to conclude pragmatic analysis by a communication of the analytical results (Step 6 in Figure 1). It is important to relate the analysis to its context: (a) important analytical parameters (assumptions, initial, boundary conditions, domain-specific parameters etc.), (b) information about modeling scale, (c) accuracy of the proposed solution, (d) estimates of representability, (e) predictive power, (f) computing and experimental resource consumption, etc. This information is needed both for the evaluation of existing models/analyses/experiments, and for their future reuse. One could even require more stringent standardization, in order to facilitate efficient collaboration, reuse and interworking (combining existing and new analyses in solution of various industrial problems).

3. DATA-CENTRIC VIEW

This section takes a data-centric view on the pragmatic modeling process sketched in Figure 1. Industrial models pose requirements such as: (a) industrial scope and perspective, (b) usefulness, (c) required accuracy and predictability, (d) simplicity of use, (e) response time and speed, (f) compatibility with other (industrial) models, etc., which were mapped to numerous SW engineering requirements in Zoric et al. (2014). We would like to motivate further development and standardization of SW engineering related to pragmatic modeling frameworks, including data and metadata:

- Standardization of data exchange formats (e.g. based on XML, JSON, HDF5 and other standards).
- Standardization of modeling metadata specifying e.g. problem definition, analytical context, accuracy, sensitivity and quality of models and simulations (organized as a part of the analytical context container), syntactic and semantic description of the data entities, information and knowledge, etc. discussed below.
- Standardization of data processing, management and retrieval functionality.

4. ORGANIZATION & STRUCTURE OF DATA

Pragmatic modeling is often part of a complex analytical and/or design process and therefore often conducted as team work and driven by analytical workflows and described by Zoric et al. (2014). Solution architects specify a set of data/information exchange standards, protocols and interfaces, and design tools. Designers and analysts generally verify the results by use of various models, analyses, experiments and fine-tuning techniques (e.g. sensitivity analyses evaluated against physical experiments).

Figure 2 illustrates the overall organization of modeling data and information on a "per-case" basis (i.e. a project divided in hierarchically organized subcases with related data/information sets), and based on the case study described in section 5. The grey area describes the work process and the white part describes associated data structures.

The work process starts with a comprehensive specification of variables and parameters as part of a detailed case description derived from the modeling customers' needs created by the solution architect. Besides the quantities of interest, this includes numerical values of all parameters, dimensions, ranges of validity, boundary conditions, fluid and particle properties, etc. This input data serves as specification for one or multiple simulation and experimental setups. Subsequent simulation and experimental runs produce solution data. In the particular industry case described in section 5 the experimentally obtained raw data using PIV had been transferred by image processing techniques to the final solution data format. By comparing the simulation result with the experimental result, and calibrating the model if necessary, the model is validated.

Simulation and experimental activities generally share some instances of associated data, e.g. the above mentioned specification and case description. For validation purposes they must also share the same class of solution data (Results) as part of the specification such as e.g. velocity profiles/fields and particle trajectories in our use case. But as opposed to the specified input data, the particular solution data sets are specific to the two approaches in the sense that one will end up with e.g. one particle trajectory as a result of the simulation and a second one as a result of the experiment. Only the data structure of the solution data is inherited from the case description data class.

Both simulations and experiments also produce their specific sets of data: During setup of the CFD case one usually has to deal with data related to e.g. grid properties, physical models and numerical schemes, whereas in the experimental case one has to deal with data related to e.g. rig information such as test section, inlet channel, pump and measurement instruments.

The very same setup of data classes is then being used over and over again for multiple test runs with varying parameters accounting for different operating/design points of the system under consideration ideally following some sort of Design-of-Experiments (DoE) approach. In order to account for natural scatter one usually also performs multiple experimental test runs for the same set of input parameters.

Figure 2 illustrates the workflow and SW engineering view on modeling. We will detail it by describing a dedicated example in Section 5. We use a data-centric view to discuss the requirements and SW Engineering issues related to model interaction, data/information exchange interfacing, standardization and other important elements for design of pragmatic models. We follow an industry case and its modeling, experimental and analysis parts, and follow creation, usage and flow of data and information.

