Proceedings of the 1st International Symposium on AI and Fluid Mechanics
Paper No S6 P7

Reduced Order Models:
A Step Toward Aircraft In-Flight Icing Certification by Analysis

Wagdi G. Habashi
McGill University, Montreal, Canada H3A 2M7

Keywords: In-flight icing, computational fluid dynamics, certification by analysis, reduced order modeling, proper orthogonal decomposition, design of experiments, iterative sampling.

Abstract. While advanced CFD has established itself as the backbone of aerodynamic design (CFD-A), its application in the certification of in-flight icing protection systems lags considerably behind, as aircraft manufacturers follow a path of least resistance to obtain such certification. A Reduced Order Modeling (ROM) methodology is proposed to drastically reduce the time and cost of applying fully 3-D viscous turbulent CFD to icing simulations (CFD-I). ROM uses a limited number of CFD solutions (snapshots) to extract, via proper orthogonal decomposition (POD), modes that represent the fundamental physical features of the icing process. Multi-dimensional interpolation is then applied to a linear combination of these modes to yield, in mere seconds (in 3-D, for a complete aircraft) and in seconds (2-D, for aircraft sections), flow, water impingement, ice accretion, and performance penalties for any conditions within the design space other than the snapshots. Such drastic shortening of solution times paves the way for a complete exploration of the combined flight and aerodynamic envelopes, analyzing the aircraft as a system rather than an assemblage of components as it is often done.

1. Introduction

For a commercial aircraft to obtain a type design certification it must demonstrate safe flight into known or inadvertent icing conditions. The icing certification process involves the use of CFD, wind/icing tunnel testing (EFD: Experimental Fluid Dynamics), and flight testing with artificial ice shapes and in natural icing as a final demonstration of compliance (FFD: Flight Fluid Dynamics).

In applying CFD for certification purposes, some 2-D methods oversimplified as far as inviscid panel methods [1], are still in use in the aerospace industry or have been the basis of in-flight icing certification of a large class of aircraft still in operation today. This is mostly due to a chasm in technology appreciation between regulators and manufacturers, discouraging the latter from applying to icing analysis the same CFD used for aerodynamics lest it results in delaying the certification process. Thus, tried-and-tested legacy approaches that previously won regulators' approval become the norm, causing a stagnation of innovation for in-flight icing certification.

It stands to reason that *airfoils don't fly; only aircraft do* [2], causing 2-D approaches to be of limited use, if at all. High-performance computing allows CFD-I methods to creepingly make inroads into in-flight icing, analyzing an aircraft (fuselage, wings, engines, nacelles, empennage, auxiliary power unit, cockpit windows, sensors, probes, antennae, etc.) as a monolithic system and not an assemblage of isolated components. Using identical technologies for aerodynamics and icing permits, from the blank sheet phase, the assessment of performance improvements and icing penalties for each change and not as an afterthought on a final "frozen" aircraft configuration. However, the cost of the large parametric CFD simulations needed for certification is considered high by manufacturers, even though it pales in comparison to EFD and FFD costs. Simplifications are then accepted under a heuristic-ROM umbrella reducing dimensionality and/or simplifying physics. Dimensionality may be reduced from 3-D to 2-D and/or physics simplified by successively neglecting turbulence, viscosity, compressibility, and rotationality, ending up with the Euler equations, potential flow equation, and, in the lowest possible rung, panel methods. Such simplifications make little sense in the context of simulating icing effects to prevent massive flow separation.

The ROM approach presented here applies to both aerodynamics and icing with no sacrifice in dimensionality or physics [3]. The process starts by deriving continuity, momentum, and energy equations for icing (CFD-I) that are in the same class as the Navier-Stokes equations for aerodynamics (CFD-A), enabling a transparent exchange of information between the air, water droplets impingement, and freezing processes [4]. ROM then uses a limited number of CFD-A and CFD-I solutions (snapshots) well distributed within the design space, each with millions of nodal unknowns, to extract a limited number of modes [5] that capture the salient features of the system. These modes, computed by Proper Orthogonal Decomposition (POD) [6], give an optimal linear representation of the dominant features/physics that can be truncated to a specified accuracy level. Finally, a linear combination of the modes is used to obtain solutions throughout the design space for any other conditions than the snapshots, at a trivial cost.

