Unstructured Grid Adaptation: Status, Potential Impacts, and Recommended Investments Toward CFD Vision 2030

Michael A. Park*
NASA Langley Research Center, Hampton, VA 23681, USA

Joshua A. Krakos† and Todd Michal‡
The Boeing Company, St. Louis, MO, USA

Adrien Loseille§
INRIA Paris-Saclay, Alan Turing Building, 91120 Palaiseau, France

Juan J. Alonso¶
Stanford University, Stanford, CA 94305, USA

Unstructured grid adaptation is a powerful tool to control discretization error for Computational Fluid Dynamics (CFD). It has enabled key increases in the accuracy, automation, and capacity of some fluid simulation applications. Slotnick et al. provides a number of case studies in the CFD Vision 2030 Study: A Path to Revolutionary Computational Aerosciences to illustrate the current state of CFD capability and capacity. The authors forecast the potential impact of emerging High Performance Computing (HPC) environments forecast in the year 2030 and identify that mesh generation and adaptivity continue to be significant bottlenecks in the CFD workflow. These bottlenecks may persist because very little government investment has been targeted in these areas. To motivate investment, the impacts of improved grid adaptation technologies are identified. The CFD Vision 2030 Study roadmap and anticipated capabilities in complementary disciplines are quoted to provide context for the progress made in grid adaptation in the past fifteen years, current status, and a forecast for the next fifteen years with recommended investments. These investments are specific to mesh adaptation and impact other aspects of the CFD process. Finally, a strategy is identified to diffuse grid adaptation technology into production CFD work flows.

I. Introduction

Advancements in both computers and algorithms over the preceding decades have resulted in steady Computational Fluid Dynamics (CFD) tool improvements, which have impacted analysis and design processes for aerospace vehicles. However, issues persist that require intense manual interaction and expert judgment to produce accurate and timely results. Slotnick et al.1 provides a number of case studies to illustrate the current state of CFD capability and capacity and the potential impact of emerging High Performance Computing (HPC) environments forecast in year 2030. They provide a number of findings and a comprehensive research strategy to enable the use of CFD outside of the small but important regions of the operating design

*Research Scientist, Computational AeroSciences Branch, AIAA Senior Member.
†Engineer, AIAA Senior Member.
‡Technical Fellow, AIAA Senior Member.
§Researcher, GAMMA3 Team, AIAA Member.
¶Professor, Department of Aeronautics and Astronautics, AIAA Associate Fellow.
space where CFD is currently employed. [Appendix A: Vision CFD 2030 Capabilities and Roadmap] quotes the key features and proposed timeline of this research strategy to address these findings. Past efforts at NASA identified roadblocks included in these findings that prevent inroads of large-scale simulation into the design process.

Three major roadblocks limit further inroads of CFD into the design process: (1) the lack of reliability of physical models (e.g., transition, turbulence, or gas kinetic models), (2) the long turnaround time of the numerical simulation, and (3) the lack of reliably automated functions (objectives and constraints) and derivatives for use in design optimization.

An approach was established to list and address these roadblocks.

The roadblocks, opportunities for breakthroughs, and long-term goals are presented in the selected areas of geometry modeling and grid generation, error assessment and grid adaptation, convergence acceleration, physical models and model synthesis, and design optimization. To realize the opportunities discussed, a much more adaptive approach, in terms of grids, solvers, and physical models, needs to be taken; the software development process is key to a successful implementation...

Progress has been made in these areas over the past decade and a half. However, a current stagnation in CFD development may limit the full potential of modeling and simulation for aerodynamic flows in coming decades. Developments toward the CFD Vision 2030 goals are expected to be complicated by an ongoing change in computer architecture summarized by Strohmaier et al. and Appendix A of Slotnick et al.

There are a number of findings in the CFD Vision 2030 Study that are interconnected and therefore require inherently multidisciplinary efforts to address these issues. One of the significant bottlenecks identified in the study is adaptive grid techniques and error estimation for complex flows and geometries. In particular, finding 3 of the executive summary states:

Mesh generation and adaptivity continue to be significant bottlenecks in the CFD workflow, and very little government investment has been targeted in these areas. As more capable HPC hardware enables higher resolution simulations, fast, reliable mesh generation and adaptivity will become more problematic. Additionally, adaptive mesh techniques offer great potential, but have not seen widespread use due to issues related to software complexity, inadequate error estimation capabilities, and complex geometries.

Unstructured grid adaptation status and the investments required to address this finding are presented here. These investments are motivated by identifying the impacts that robust automated unstructured grid adaptation would have on production CFD applications, modeling development, design, and uncertainty quantification. The current status of unstructured grid adaptation is identified with the deficiencies that impact routine application. Recommendations for investment are combined with a forecast for the next fifteen years to realize these capabilities and impacts. These forecasts are anchored by a prediction of the relative increase in computer performance and architecture changes. This forecast is also supported by a summary of progress made in the last fifteen years including a bibliography of published methods. This progress shows the steady introduction of novel methods and the enhancement rate of published techniques. The status, forecast, and recommended investments are supported by a detailed description of the elements of grid adaptation. Finally, recommendations are made to provide the critical step of impacting production workflows and projects by facilitating the adoption of newly developed and matured unstructured grid adaptation technology. The detailed unstructured grid adaptation descriptions and recommendations are made here with a clear intent of contributing to the CFD 2030 Vision Study capabilities.

I.A. Impacts of Mature Unstructured Grid Adaptation

Widespread use of improved unstructured grid adaptation techniques has the potential to make a considerable impact on the practical application of CFD-based analysis and design. Grid adaptation may allow the demonstration of asymptotic convergence rates sooner on smaller meshes. This includes high-order accurate schemes. It can also show robustness to initial grids. Observing design-order asymptotic spatial and temporal convergence rates through grid adaptation that are independent of the initial grid can dramatically improve the confidence in a simulation method and application. This goal has remained elusive on manually
created grids as shown by the Drag Prediction Workshop series. Unstructured grid adaptation is proposed as a method to demonstrate asymptotic convergence rates. Error estimation and control is a key technique for formal verification and validation of CFD. The combination of these factors could enable certification by analysis.