Our process is structured as a data-focused interview of participants during the six analytical phases (Figure 1), by means of templates for description of use-cases, metadata lists (attribute-value pairs), and descriptions of other data types

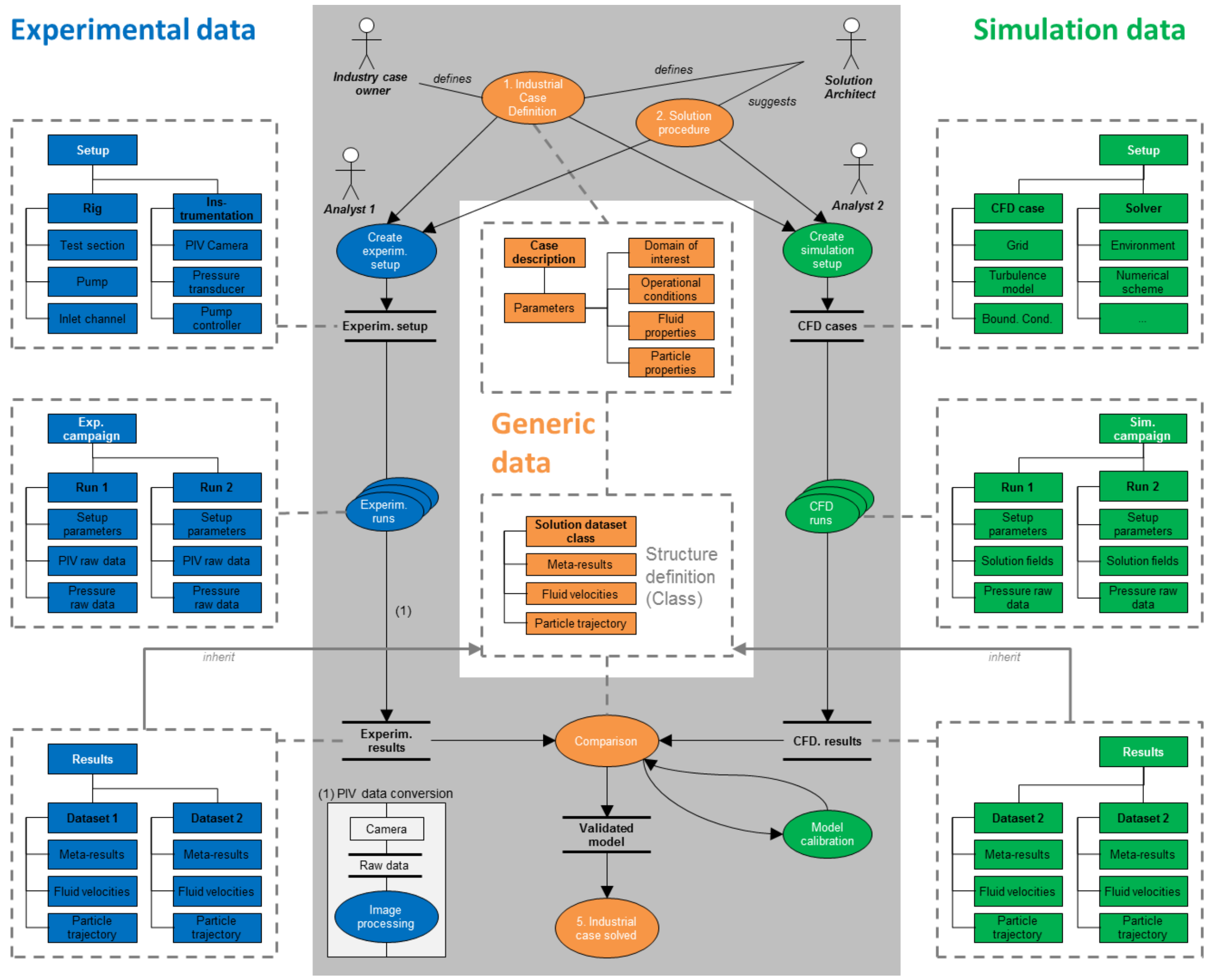

Figure 2: Organization of data flow and structure in the case study (section 5), see also Figure 1 and Zoric et al. (2014).

(hyperlinks, raw-data containers, figures, documents etc.). Some attribute-value-pairs are, as examples, indicated by colored references i.e. "[1]" where the respective attribute-value-pair is given in Table 1with a color code based on the data structure of Figure 2.