POD solutions can, however, be significantly affected by problem nonlinearity caused by snapshots with markedly different physical characteristics. This could be the case of mixed subsonic and supersonic flows in CFD-A or greatly varying shapes of trace, glaze, and rime ice in CFD-I. A local POD approach is thus implemented to subdivide the solution space into clusters characterized by similar or sufficiently close physical characteristics. Here, k-means clustering [7] is used for cluster identification during the construction of the local POD, giving excellent results using a few modes.

Once the modes/basis vectors have been determined, the scalar coefficients in their linear combination can be obtained by defining a response surface for each coefficient through Kriging interpolation. The next step is to optimize the number and positions of snapshots using design of experiments (DoE) [8]. The accuracy of ROM solutions is estimated via leave-one-out cross-validation (LOOCV), with sampling techniques based on Centroidal Voronoi Tessellation (CVT) [9] to judiciously distribute the snapshots in the parameter space according to a prescribed density function based on the error or an a priori knowledge of the physics.

The following sections describe the ROM framework, with Section 2 outlining the machine learning concepts to ensure localization and an error-driven iterative sampling method for LOOCV and delineating online from offline calculations. Section 3 presents a convincing application combining airflow, impingement, and accretion over an aircraft for a complete sweep of the Appendix C Continuous Maximum certification rules (APP C-CM) [10] in mere seconds, while Section 4 concludes.

2. Methodology

2.1. Reduced order modeling applied to in-flight icing

For an ensemble of N_S observations "U_1 , ... , U_{N_S} ', where $U_i = U(\mathbf{x}_i) \in \mathbb{R}^{NP}$, and $\mathbf{x}_i \in \mathbb{R}^{ND}$ specifies input parameters defining observation conditions. N_P denotes the number of grid points from which the observation is defined.

For CFD-A, U can be the solution at all grid points, e.g., velocity, pressure, temperature or integral quantities such as shear stress or heat flux, while $\mathbf{x}$ defines flight conditions such as angle of attack (AoA), speed, altitude, etc.

For CFD-I, U can be the ice shape given by the coordinates of the N_P surface grid points or integral quantities such as the mass of ice, while $\mathbf{x}$ defines icing conditions in terms of, droplet Mean Volumetric Diameter (MVD), Liquid Water Content (LWC), etc.

POD then yields a set of basis vectors $\boldsymbol{\varphi}_j \in \mathbb{R}^{NP}$, $jj = 1, \dots, N_S$ that represents the dominant physical features within the snapshots, and the corresponding eigenvalues λ indicate the fraction of the total energy associated with each mode $\boldsymbol{\varphi}_j$ (i.e., $E_j = \lambda_j / \sum_{i=1}^{Ns} \lambda_i$), after which any solution for a case other than the snapshots is obtained via a linear combination of the modes.

$$U(\mathbf{x}^{\delta}) = \sum_{j=1}^{M \leq N_S} \alpha_j^{\delta} \boldsymbol{\varphi}_j \tag{2.1}$$

where M indicates the truncation of the expansion at the desired level of energy content. For each mode $\boldsymbol{\varphi}_j$, the projection coefficient at each snapshot location $\mathbf{x}_i$ is expressed as:

$$\alpha_j^i = U_i \cdot \boldsymbol{\varphi}_j \tag{2.2}$$

The $\alpha_j^i, i = 1, \dots N_S$ form a multi-dimensional response surface (hypersurface) for each mode $\boldsymbol{\varphi}_j$, shown in Figure 1, projected on APP C-CM, having as input the parameters $\mathbf{x}_i$ of the analysis and as output the α_j^i coefficients. Following this, the mode coefficient α_j^{δ} can be obtained by multi-dimensional interpolation for any untried input parameter $\mathbf{x}^{\delta}$, here by Kriging.