Dalle and Rogers detail the build up of an aerodynamic database for the complex issue of booster separation with thousands of automated simulations. They used an Euler output-adaptive Cartesian grid approach, which produced solutions with shock topologies that were not observed in the hundreds of viscous calculations used for uncertainty quantification. Dalle et al. indicated that “The inviscid assumption is tentatively acceptable due to the high supersonic Mach number (greater than 4.0) at separation and low dynamic pressure.” An adaptive unstructured grid viscous simulation tool with the same robustness and throughput as the adaptive Euler method would have great impact for building aerodynamic databases. Dalle et al. may show some evidence of nonunique solutions where the output-based technique is refining the solution related to the closest attractor or trying to extract a steady solution out of an inherently unsteady problem. These issues may be hard to discern without proper physics, discretization error controls, and strong nonlinear solvers. Database development is a situation where the capacity to do a large number of cases is as important as the capability of doing extremely large and complex single simulations.

Knowledge gained from grid adaptation has also impacted initial grid generation. Output-adapted grids and plots of the error estimate have been employed to modify the manually specified spacing field to resolve areas targeted by the output-based approach. This includes some regions of the flow that were intuitive (leading and trailing edge singularities) or not intuitive (stagnation streamline).

Slotnick et al. provide a number of case studies that show how controlling discretization error can positively impact developments in other CFD disciplines. For example, they summarize effects of grid resolution and solution scheme in assessing turbulence models.

A key gap in the effectiveness of current and future turbulence models is the effect of grid resolution and solution scheme on both the accuracy and convergence properties of the models. Studies show that adequate grid resolution is required to capture the full range of turbulence structures in models ranging from simple eddy-viscosity formulations to full LES [Large Eddy Simulation] and DNS [Direct Numerical Simulation] simulations. Park et al. summarized the impact of grid resolution on the side-of-body separation bubble seen in simulations of the Common Research Model. Without adequate grid refinement, the side-of-body modeling issue is masked by discretization error, i.e., coarse wing body junction grids tended to suppress the prediction of an enlarged separation bubble. This enlarged bubble is not observed in wind tunnel tests. Having an adaptive grid method that can eliminate discretization error as a concern may have accelerated the adoption of improved modeling techniques such as the Quadratic Constitute Relation (QCR) that dramatically improved results.

Turbulence models may perform poorly when under resolved at the wall or in outer regions of the boundary layer. This is often the result of using a geometric growth rate or an accelerating geometric growth rate for the boundary layer cell heights. Some boundary layer transition prediction methods require well resolved boundary layer profiles including accurate higher order derivatives and inflection points. The properties of the boundary layer are not known before the solution is computed. This makes grid adaptation uniquely suited to ensuring that the first cell height and the outer portions of the boundary layer are properly resolved. Increased resolution of the outer portions of the boundary layer can result in a dramatic increase in turbulent eddy viscosity.

The limited use of existing error estimation and control methods and the inadequacy of current error estimation techniques are also identified by the CFD Vision 2030 Study. Larsson and Wang discuss the feasibility of using turbulent eddy resolving methods in an industrial design process. The need for grid adaptation is specifically mentioned. It is illustrated by showing how manually-specified grid refinement enhanced predictions elsewhere in the simulation. “It is difficult even for experienced users to anticipate exactly what grid resolution is required in different regions, and thus a more mathematical approach to this problem would quite clearly be very useful.” Using a rigorous mathematical approach to automate this discovery would have a dramatic impact on complex design and analysis applications. Full aircraft airframe noise prediction and control is another example that requires eddy resolving methods and would impacted by solution adaptive technology. The analysis and design of open rotor systems is another critical noise application that would be impacted by improved grid adaptation methods.
Robustness and automation of CFD analyses enables optimization and multidisciplinary analysis environments. Poor mesh generation performance, low mesh generation robustness with a high degree of manual intervention, and inadequate linkage with Computer-Aided Design (CAD) are all listed as impediments to autonomous and reliable CFD simulation that improved grid adaptation techniques can address. Unstructured mesh adaptation yields a systematic procedure to generate meshes where the discretization error is controlled. Coupling mesh adaptation with shape optimization leads to better functional and gradient evaluation by ensuring that the solutions have equivalent accuracy during the design cycles. Response surfaces are often constructed for uncertainty quantification and design. Mesh adaptation can provide a fast way to generate coarse grid solution samples or very accurate, fine grid solutions. This insures that the space of the uncertainty and design parameters are rapidly populated with consistent accuracy or multi-fidelity methods.

The integration with the other processes required for multidisciplinary analysis is improved when discretization error is controlled and may present the need for also controlling errors that are introduced by the coupling terms. Design optimization is often performed with a parameterization of the discrete grid, which requires the optimized shape to be reverse engineered into the original CAD model for analysis by other disciplines or for manufacture. Direct optimization with parameterized CAD models would alleviate this difficulty, but requires the sensitivities of these parameters to be propagated through the CAD model, adapted grid, and multidisciplinary analysis.