Table 1: Attribute-value pairs

[#]	Attribute	Value
[1]	Application	Cuttings transport in drilling wellbores
[2]	Fluid	H2O; PAC
[3]	Fluid density	1000 kg/m³
[4]	Particle diameter	1; 2; 3 mm
[5]	Particle density	2560 kg/m³
[6]	Unknown 1	Particle trajectory
[7]	Unknown 2	Point of impingement
[8]	Overall accuracy	5 %
[9]	Delivery time	1 Month
[10]	Concentration	2; 4 g/L
[11]	Rheology model	Power-Law
[12]	Uncertainty	1-3 %
[13]	Uncert. anal.	Polynomial Chaos: Xiu (2010)
[14]	Assumptions	Incomp.; Steady; Isothermal; Laminar
[15]	Camera	SpeedCam MiniVis
[16]	Frame rate	2500fps
[17]	Resolution	512 x 512 pixels
[18]	Fluid bulk vel.	0.042; 0.085 m/s
[19]	Impingement	5 mm
[20]	Calibration 1	Spatial dimensions = 3
[21]	Calibration 2	v_{x0} = 0.119 m/s; v_{y0} = -0.476 m/s
[22]	Uncert. interval	-0.09 … -0.47 m/s
[23]	Uncert. interval	-0.03 … -0.06 m/s
[24]	Uncert. interval	-0.01. -0.06 m/s
[25]	Further work 1	Redesign the particle inlet geometry…

5. CASE STUDY

Adequate cuttings transport in drilling wellbores [1] is a major challenge. This case study is limited to a sub-problem that might become an element in a larger and more complex model describing the physics of cuttings transport in a wellbore. The task of this sub-problem is to predict motion and particle trajectories in a non-Newtonian, laminar flow through a horizontal channel according to Figure 3.

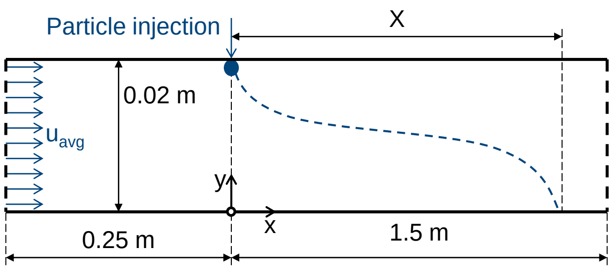

Figure 3: Particle settling domain.

The fluid is either Newtonian (Water) or non-Newtonian (Water with Poly-Anionic Cellulose added) [2] with a density of 1000 kg/m³ [3]. A spherical particle [4], [5] is released at the centerline of the top wall. Besides predicting the particle trajectory [6] there is a particular interest in the position X at which the particle impinges on the lower horizontal wall [7]. Further major requirements are that X is predicted by certain accuracy [8] and that the model is available within certain time [9].
Both experimental and CFD activities will be part of the solution approach acc. to Figure 2.

Fluids rheology & uncertainties

The rheology of the two PAC solutions [10] is represented by a power-law model for the effective viscosity [11] according to Eq. (1).

$$\mu = A\dot{\gamma}^{n-1} \quad (1)$$

Here $\dot{\gamma}$ is the shear rate, while the coefficients A and n are obtained from rheometer measurements of Time and Rabenjafimanantsoa (2012), which represents a data flow from a previously conducted experimental result to the CFD model parameters (See Figure 2: Experimental runs, results, model calibration). The uncertainties in such measurements should always be stored along with the results [12]. However, as indicated by Figure 4, the largest uncertainty here is that the chosen power-law model does not represent a perfect fit to the rheometric data.

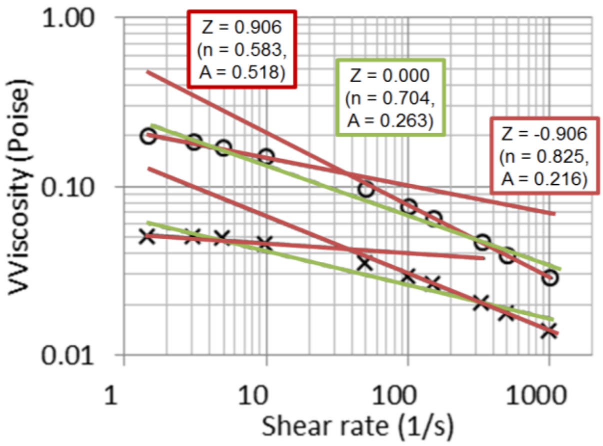

Figure 4: Rheology of PAC 2 (×) and PAC 4 (o); green line is best fit, red lines are extremes to assess uncertainty.