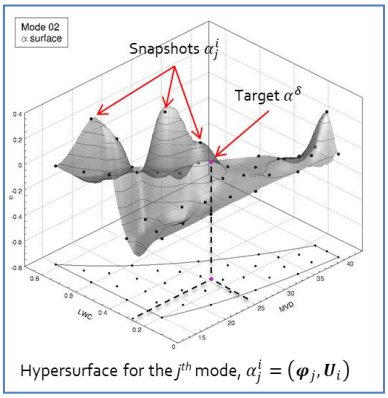


Figure 1: Response surface created by coefficients of the j[th] mode, projected onto APP C-CM.

2.2. Localization via machine learning, of trace, glaze, and rime ice shapes

Given that icing occurs at subsonic speeds, little nonlinearity exists within CFD-A calculations. Compliance with the icing certification envelope, however, can result in trace, glaze (horns), or rime (aerodynamic) ice, of significantly different characteristics. Such CFD-I nonlinearity requires grouping similar snapshots into clusters with clearly delimited boundaries and no void regions within. POD is then applied to each cluster to generate a local set of modes. The most suitable cluster for a new solution of an untried condition is represented as a linear combination of the local POD of that cluster.

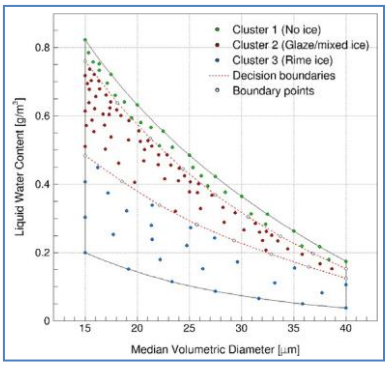


Figure 2: CM k-means clustering into 3 distinct regions with sharp decision boundaries.

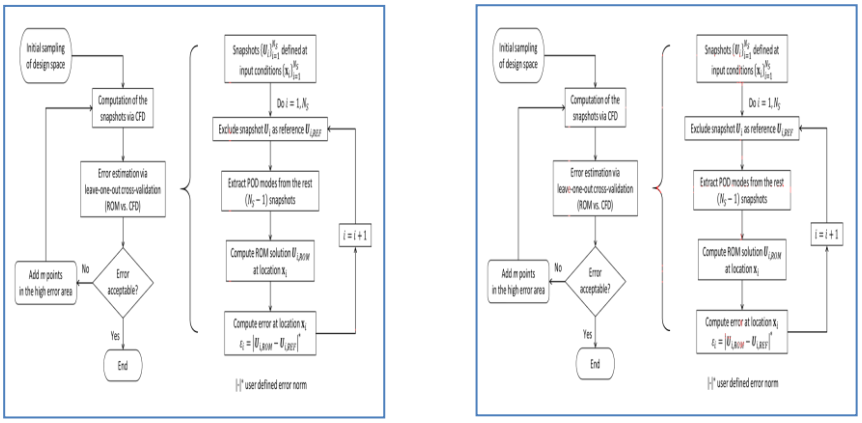
In k-means clustering, the snapshots "$U_1, \dots, U_{N_S}$' are taken as input, and the desired output is the class label $\mathcal{CC}_k \in \mathbb{Z}, k = 1, \dots, K$, which identifies a snapshot with a cluster k. The identification of the proper

cluster for an untried condition $\mathbf{xx}_\delta \in \mathbb{R}^{ND}$, and the definition of the boundaries of each cluster in the parameter space are carried out by logistic regression. To assign an input vector $\mathbf{xx}_i \in \mathbb{R}^{ND}$ to one among K available classes $\mathcal{CC}_k$, $k = 1, \dots, K$, a classifier is trained using a set $S = \{(\mathbf{xx}_1, t_1), \dots, (\mathbf{xx}_{N_S}, t_{N_S})\}$, where $t \in \mathbb{Z}$ is the label of the class. Since the classes are disjoint sets, the input space can always be divided into decision regions, with clearly defined decisional boundaries. Figure 2 shows the resulting subdivision of Appendix C's CM into the 3 distinct ice-type regions, with clear decisional boundaries separating them and each with its local reduced-order basis.

2.3. Snapshots' optimization via iterative sampling and leave-one-out cross-validation

To optimize the number of snapshots and their location in the parameter space an iterative sampling strategy based on a "greedy" application of CVT is applied, with the error estimate between the ROM and CFD snapshot based on LOOCV.