The impact of CFD tool development on NASA science and space exploration missions is detailed by Slotnick et al.\textsuperscript{1} The commercial and military space programs are equally impacted by improved CFD techniques. These opportunities have been pursued for the last decade and a half for unstructured grids\textsuperscript{2} and include the maturation of advanced entry, descent, and landing (EDL) concepts such as deployable (inflatable) heat shields and supersonic retropropulsion (SRP). Adaptive grid methods are needed that improve the critical technology of heating prediction.\textsuperscript{32} Adaptive unstructured grids are uniquely suited to resolving the rapidly varying characteristic directions of SRP and the interaction of shocks and boundary layers if the accuracy of computing gradients on the boundary (heating and skin friction) is maintained. The elevated levels of uncertainties in the prediction of heating often result in the overly conservative design of heat protection systems by multiple factors. This increases the mass of the system, which dramatically reduces mission performance. Uncertainties in predicting the transonic launch environment\textsuperscript{33} of a launch vehicle can result in an overly conservative design or a design that fails in flight.

## II. Current Status and Limitations of Unstructured Grid Adaptation

After two decades of research and development, anisotropic grid adaptation has yet to be adopted in mainstream commercial or industrial CFD use. There are examples where grid adaptivity has found its way into production usage, but these are primarily for simplified geometries or for applications where isotropic adaptation can be utilized.\textsuperscript{34,35} There are several reasons for this slow adoption, but some of the main impediments are stringent geometry requirements, poor robustness (particularly for high Reynolds number turbulent flows on complex geometry), efficiency relative to fixed grid approaches, limitations of error estimation methods, and inadequate demonstration on verification and validation databases. The past decade has seen a gradual maturation in several critical technologies that are necessary for a robust and efficient adaptive capability.

Alauzet and Loseille\textsuperscript{36} surveyed the status of anisotropic mesh adaptation, current limitations, and progress made in the last decade. More details of the elements of unstructured grid adaptation is provided in section \textbf{IV.D} and \textbf{Appendix B: Partial Bibliography of Unstructured Grid Methods}. Robust grid adaptation mechanics that produce and modify anisotropic elements with aspect ratios on the order of tens of thousands are required for high Reynolds number viscous flows. Many variants of local operator adaptive mechanics are currently employed.\textsuperscript{37–44} This has been shown to be unified by the cavity operator,\textsuperscript{45} which has yet to be widely implemented but shows promise for producing semi-structured regions in the boundary and interior of the domain.\textsuperscript{46,47} The differences between a mapped isotropic method and a metric-orthogonal method are shown in Fig. 1. Grids created with standard anisotropic procedures, Fig. 2(b), can result in large gradient errors if they are slightly misaligned with the gradient direction.\textsuperscript{48} An example of the metric-orthogonal method applied to the Trap Wing configuration is shown in Fig. 2 where the wake region is constructed of elements that have an orthogonal construction in both views.
For many flow solvers, incorporating semi-structured stacks of prismatic elements near wall bounded shear layers improves the accuracy of computed normal derivatives. An example of grid adaptation with prismatic element stacks is shown in Fig. 3. Semi-structured stacks of elements also provide benefits by constructing hybrid schemes or line relaxation, which can increase accuracy and residual convergence rates. The semi-structured nature of prismatic stacks has also been leveraged to enable $y^+$ adaptation, which can augment error estimation procedures that degrade near the boundary. A common approach for mixed element grid adaptation is to maintain the prism element stacks found in the initial grid as the grid adaptation progresses. This hybrid prismatic approach appears to have benefits for the near term until the cavity operator approach to semi structured element insertion fully matures.
Adaptation on complex industrial CAD geometry models presents difficulties beyond those of standard grid generation methods. There have been a number of published CAD coupling techniques, but no standard or preferred method has emerged for coupling or interchange of geometry. Employing a different geometry source than CAD for design and CFD application is common place. Translation between these geometry sources introduce translation errors, require manual steps, and is influenced by human judgment. This can be especially problematic for grid adaptation where localized curvature or tangency errors may become an attractor for error estimates resulting in pockets of excessive refinement. Requirements for the geometry fit tolerance (maximum allowable gap/overlap between neighboring surfaces) are also much more stringent and difficult to predict for adapted grids as the required tolerance is directly related to the final grid resolution, which is not known a priori. Problems with the geometry may not be discovered until the adapted mesh is sufficiently resolved to highlight the problem area, see Fig. 4 where excessive refinement is caused by misalignment of geometry surfaces not resolved on a coarse mesh. As a result, preparation of geometry models suitable for adaptation can be time consuming and prone to labor intensive rework. Another weakness of using a typical CAD system for grid adaptation is that they may employ sequential execution on different hardware than the parallel CFD and grid adaptation processes. Supporting a different computer architecture and communicating with it from an HPC system can complicate direct geometry coupling. Surrogate geometry models can address many of the geometry modeling issues, but a standard approach for construction, evaluation, and interchange of surrogate geometry models has yet to be adopted.

Adaptive meshing poses unique challenges to efficient execution on parallel computing systems. As the adapted grid size is continually evolving during the application, careful attention must be given to load balancing. Also, adaptation of the surface grid requires distributed access to the geometry model, a capability that is rarely incorporated in existing adaptation tools. Several published grid adaptation methods incorporate parallel execution on small- to moderately-sized systems of thousands of cores. Extension to massively parallel environments envisioned in the 2030 report will require additional improvements in load balancing techniques and likely will require multi-level parallelization to attain the full performance available on hybrid architectures.

For smooth areas of the flow field, higher-order CFD methods and the use of h-p adaptivity has been shown to have advantages over second-order methods. While adaptive meshing for higher-order solvers is less mature than linear meshes suitable for second-order methods, several researchers have published work...
in this area, which is reviewed by Gargallopeiro. A curved mesh example is shown in Fig. 5 where mid-nodes are projected to the underlying CAD model producing quadratic triangles. These mid-nodes can also be projected to a reconstructed surface. Adapting for higher-order methods introduces the additional complexity of generating curved element grids. A standard approach for curving higher-order grid elements is to introduce nodes at the collocation points of the linear elements and then use a deformation method to diffuse the displacements of the nodes on curved geometry into the volume of the domain. For complex geometry models, the deformation approaches are often insufficient to avoid crossing of the elements. In these situations, untangling methods are often employed to restore the mesh to a usable format. Local mesh modifications can also be used to recover curved meshes and directly adapt moving curved meshes.