To assess the uncertainty introduced by the power-law model, we vary the number of points that are used to fit the power law to the data obtaining a one-parameter family of power laws. The parameter value Z is then uncertain and is regarded as a uniformly distributed random variable. Following the pseudospectral collocated generalized polynomial chaos approach [13], an approximation to the particle path can be found through the series

$$[x(t,Z), y(t,Z)] = \sum_{k=0}^{n} [\hat{x}(t), \hat{y}(t)] L_k(Z) \quad (2)$$

where L_k is the k-th order Legendre polynomial. The coefficients $\hat{x}$ and $\hat{y}$ are found by exploiting the property that the Legendre polynomials are orthogonal under the uniform distribution Z and approximating the orthogonality integral through a high order quadrature. For this case study we have opted to use five points in the quadrature, resulting in five simulations per fluid to find the approximate solution to the stochastic problem. It is paramount to the accuracy of the method to use the zeros of the 5th order Legendre polynomial in this case. For PAC 4 the parameters corresponding to these zeros are given in Figure 4. Once the parameters in the expansion (2) are found it is trivial to extract mean value, standard deviation (SD), or even the probability density function for the particle path.

Particle trajectories

Numerical simulation

Numerical simulations were conducted with FLUENT using an Eulerian-Lagrangian concept. The fluid flow model is simplified based on various assumptions [14]. Using the Discrete Phase Model (DPM), the particle is subject to gravity and drag only and is coupled to the fluid just one-way. In order to obtain fully-established velocity profiles the entrance length of the computational domain was increased to 2 m. Initially a 2D and finally a 3D Cartesian staggered grid of 1 mm cell size were found to produce grid-independent solutions after 300 iterations. Numerical schemes used were the SIMPLE method for pressure-velocity coupling, implicit backwards Euler for temporal discretization and QUICK (Leonard) for spatial discretization.

Experimental work

Experimental studies were conducted using a small-scale flow loop where the test section (1.5 m long, in Figure 3) is transparent in order to obtain particle trajectories using PIV (Particle Imaging Velocimetry) measurement techniques. This was done using a high-speed camera [15] with a frame rate of 2500 fps [16] and a resolution of 512 x 512 pixels [17]. Subsequent data processing involved an in-house Matlab code based on a time-resolved digital PIV tool developed by Thielicke and Stamhuis (2015).

Comparison of num. and exp. preliminary results

Both simulations and experiments were conducted for multiple combinations of fluid inlet velocities [18], fluid types [2], [10] as well as particle diameters [4].

Figure 5 shows experimentally and numerically obtained trajectories for H_2O, where Y is the height- and X is the length-coordinate of the flow channel.

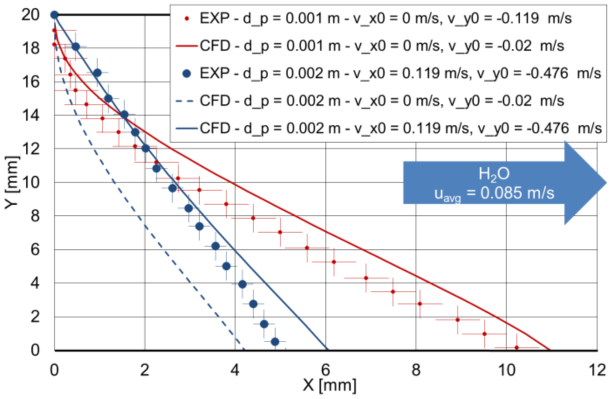

Figure 5: Trajectories for H_2O; ±SD

The smaller particle is being transported further downstream resulting in an experimentally obtained impingement of 10.34 mm (red circles). The numerically obtained impingement position X = 11.03 mm (solid red line) is just outside the SD of the experimentally obtained value. The bigger particles settle faster [19] (blue circles). The experimentally obtained particle trajectories for H_2O show different shapes very close to the injection point (x = 0 mm & Y = 20 mm). Zooming in here reveals, that the trajectory for the larger particle shows a remarkably larger initial x-velocity-component compared to the smaller particle's trajectory. Figure 5 shows two different numerically obtained trajectories for the larger particle: One where the initial particle velocity is very small (dashed blue line), and one where the initial particle velocity was calibrated according to experimental data (solid blue line) with the respective numerically obtained values outside the experimental SD.