Flowchart 1 (left) demonstrates the error-driven iterative sampling process, while Flowchart 1 (right) details the leave-one-out cross-validation procedure. Given a set of N_s snapshots, a snapshot is excluded from the set and treated as the "reference" solution. A ROM is then defined based on $N_s - 1$ snapshots and compared with the reference. This is used N_s times, excluding at each time a different snapshot, giving an error indicator over the entire parameter space.

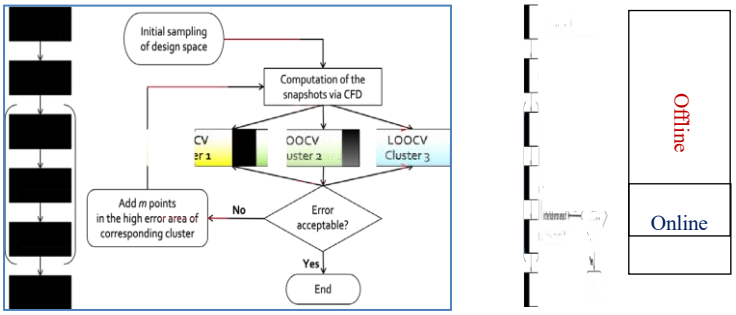


Flowchart 1: General flowchart of the snapshots' optimization approach.

2.4. ROM offline vs. online computational cost

The ROM computational cost is split into offline and online costs as shown in Flowchart 2. The snapshots, adaptive sampling, and the training of local ROMs are considered the offline cost (Flowchart 2, left), and Kriging is the online cost (bottom of Flowchart 2, right). When the LOOCV error is reduced to an acceptable level via iterative sampling, the set of snapshots is considered fixed. Then the extracted POD bases and their associated projection coefficients can be stored in a database for later use.

With Kriging as the interpolation method, the maximization of the likelihood function can also be done offline since the Kriging model weight parameter is fixed once the snapshots are fixed. The optimized model parameter can be stored in the same database, such that the time-consuming optimization process is avoided during the online stage.



Flowchart 2: Offline vs. Online cost.

To compute a target solution, the online cost consists only of the identification of the corresponding local reduced-order basis (logistic regression) and interpolation of coefficients for the modes (Kriging). Depending on the size of the problem, i.e., the number of modes M and dimension of mode N_P, the cost of the online phase for a 3-D problem is from a fraction of a second (real-time) to a few seconds, as opposed to the hours required by CFD-I.

3. ROM for a sweep of APP C-CM for an aircraft

ROM is applied to ice build-up on unprotected aircraft surfaces during a holding pattern of a Regional Jet (ARJ21) [12]. An unstructured mesh of 10M nodes, 12M tetrahedra, and 15M prisms is used. The airspeed and altitude are representative of a holding condition at 5,000 m at an Angle of Attack of 3.7°, and a total air speed of 268 knots. The exposure time considered for ice accretion is 25 minutes. The purpose is to demonstrate that ice shapes and mass within the CM icing envelope can be completely explored via ROM in mere seconds for the entire aircraft at once.

36 snapshots are initially used, each representing an APP C-CM condition in terms of MVD, LWC, and, implicitly, external temperature. One ought to be reminded that a snapshot includes a CFD-A solution, followed by water impingement and ice accretion CFD-I solutions, each using 10M nodes. LOOCV iterative sampling is applied resulting in 103 snapshots delineated by 3 distinct regions for trace, glaze, and rime ice. Figure 3 demonstrates the remarkable agreement between the ROM and CFD-I solutions at all locations of the aircraft (wing, engine, and empennage).