Output error estimation is most commonly based on adjoint-weighted residual formulations. Hessian reconstruction is commonly used to determine anisotropy for second-order (and sometimes higher-order) discretizations but reconstruction is problematic near boundaries and in other situations where gradient reconstruction is degraded. A technique that optimizes local error-metric pairs is suitable for higher-order discretization and time-accurate formulations. This optimization technique avoids the Hessian but requires an implied metric of the current grid that can contain noise due to irregular grids. The continuous mesh framework has a rigorous mathematical formulation for Euler and Laminar steady and time-accurate problems. It also depends on a reconstructed Hessian but is free from the implied metric. This continuous formulation establishes a direct connection between the metric field visualized as ellipses in Fig. 6(a) and the mesh with unit edge lengths in the metric space shown in Fig. 6(b). Edge sampling or searches for the direction of the next higher-order solution derivatives can be used to form the metric but these methods have not been evaluated on turbulent boundary layer applications. Feature-based error estimation is used for problems where the adjoint is not available or where refinement of a particular feature is desired. Feature-based error estimation appears to improve the resolution of these features, but there is no guarantee that the features are not displaced by discretization errors elsewhere in the domain or that resolving these features improve the accuracy of integrated forces and moments.

Much progress has been made recently in increasing the robustness of the flow solver but this adaptive solver technology has not been implemented uniformly in all CFD solvers. Incipient or large scale unsteadiness may require the use of time-accurate simulation or continuation to obtain the desired solution when multiple solutions or hysteresis are observed. For example, Nishikawa et al. demonstrate this behavior for the Drag Prediction Workshop configuration in Fig. 7. A standard linearized adjoint is unstable for eddy resolved turbulent calculations. This currently requires a stabilized adjoint or heuristic approach for an error estimate.

Grid adaptation research has been focused on steady analysis. Some pioneering efforts have been applied to time-accurate simulations, but all of the issues present in steady grid adaptation are compounded. These include execution time, robustness, parallel execution, and error estimation. Time-accurate adaptation adds the requirement of interpolation to maintain accuracy, particularly mass conservation. While the grid can
be adapted at each time step or a single grid can be adapted and held fixed for all time steps, using a number of interpolated sub-intervals may be the optimal solution.86

III. Forecast and Recommended Investments

The success of unstructured grid adaptation can be measured in many ways. Beyond the capabilities demonstrated in publications, impacting production workflows is ultimately the measure of success for any candidate technology. Johnson, Tinoco, and Yu29 illustrate how CFD developments are adopted by application engineers at Boeing Commercial Airplanes. This adoption or technology diffusion process has been studied in many fields and Appendix C: Strategy for Diffusion of Unstructured Grid Adaptation Technology details this diffusion process and summarizes strategies for facilitating the adoption of unstructured grid adaptation into production workflows. Some of these recommendations to accelerate diffusion backed by research are more intuitive than others. Facilitating adoption and technology diffusion is a complementary effort to the technological developments described in this section.

Each forecast time period that follows has a prediction for relative improvement in computer performance to provide context for the described applications. These forecasts are based on estimates by Strohmaier et al.5 and Appendix A of Slotnick et al.,1 where a given simulation is predicted to execute approximately ten times faster every five years on the new hardware as it becomes available. Though this is a lower rate than historical observations before 2010, this would still result in an increase in speed over the next fifteen years by factor of 1,000. The National Strategic Computing Initiative (NSCI)87 may help to sustain or accelerate these performance trends. These gains are expected to follow past treads of increased concurrency where “increasing number of cores per socket has compensated for stagnant core performance in the latter half of the past decade.”3 While the majority of the data that these forecasts are based on is theoretical peak and Linpack benchmarks, the increases demonstrated by Gorden Bell Prize winners are well correlated, indicating that these relative performance forecasts may be applicable to actual application performance.3 These performance forecasts help to establish the increase in capability and capacity that may be realized in the next fifteen years, which is decomposed into five year intervals.

III.A. Five Years: 2020

A CFD simulation in 2020 is predicted to execute ten times faster than it did in 2015. Cooperation between CFD solver technology development and adaptive grid methods is critical during this five year period. This cooperation should include participation in the AIAA Special Sessions on the Evaluation of RANS Solvers on Benchmark Aerodynamic Flows88,89 organized to verify and characterize turbulent solvers. There are a number of promising techniques90–92 that increase the automation, reduce execution time, and obtain lower iterative error than typical iterative solver and solution update schemes. The diffusion of these techniques and hardening of improved automated solvers will positively impact the unstructured grid adaptation process where many error estimates assume the discrete residual is zero. Other disciplines (e.g., time-accurate
simulations, multidisciplinary analysis, design optimization, aerodynamic database creation, high-energy reacting gas simulations, or uncertainty quantification) will also be beneficially impacted by improvements to solver technology.

The management and interaction of the geometry definition with the adaptive process will require significant investment. Mesh and geometry databases (e.g., MOAB or PUMI) should be evaluated to leverage the effort placed on the efficiency and reduced memory footprint of these packages. Some of these databases also include linkages to CAD and CAD surrogate geometry. This evaluation also has the potential to provide the mesh database developers with feedback on the typical stencils of anisotropic grid adaptation kernels and the relative performance of different implementations. The underlying geometry is repeatedly accessed during the boundary mesh modification phase of adaptation. As problem size and concurrency increase, parallel access to underlying geometry will become a requirement of the adaptive process. Some of the parallelization tasks may be implicitly handled by using a parallel mesh database. Continued collaborative development with these databases’ developers may help to ensure that unstructured grid adaptation is efficient on current and future architectures.