In the case of the two PAC fluids, multiple simulation runs using the different combinations of model coefficients obtained with the Polynomial Chaos Approach were conducted. Figure 6 shows experimentally and numerically obtained trajectories for PAC, with the Polynomial Chaos results using 2D CFD represented by the reddish area for the d_p = 2 mm case. Because the respective experimental trajectory (red circles) shows an impingement at X=107.4 mm, a second CFD run using a 3D domain [20] was conducted; the results of which could reproduce the experimental result within its SD (red dashed line).

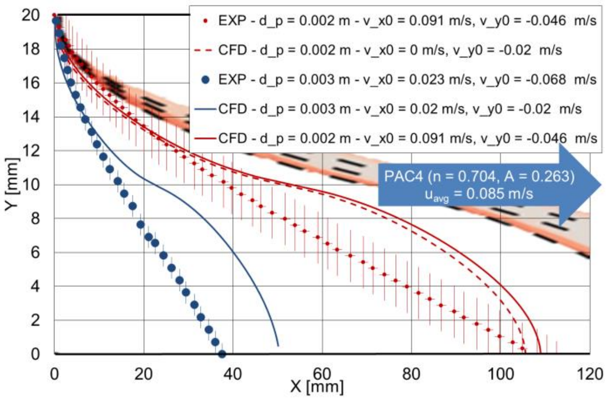

Figure 6: Trajectories for PAC 4; ± 3SD

As with the H_2O case, the CFD model was calibrated regarding the initial particle velocities (red solid line). Bigger particles (d_p = 3 mm) settles faster with an impingement at X=36.09 mm (blue circles). The respective numerical trajectory is far off the experimental SD (blue solid line).

Discussion of preliminary results

For both the Newtonian and non-Newtonian case, the experimentally obtained trajectories indicate different initial particle velocities. A rough estimate using the gradient of the first two sample points of each experimental series revealed some considerable scatter. The numerical values were then used as improved initial particle velocity in the simulations [21]. This lead to slightly improved predictions in the case of H_2O (Figure 5, solid blue line) but not on the PAC case (Figure 6, solid red line).

The initial particle velocities seem to have a major influence on the trajectories in the Newtonian case (H_2O). In the non-Newtonian case (PAC) with a high effective viscosity the impact of the initial velocity is significantly smaller than for H_2O. Thus, initial particle velocities are a major uncertainty in the CFD model with approximate uncertainty intervals depending on the fluid type; H_2O [22], PAC 2 [23] and PAC 4 [24].

Another major uncertainty is the fit of the power-law rheology model to the experimental data. Even though the 2D CFD results based on the power-law coefficients obtained with the Polynomial Chaos Approach do not result in correct trajectories, the results could be used to estimate the effect of the rheology model on the impingement accuracy.

For the lower part of the channel, using the PAC fluids, a significant mismatch between the trajectories is found; numerically obtained trajectories show an S-shape, whereas experimentally obtained trajectories show a more-or-less straight line. The difference in the trajectories in the lower half of the channel is considered to be due to the fact that with the current modeling approach the local flow field of the particle is unresolved. Hence, the computed fluids viscosity, purely based on the main fluid flow velocity and the associated shear-rate, will not be sufficient to represent the local viscosity experienced by the particle. As the particles fall through the liquid, we expect that increased local shear rate will reduce the viscosity further, and thus reduce the drag force and enhance the fall speed.

Based on the preliminary results the following further work is specified:

- Redesign the particle inlet geometry in order to reduce natural scatter of initial x- and y-velocity components (direction and magnitude) and associated uncertainty [25].
- Generate a more sophisticated model which accounts for local shear conditions around the particle and respective viscosity in order to improve the shape of the computed trajectory.