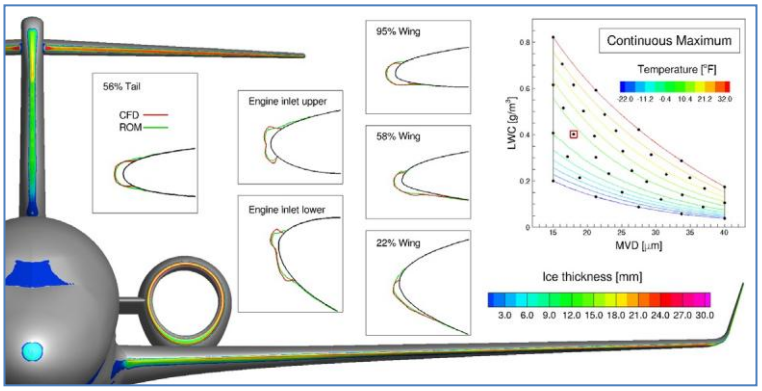


Figure 3: APP C-CM ARJ21 ice shape comparisons between ROM and CFD-Icing.

The bottom line of the approach is how much time is saved by using ROM: the computational time of a ROM solution is a mere 1.4 seconds using 6 cores (online cost), whereas a CFD-Icing solution requires 16 to 32 hours using 128 cores (offline cost).

4. Conclusions

The paper presents a ROM framework specific to in-flight icing problems. The methodology is based on response surface methods, proper orthogonal decomposition, multi-dimensional interpolation, and machine learning algorithms, along with an automatic error-driven iterative sampling approach to adaptively define and place an optimal number of snapshots. The introduction of local POD using machine learning algorithms for strongly nonlinear problems allows an equal exploration of all regions of the in-flight icing certification envelope.

The ROM methodology dramatically reduces the cost of 3-D CFD simulations. Other than the offline cost of obtaining the snapshots, all online 3-D complete aircraft results are obtained almost "gratis" and are even cheaper than using 2-D heuristic methods. This, by itself, should be a major incentive to move to realistic 3-D CFD-Icing analyses.

In addition, it should be remarked that the elapsed time for any number of snapshots is the time for one, as snapshots are independent of one another and can be run simultaneously in an embarrassingly parallel manner on a large enough cluster. Thus, 100 snapshots can be easily computed in a day.

Such inexpensive ROM provides a structured pathway to identifying the critical ice shapes for each phase of the flight, opens the door to systems optimization for which thousands of parametric results may have to be considered, and most importantly, can bridge the chasm between manufacturers and regulators if provided as a desk simulator allowing non-experts to carry out rapid verification over the entire certification appendices without the need for a CFD code: no meshing, no parameters adjustment, no large computers, and no need for local CFD expertise.

5. References

[1] https://software.nasa.gov/software/LEW-18573-1.

[2] Habashi, W.G., 2023. A Path to Enabling a Wider Use of Controlled-Accuracy 3D CFD in Industry and Academia, Special Issue "Crossing Computing Frontiers", International Journal of Unconventional Computation, 18, pp.15-48.

[3] Zhan, Z., Habashi, W.G., Fossati, M., 2015. Local Reduced-Order Modeling and Iterative Sampling for Parametric Analyses of Aero-Icing Problems. AIAA Journal, 53(8), pp. 2174-2185.

[4] Habashi, W.G., 2009. Recent Advances in CFD for In-Flight Icing Simulations. Journal of Japan Society of Fluid Mechanics, 28(2), pp. 99-118.

[5] Lucia, J.D., Beran, S.P., Silva, A.W., 2004. Reduced-Order Modeling: New Approaches for Computational Physics. Progress in Aerospace Sciences, 40(1), pp. 51-117.

[6] Sirovich, L., 1987. Turbulence and the Dynamics of Coherent Structures: I-Coherent Structures. Quarterly of Applied Mathematics, 45(3), pp. 561-571.

[7] Amsallem, D., Zahr, M.J., Farhat, C., 2012. Nonlinear Model Order Reduction Based on Local Reduced-Order Bases. International Journal for Numerical Methods in Engineering, 92(10), pp. 891-916.

[8] Statnikov, R.B., Matusov, J.B., 1995. Multicriteria Optimization and Engineering. Chapman and Hall, New York, pp. 225-261.

[9] Du, Q., Faber, V., Gunzburger, M., 1999. Centroidal Voronoi Tessellations: Applications and Algorithms. SIAM Review, 41(4), pp. 637-676.

[10] https://www.ecfr.gov/current/title-14/chapter-I/subchapter-C/part-25/appendix-

[11] http://english.comac.cc/products/rj/