During this time frame, the tighter geometric tolerances required by adaptation will conflict with existing geometry handling best practices for fixed grid approaches. Without watertight representations of the geometry sources, adaptive methods may fail as they resolve important flow physics near gaps or inconsistencies in neighboring boundary faces that a fixed grid may not resolve. The need for robust boundary grid adaptation will drive robust connections between geometry sources and the adaptive methods that use them. In addition, mesh surrogate models will be matured to maintain watertight and smooth geometry where a higher-fidelity geometry representation is unavailable or does not meet the required tolerances. Implicit geometry representations may be used to address localized regions with inadequate CAD tolerances or may be combined with explicit boundary representations to attain robustness without a reduction in geometric accuracy, i.e., maintaining sharp corners.

Research during this five year period will address the weaknesses of the currently available metric-based error estimates while unifying their strengths. A number of current error estimates rely on an implied metric of the current grid as a baseline metric field. This makes these estimates sensitive to artifacts present in the current grid (e.g., poorly shaped elements or nonsmooth gradation). Additionally, reliance on a reconstructed Hessian can be problematic where boundaries, grid quality, and solution error can degrade the reconstruction. The ideal error estimate should include CFD solver element quality requirements, which can vary from solver to solver. Examples include element orthogonality, dihedral angles, and neighbor element volume ratios. As higher-order methods begin to displace traditional second-order solvers, error estimates that extend naturally to arbitrary order will be in demand. The increased use of time-accurate simulations will require extension to time-varying grids and chaotic systems. While some applications will need to resolve a particular output at a given set of conditions, many will require the adaptive process to account for multiple output functions at multiple conditions, and even global solution error throughout the domain.

The robustness of adaptive grid processes will improve in this time frame to produce more accurate solutions. Improved metric-based error estimates will provide the basis for grids that demonstrate a reduction in error. Local cavity operators and marching techniques will allow for semi-structured regions in near-wall boundary layers, shock waves, shear layers, and wakes as in Fig. 2(c). The improved orthogonality of the elements can improve the accuracy of gradient quantities, especially adjacent to the boundaries (e.g., skin friction or heating).

A balance will continue to be made between 2D test cases that execute quickly to permit rapid feedback and 3D test cases that are inherently more difficult. Demonstrating progress on complex 3D geometries is critical to the relevancy and adoption of the methods. This demonstration will include verification and validation databases to increase confidence in available methods and gain a foothold in production CFD communities. A similar balance will be made between sequential and parallel implementations.

At the end of this five year period, adaptation will likely remain a niche capability used for particular problems or when discretization error is identified as a dominant error source. In the mean time, it will enable the accurate solution of problems where grid convergence studies with fixed grids would be prohibitively expensive. Verifying large fixed grid databases with a sample of adaptive grids and using the adaptive results to guide more traditional grid generation processes will become increasingly common. Applications where existing fixed grid approaches do not produce suitably accurate results to support engineering decisions will begin adoption of adaptation in earnest. Multiple commercial vendors will offer feature-based grid adaptation
technology with a small subset providing more rigorous error estimation techniques to their customers.

III.B. Ten Years: 2025

A CFD simulation in 2025 is predicted to execute one hundred times faster than it did in 2015. The increase in concurrency will require all practical unstructured grid adaptation schemes to execute in parallel and may create situations where fault tolerance and resiliency become critical elements of an adaptive simulation. Time accurate simulation will become more routine and require reliable error estimates for unsteady chaotic flows to drive grid adaptation.

Reliable metric-based error estimation extensions will include other disciplines like structural analysis and acoustic propagation. These include the exploration of synergies between continued development of boundary layer modeling methods (e.g., wall functions, cut cells, immersed boundary, or integral boundary) with error estimates. The uncertainty quantification issues raised by Barth will be addressed and unstructured grid adaptation will be extended to include the control of modeling errors. Methods such as Palacios et al. and Yano will mature to produce grids with low discretization error over a range of inputs with potential benefits for uncertainty quantification and design. This is a dramatic improvement over the current practice of uncertainty quantification and design optimization based on solutions with unknown and highly variable discretization error. These developments will combine to augment multidisciplinary design optimization with quantification and control of discretization error.

At the end of this five year period, automated grid adaptation will be ready for simulations that have limited application in the past. Adaptation will be used for steady and eddy resolving simulations, high lift calculations, and bluff body flows. Design optimization based on adapted grid solutions will be achieved at comparable or superior efficiency to fixed grids, with higher confidence in the accuracy of the results. As this is accomplished, other error contributions such as modeling error can be attacked in a manner isolated from discretization errors in space or time.

Adaptation will enable heating for reentry flows to be predicted to a lower uncertainty, which will reduce the required conservatism and mass penalty of thermal protection systems. CAD exchange formats may mature by this point to facilitate interchange of parameterized geometry for design, which is also compatible with additive manufacturing. This will continue to break down the walls of analysis and design silos where geometry representations are imported into a particular system and difficult to retrieve. During this time frame, the adaptive process, including error estimation and metric generation, will be validated for wider classes of problems enabling greater adoption. For example, recovery of asymptotic convergence rates on a grid converged solution for the Drag Prediction Workshop Common Research Model configuration will be demonstrated. These demonstrations will encourage customers to request grid adaptation technology from commercial vendors, which will make this technology more common in commercial offerings.