- Use a more sophisticated rheology model (e.g. Cross, instead of power-law) reduce the associated uncertainty.
- Use the Polynomial Chaos Approach on other important but uncertain parameters such as the flow rate.
- Finalize 3D simulations and compute an overall uncertainty of the model by i.e. using the validation metric of Oberkampf and Barone (2006) on both the impingement distance and the trajectory.

Framework design considerations

Following a data-centric approach resulted in some interesting experiences that we discuss in this section.

Preparing a case description for the database: The project team understood and supported the documentation of the case using templates and structured metadata (Figure 2 and Table 1) and the need for an iterative improvement process (four iterations). It was beneficial to have six participants in the case: one solution architect guiding the organization and structure of the case and providing the templates for description of the data and information and five analysts: two modelers, two experimentalists and one person dealing with uncertainties. During the project some analysts occasionally showed the tendency to take experience-based shortcuts and use the quickest way to conclude some phases, without describing the whole procedure. Here the templates proved beneficial as they forced the analysts to document their approaches and findings. Different mind-sets of "modelers" and "experimentalists" will always challenge the multi-expertise projects. It always takes some time to establish "a common dictionary". It is important that the agreed dictionary becomes a part of the case description (dictionary metadata table). We suggest using two levels of dictionaries (lookup tables): a *domain specific dictionary*, which is not part of the particular use-case data structure as it is of generic nature, and a *case-specific dictionary*, which is an integral part of the case data structure as it includes the case-specific definitions, terms, abbreviations, symbols, etc. While the first may have thousands of terms, the second one should be limited just to the nomenclature used in the case. That will also allow for some case-specific definitions. The templates were useful for good description of the case, and in particular media-to-metadata mapping tables (each information/media item, e.g. text, image, schema, formula was accompanied by a descriptor metadata-table with attribute-value keys, i.e. Table 1). These metadata tables were later used to insert the information in the database tables. Metadata is integral part of the queries for later retrieval of cases. We are evaluating several solutions for mapping the geometry descriptions, various design information items, (e.g. technical drawings), formulas (equations sets) and other heterogeneous media to their metadata representatives. They are important for future automation of the case management and retrieval.

Retrieval of the cases from the database: We anticipate the need for several levels of queries: (a) queries for the case (based on the domain-specific metadata), (b) search for subcases and their descriptors (by use of case-specific metadata) and finally (c) search for the case results and various raw data.

Organization of the data/information: It is important to cater for differences in topology, structure and organization of the cases, i.e. ensure the standardization of data/information and metadata containers, while allowing the flexibility in their hierarchy and relations. Our ambition goes definitely towards a database management system, if possible in a standard way, including interworking with major analytical (experimental and modeling) tools and FWs.

6. DISCUSSION AND CONCLUSION

Assessment of the uncertainty in predictions is a critical part of pragmatic modeling. We categorize the uncertainty into two classes: (1) Known uncertainty, where the uncertainty lies into model input and numerics and is possible to be assessed with well-defined procedures. (2) Unknown uncertainty, where the uncertainty lies into the model concepts and physics. Only new experiments can help to reduce this uncertainty.

The presented case study illustrates the importance of addressing the a priori unknown uncertainty. This is done by exploitation of experimental data as an integral part of pragmatic modeling. Because of the experimental data the numerical model was improved with respect to the initial velocities and the constitutive equation of the fluid. Also vice versa, the experimental setup received improvement because the simulations revealed the importance of the initial particle velocity. In addition, it was clarified that the physical model for the drag-force is unsatisfactory for the non-Newtonian fluids tested, as significant deviations between experimental and theoretical particle trajectories were found. This allows us to go back and improve the physical model in order to reduce the uncertainty.

We believe that almost all the six phases in a typical industrial (pragmatic) analytical process illustrated in Figure 1 can to some extent be standardized. We can standardize the structure of the processes and the data as e.g. in Figure 2 and Table 1 as well as the tools that are used, the quality assurance methods, and how the input, output / results and analytical context are presented and described. In particular the standardization of data types and flows is of major importance and both a challenge and major opportunity. Establishing a template-based database including all data, metadata, dictionaries and findings will enhance the interoperability and reuse of models as well as consistency of various models and/or experiments. This paper tries to contribute to these ideas, by taking the metadata and data / information perspective and providing some considerations how such a solution could look like.