As adaptive simulations grow in size, frequency, and complexity, completely testing all aspects of the integrated CFD process during development will no longer be possible. While each component must be made robust, hardening all components against all the conditions encountered in production is not possible. Therefore, robustness must also be incorporated into higher levels of the system to allow recovery from component failure. This situation has been identified by large companies that employ hundreds of thousands of servers in multiple data centers to stream video and search databases at Internet scale. “[W]hen you set out to engineer a system at Internet scale, the best you can hope for is to build a reliable software platform on top of components that are completely unreliable. That puts you in an environment where complex failures are both inevitable and unpredictable.” This realization has shifted the emphasis from pre-deployment testing to monitoring the application in production. Some organizations intentionally cause these subsystem failures in a controlled manner to measure and increase resiliency of the entire system. “Our plan is to preemptively trigger the failure, observe it, fix it, and then repeat until that issue ceases to be one.” The initial production applications of grid adaptive methods will provide environments that can be statistically probed for weakness and hardened (e.g., failure to evaluate a CAD geometry query, rebooting a server, network failures, or flow solver divergence).

III.C. Fifteen Years: 2030

A CFD simulation in 2030 is predicted to execute one thousand times faster than it did in 2015. Computer architecture is expected to become more heterogeneous with elements that can throttle up and down based
on load or resources (e.g., electrical power). Fault tolerance and resiliency, as well as adaptive load balancing that can account for time-varying processor performance, will become critical to efficient resource utilization.

In this time frame, adaptive grid computations will have displaced fixed grids as the default, to the degree that the practitioner will generally never visualize the grid directly. The maturation of error estimation techniques and grid mechanics will effectively make the grid invisible to the CFD analysis process. This confidence is supported by demonstration of production CFD relevant test problems. Standard practice will include adaptive discretizations, adaptive solvers, and adaptive load balancing. Rather than depending on specialized knowledge of the practitioner to generate a grid, the focus will shift to the design of input geometry and uncertainty quantification of multidisciplinary analysis. Ultimately, success of grid adaptation will be measured by removing spatial and temporal discretization as a concern. As this is accomplished, other error contributions such as modeling error will become dominant. Once other sources of error are addressed, the continued advancements to unstructured grid adaptation and experience with verification and validation databases will set the stage for certifiable analysis and certification of systems by analysis.

Interactions with other disciplines will form strong collaborations whereby estimation of discretization, coupling, modeling, and manufacturing errors will be quantified, controlled, and balanced to increase the robustness of aerospace vehicles. This will allow the reallocation of human interaction to focus on creative solutions and accelerate the market entry of new concepts that make improvements through multidisciplinary interaction. The automation that results from removing the grid as a concern for the CFD application engineer will present a clear competitive advantage to the majority of commercial vendors that offer grid adaptation technology.

IV. Elements of Grid Adaptation

In this section, the components that compose a typical adaptive process are reviewed. Each component has its own issues and is generally considered a field of research on its own. For example, many of the developments of these components are listed in Appendix B: Partial Bibliography of Unstructured Grid Methods. The optimality of the whole adaptive process in term of convergence and solution accuracy can be only achieved once the complex interaction of these components is understood completely. This decomposition is illustrated in Fig. 8.

Starting with an initial grid, a flow solution is computed. Optionally, an adjoint solution can be computed based on the flow solution and an output of interest to identify impacts of local error on that output. The information from the flow (and possibly adjoint) solution are used to estimate interpolation error, truncation error, discretization error, or identify features of the flow. This information is used to specify a new grid resolution request. If the estimated errors are larger than limits specified by the practitioner, the current grid system is modified by grid mechanics to adhere to the grid resolution request. Once the adapted grid is available, the previous flow solution is interpolated to the new grid and the process is repeated until an exit criteria is met (e.g., accuracy requirement or resource limit). There are potential interactions between each of these elements that impact the overall convergence and efficiency of the adaptation process. The computational environment is also crucial to system performance. For instance, the execution and communication between these steps must be efficient on the target architecture to enable a high throughput of analysis.
This review is focused on the computation of steady flow phenomena of a single discipline analysis. For unsteady simulations, the flowchart in Fig. 5 must be modified with an additional inner loop to integrate the solution in time. Each of the grid adaptation elements becomes more critical for unsteady simulations. The transfer of the solution on a new grid has to be conservative to avoid the loss of mass.\textsuperscript{103,104} The regularity of the adapted grids has a major impact on the time-stepping of the flow solver, particularly for explicit time advancement methods. If an adjoint-based error estimate is used, a stable backward adjoint solver is needed. Multidisciplinary analysis and design optimization further complicate the flowchart by adding inner and outer loops.

IV.A. Grid Mechanics

This review focuses on unstructured grid adaptation, primarily with simplicies and pyramidal, prismatic and hexahedral elements. Hybrid elements are generally restricted to the boundary layer region and (adapted) tetrahedral elements to the outer field. This narrow view is based on several observations. First, we have the opinion that high aspect ratio elements with general alignment and gradation\textsuperscript{77} are required to efficiently resolve Euler\textsuperscript{105} and high Reynolds number fluid flows.\textsuperscript{106} There exist fully automatic tetrahedral-based mesh generation methods with a high-level of maturity and robustness for very complex geometries. Simplicial elements have the necessary flexibility required for adaptivity in size and orientation that is lacking in structured grids or hybrid elements.

For the sake of completeness, we first review the Cartesian-based adaptive methods based on structured and nonconformal grids. Then, we discuss the standard strategies to generate adapted meshes. We also describe the standard mesh modification operators. Finally, three crucial issues especially for viscous flows namely the CAD interaction, parallel computing, and boundary layer mesh generation are discussed.

Cartesian-based adaptation. There is a class of hierarchical subdivision techniques that have been extremely successful for isotropic resolution,\textsuperscript{35,107} especially when coupled to error estimation techniques that target functional errors.\textsuperscript{34} Even anisotropic refinement in Cartesian directions typically requires coupling to a different technique for anisotropic resolution of boundary layer flows (e.g., hybrid grids),\textsuperscript{108} near-body mixed-element unstructured solver,\textsuperscript{35} or near wall model\textsuperscript{109,110}. The subdivision of block structured grids by Hartman\textsuperscript{111} and Ceze and Fidkowski\textsuperscript{112} provides arbitrary alignment, but this alignment must be specified during initial grid generation because it is restricted to the local curvilinear grid orientation. Ceze and Fidkowski\textsuperscript{113} compare this hanging node subdivision of structured grids to unstructured grid adaptation. Both of these methods produced comparable asymptotic results in 2D, but the initial unstructured grids were easy-to-generate with isotropic elements and the initial structured grids were more meticulously tailored. High quality structured grids are more difficult to create in 3D.\textsuperscript{112}

Unstructured-based adaptation. Grid adaptation can be accomplished by simple regeneration of the entire grid to adhere to the interpolated spacing field from a background grid.\textsuperscript{114–116} This regeneration can also be limited to specific regions of the grid.\textsuperscript{117} Unfortunately, this can be an inefficient method for small, localized changes to the grid resolution request and may require the same level of manual intervention that initial grid generation often requires. In addition, global adaptive procedures become less robust when complex geometries and a high-level of anisotropy are required. For example, boundary recovery becomes increasingly difficult for constrained-Delaunay methods and closure of the front becomes increasingly difficult for advancing front methods. These difficulties have limited the use of global methods to the generation of isotropic meshes or slightly anisotropic meshes. Consequently, local remeshing approaches have become the method of choice in order to guarantee that a mesh is always generated (where failure of the mechanics results in a suboptimal but valid grid). Most of the local strategies are based on a combination of simple mesh modification operators, i.e., node movement, splitting, collapse, and edge/face swaps.

Iterative node movement is suitable for structured and unstructured grids as the topology of the grid is kept unchanged. This method iteratively moves nodes to cluster and align elements for satisfaction of the resolution request or to improve element quality.\textsuperscript{118,119} The node movement can also be driven by modeling physical processes.\textsuperscript{120,121} There are many possible splitting stencils,\textsuperscript{122–124} but often the simple approach of inserting a single node or splitting a single edge is preferred. Edge collapse merges two nodes connected by an edge by deleting one node, deleting the elements incident to that edge, and updating the connectivity of the elements to include the surviving node. Freitag and Ollivier-Gooch\textsuperscript{118} describe the canonical configurations of tetrahedral face and edge swapping and how to reduce the cost of evaluating
potential swapped configurations. The elemental operations of split, collapse, swap, and node movement can be combined to iteratively modify a grid. When an anisotropic mesh is sought, the description of the desired size and orientation of the mesh is prescribed via a metric field. All the previous operators are then recast in a metric-based framework to satisfy the input metric. The scope of the remeshing algorithm is to produce a uniform mesh in the metric, i.e., where the lengths of edges are close to one and the volume of the elements are close to the volume of the unit regular tetrahedron.

This approach examines each edge in the mesh and splits or collapses the edge if it is too long or too short as compared with the mapping metric, see Fig. 6. Michal and Krákos produce metric-conforming grids using only edge split and collapse operations. Loseille and Menier have shown that a cavity based operator unifies these insert, collapse, swap, and node movement operators into a single framework. The flexibility of the cavity operator has enabled metric-orthogonal grid adaptation, which produces grids with locally structured regions and improved dihedral angles. This metric-orthogonal approach shows the potential to produce fully adaptive boundary layer grids and semi-structured regions to resolve shear layers and shocks in the volume.

CAD interaction. The interrogation of geometry is challenging. Geometry representations are typically maintained in a computer-aided design (CAD) system that may not be intended for parallel execution on a HPC machine. The presence of curved geometry imposes constraints on the grid mechanics and maintaining high quality elements near these constraints is difficult. This is compounded by geometry fit tolerances that can be larger than the final grid resolution, which is not known before grid adaptation is attempted. Implicit geometry representations can address some geometries with inadequate tolerances, but have not found widespread use in adaptive procedures. Numerical evidence indicates that a high-order geometric approximation of curved boundaries is mandatory if accurate numerical solutions are sought. These curved grids are often created by displacing linear generated or adapted grids or local mesh operators. These high-order grids are untangled or optimized to improve a distortion measure. Iterative methods have been developed to project the surface node to the “first problem plane” past the first element that becomes inverted during projection) with various combinations of edge swapping and node movement. However, local grid regeneration is reported to be necessary when this iterative method fails. Loseille and Löhner illustrate how the cavity based operator can readily be extended to include boundary recovery during grid adaptation. This operator may help to address a major issue for high-Reynolds viscous flows where projecting or inserting a new point on the geometry can be prevented by the presence of the boundary layer mesh or require the replacement of a large number of boundary layer elements.

Boundary layer meshing. The previously described meshing strategies to generate anisotropic meshes have shown a limited success when applied to complex turbulent flows and complex geometries. This fact is explained by the difficulty to provide accurate error estimates within the boundary layer and the difficulty to generate highly anisotropic grids with the properties preferred by standard numerical schemes (e.g., right angle elements). The combination of these difficulties (e.g., high aspect ratio adaptation, curved boundary, and robustness of flow solvers on general unstructured grids) has resulted in the segregation of a stack of prismatic elements from an unstructured tetrahedral grid. The heights of the elements in these stacks can be adjusted to perform y+ adaptation. The stacks can be modified in the tangential direction with subdivision or anisotropic adaptation.