ACKNOWLEDGEMENTS

The project *Advanced Wellbore transport Modelling (AdWell)* with its sponsor, the Research Council of Norway and its partners Statoil, GDF SUEZ E&P Norge, IRIS, UiS, NTNU and SINTEF are gratefully acknowledged for funding and supporting part of this work.

REFERENCES

AIAA, (2005), "AIAA R-101A-2005 Recommended Practice - The CFD General Notation System - Standard Interface Data Structures." American Institute of Aeronautics and Astronautics.

AIAA, (1998), "Guide: Guide for the Verification and Validation of Computational Fluid Dynamics Simulations (AIAA G-077-1998(2002))." American Institute of Aeronautics and Astronautics, Inc., Reston, VA.

ANSYS - CFD, http://www.ansys.com/Products/Simulation+Technology/Fluid+Dynamics [Accessed June 9, 2015].

ANSYS - STRUCTURAL, http://www.ansys.com/Products/Simulation+Technology/Structural+Analysis [Accessed June 9, 2015].

ANSYS WORKBENCH PLATFORM, http://www.ansys.com/Products/Workflow+Technology/ANSYS+Workbench+Platform [Accessed June 9, 2015].

ASME, (2009), "Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer: ASME V V 20." American Society of Mechanical Engineers (ASME).

BABUSKA, I., AND ODEN, J. T., (2004), "Verification and validation in computational engineering and science: basic concepts." *Computer Methods in Applied Mechanics and Engineering* 193:4057–4066.

BULATEWICZ, T. ET AL., (2014), "A distributed data component for the Open Modeling Interface." *Environmental Modelling & Software* 57:138–151.

CFD GENERAL NOTATION SYSTEM, http://cgns.sourceforge.net/ [Accessed June 9, 2015].

DE LA CRUZA, R. ET AL., (2011), "Implementing a XDMF/HDF5 Parallel File System in Alya."

DAKOTA, https://dakota.sandia.gov/ [Accessed June 9, 2015].

DONN, M., (2001), "Tools for quality control in simulation." *Building and Environment*:673–680.

ERDEMIR, A. ET AL., (2012), "Considerations for reporting finite element analysis studies in biomechanics." *Journal of Biomechanics* 45:625–633.

FREIBERG, M. ET AL., (2012), "Extensible Prototyping for pragmatic engineering of knowledge-based systems." *Expert Systems with Applications* 39:10177–10190.

FRENKLACH, M., (2007), "Transforming data into knowledge—Process Informatics for combustion chemistry." *Proceedings of the Combustion Institute* 31:125–140.

GAYER, M. ET AL., (2010), "CFD modelling as an integrated part of multi-level simulation of process plants: semantic modelling approach." *In* Proceedings of the 2010 Summer Computer Simulation Conference pp. 219–227. Society for Computer Simulation International.

GOPSILL, J. A. ET AL., (2013), "A Social Media framework to support Engineering Design Communication." *Advanced Engineering Informatics* 27:580–597.

GREENWALD, M., (2004), "Beyond benchmarking—how experiments and simulations can work together in plasma physics." *Computer Physics Communications* 164:1–8.

HELTON, J. C. ET AL., (2004), "An exploration of alternative approaches to the representation of uncertainty in model predictions." *Reliability Engineering & System Safety* 85:39–71.

HINSEN, K., (2011), "A data and code model for reproducible research and executable papers." *Procedia Computer Science* 4:579–588.

HU, K. ET AL., (2012), "A Scalable Infrastructure for Online Performance Analysis on CFD Application." *Chinese Journal of Aeronautics* 25:546–558.

ISO/TC 184/SC 4, (2012), "ISO 14306:2012 - Industrial automation systems and integration - JT file format specification for 3D visualization." International Organization for Standardization (ISO).

JARETEG, A., (2014), "Coupling of Dakota and OpenFOAM for automatic parameterized optimization." Chalmers University of Technology.

KLEMENT, M. ET AL., (2014), "Biochemical Space: A Framework for Systemic Annotation of Biological Models." *Electronic Notes in Theoretical Computer Science* 306:31–44.

LLOYD, S. ET AL., (2007), "Integrative Biology — the challenges of developing a collaborative research environment for heart and cancer modelling." *Future Generation Computer Systems* 23:457–465.

LOMBARD, M., AND YESILBAS, L. G., (2006), "Towards a framework to manage formalised exchanges during collaborative design." *Mathematics and Computers in Simulation* 70:343–357.

MALEY, K., (2012), "STAR-CCM+ / Optimate." *In* Singapore.

MALLERON, N. ET AL., (2011), "On the use of a high-performance framework for efficient model coupling in hydroinformatics." *Environmental Modelling & Software* 26:1747–1758.

MENDENHALL, M. R. ET AL., (2003), "Best practices for reduction of uncertainty in CFD results." *In* 41st AIAA Aerospace Sciences Meeting.

MODEFRONTIER, http://www.esteco.com/modefrontier [Accessed June 9, 2015].

MOOIJ, W. M. ET AL., (2014), "Serving many at once: How a database approach can create unity in dynamical ecosystem modelling." *Environmental Modelling & Software* 61:266–273.

NLOPT, http://ab-initio.mit.edu/wiki/index.php/NLopt [Accessed June 9, 2015].

NOVERE, N. L. ET AL., (2005), "Minimum information requested in the annotation of biochemical models (MIRIAM)." *Nat Biotech* 23:1509–1515.

OBERKAMPF, W. L. ET AL., (2002), "Error and uncertainty in modeling and simulation." *Reliability Engineering & System Safety* 75:333–357.

OBERKAMPF, W. L., AND BARONE, M. F., (2006), "Measures of agreement between computation and experiment: Validation metrics." *Journal of Computational Physics* 217:5–36.

PORTO, https://github.com/NanoSim/Porto [Accessed June 9, 2015].

ROACHE, P. J., (1998), "Verification and Validation in Computational Science and Engineering." Hermosa Publishers.

ROY, C. J., AND OBERKAMPF, W. L., (2011), "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing." *Computer Methods in Applied Mechanics and Engineering* 200:2131–2144.

RUSSI, T. ET AL., (2010), "Uncertainty quantification: Making predictions of complex reaction systems reliable." *Chemical Physics Letters* 499:1–8.

SANKARARAMAN, S., AND MAHADEVAN, S., (2015), "Integration of model verification, validation, and calibration for uncertainty quantification in engineering systems." *Reliability Engineering & System Safety* 138:194–209.

SILO, https://wci.llnl.gov/simulation/computer-codes/silo [Accessed June 9, 2015].

SLOTNICK, J. ET AL., (2014), "CFD Vision 2030 Study: A Path to Revolutionary Computational Aerosciences." National Aeronautics and Space Administration (NASA), Hampton, Virginia.

STRANDENES, H., (2013), "Implementation of an Efficient Storage Strategy for Fluid-Particle Simulations in OpenFOAM." *In* NTNU, Trondheim.

SURRIDGE, B. W. . ET AL., (2014), "A framework for coupling explanation and prediction in hydroecological modelling." *Environmental Modelling & Software* 61:274–286.

SWAYNE, D. ET AL., (2010), "Linking Data, Models and Tools: An Overview."

TECPLOT, INC., (2011), "Challenges of Analyzing Parametric CFD Results."

TIME, R. W., AND RABENJAFIMANANTSOA, A. H., (2012), "An Experimental Study of Gas Lift Efficiency in Polymer Systems." *Annual Transactions of the Nordic Rheology Society* 20:59–69.

VARNELL-SARJEANT, J. ET AL., (2015), "Comparing development approaches and reuse strategies: An empirical evaluation of developer views from the aerospace industry." *Information and Software Technology* 61:71–92.

WAITERS, R. W., AND HUYSE, L., (2002), "Uncertainty Analysis for Fluid Mechanics with Applications." National Aeronautics and Space Administration (NASA).

XIU, (2010), "Numerical Methods for Stochastic Computations: A Spectral Method Approach." Princeton University Press.

ZIGH, G., AND SOLIS, J., (2013), "Computational Fluid Dynamics Best Practice Guidelines for Dry Cask Applications." Office of Nuclear Regulatory Research.

ZORIC, J. ET AL., (2014), "On pragmatism in industrial modeling." *In* CFD 2014 - 10th International Conference on Computational Fluid Dynamics in the Oil & Gas, Metallurgical and Process Industries (S. T. Johansen and J. E. Olsen, eds.) pp. 1–16. SINTEF, Trondheim.