Parallel mesh adaptation. The same scalability and parallel execution issues arise as in analysis, but are compounded by the nonuniform memory access of adaptive grid algorithms. Heuristics can be used to anticipate the new size of the partitions as the grid is modified, but the change in partition size introduces more complexity than the typically static stencil and communication pattern of the CFD solver. Chrisochoides et al. proposes that a hierarchical parallel execution scheme will be required on the architectures anticipated by the CFD Vision 2030 Study. This hierarchy includes domain-decomposition approaches as well as multithreaded and shared-memory approaches. Domain decomposition requires load balancing to maintain efficiency for the nonuniform operations of grid adaptation. Devine et al. provides a review of load balancing and discussion of mesh data structures. Kavouklis and Kallinderis focus on load balancing issues related to parallel unstructured grid adaptation methods. A number of domain-decomposed approaches adapt the grid in the interior of partitions and then locally migrate a specific set of cavities or fully repartition to shift the partition boundaries into the interior to complete adaptation. Xie, Seol, and Shephard present the generic programming paradigm built on two decades of research into parallel adaptive data structures and partitioning. This allows the adaptive grid algorithm to
navigate the large associative data structures that change topology as grid elements are added, removed, or modified. Adaptive grid schemes must be capable of supporting large relative motion by adapting the grid or operating on the component grids of an overset grid system. Unstructured spatial plus time grid adaptation requires four-dimension meshes to create a time accurate discretization of the three spatial dimensions.

IV.B. Error Estimation

Mavriplis identifies a critical element of grid adaptation,

Central to the implementation of any solution-adaptive scheme is the ability to detect and assess solution error. The construction of a suitable refinement criterion represents the weakest point of most adaptive strategies.

One way to provide a natural segregation of this critical error estimation and refinement criterion from grid mechanics consists in using a metric-tensor field. It provides general orientation, stretching, and sizing. With this metric, the ideal mesh adheres to the unit mesh concept, where the edges have unit length as measured in the metric. This metric provides a natural way for error estimation techniques to specify the next grid in an iterative adaptive process. There exists different classes of error estimates that were derived with different goals in mind. We describe the following popular classes: feature-based, goal-oriented, and norm-oriented.

Feature-based and Hessian-based. A number of approaches have targeted a feature of the flow for refinement (e.g., vorticity). These methods clearly improve the resolution of these features, but there is no guarantee that the features are not displaced by discretization errors elsewhere in the domain or that resolving these features improve the accuracy of integrated forces and moments. For example, Warren et al. show that an adaptation indicator like the norm of the gradient may predict the wrong location of a shock given coarse initial grids. However, these tools have found utility for problems or simulation tools where the adjoint is not available.

The Hessian of the solution is another popular tool that forms the basis of a number of anisotropic metrics for second-order methods because it models the interpolation error between linear and quadratic functions and can be extended to systems of equations. This Hessian is recovered from the control volume averages of finite-volume solutions or finite-element solutions. The accuracy of Hessian recovery methods has been studied for analytic functions and found to be more problematic on the boundaries. When first- or second-order Hessian reconstruction methods are applied to second-order solutions, the resulting process may be nonconvergent. This is discussed further in section IV.D. The Hessian approach depends only on the numerical solution, so the governing equations are completely ignored during adaptation. From an engineering point-of-view, this class of error estimate has limitations, the Hessian of the numerical solution can not adapt the mesh to target a functional of interest like the lift or the drag. In that case, goal-oriented estimates are preferred.

Goal-oriented. Fidkowski and Darmofal provide a summary of error estimation and grid adaptation strategies to control output error of a simulation. This class of error estimation techniques is based on an adjoint-weighted residual (or truncation error estimate), where the individual schemes differ in how they estimate the adjoint, estimate the residual, and localize contributions to the error estimate. For example, Venditti combines an output estimate with the Hessian to form a grid adaptation metric. There are alternatives to the Hessian that are also valid to higher-order schemes. Fidkowski searches for the direction of the next higher derivative to set the principle direction of the metric. The element stretching is based on the derivative value in the orthogonal direction. This is similar to the approaches of Cao and Dolejši. The adjoint-weighted residual error estimate also has alternatives. Todarello et al. base an error estimate on a norm of the local grid sensitivity to an output. They produce similar adaptation patterns to a method by Dwight that targets artificial dissipation. The continuous mesh framework provides a more direct mathematical connection between estimated error and the metric through the unit mesh concept. This has produced metric formulations to control interpolation error and output error via an optimal goal-based metric. Coupez also proposes controlling interpolation error via a metric constructed from edge error estimates. Yano and Darmofal form surrogate error-metric models and minimize error with the constraint of a specified grid size. This change in the current metric (from a metric implied from
the current grid) is formulated for arbitrary order schemes and produces an anisotropic metric without the use of a reconstructed Hessian or search for the direction of higher derivatives.

**Multi-target or norm-oriented.** A weakness of output- or goal-based methods is that a very complex function may be required that is composed of a linear combination of many properly scaled forces and moments (e.g., prediction of multibody trajectories\(^\text{11}\)). Convergence properties hold for the functional of interest, but there is no convergence rate guarantee of the solution itself. Norm-oriented\(^\text{158}\) and multitarget\(^\text{159}\) approaches have been developed to address these issues without computing an adjoint for each output or a domain integral of the solution.\(^\text{160}\) The entropy-variable adjoint approach\(^\text{161}\) uses a transform of the solution to directly obtain the adjoint solution via symmetrization of the Navier-Stokes equations. This is an inexpensive error estimate for entropy flow out of the domain. The norm-oriented approach is based on the derivation of a nonlinear corrector. The deviation between this corrector and the current numerical solution is used as a functional of interest as in the goal-oriented approach.

The previous estimates are mainly derived for steady simulations. For time accurate problems, the transient fixed-point\(^\text{86,162}\) metric has been extended to include a goal-based metric.\(^\text{163}\) This approach separates time into a number of windows where the grid is constant and interpolates the solution between these fixed grids. Krakos and Darmofal\(^\text{164}\) have extended the approach of Yano and Darmofal\(^\text{77}\) to single-grid, time-accurate simulations. Properties of the time-averaged flow field have been used to adapt eddy resolving turbulent simulations.\(^\text{84}\)

There are a number of concerns that may impact error estimation for more complex problems. The adjoint solution used in many error estimates grows unbounded for chaotic unsteady flows like turbulent eddy resolving simulations. There are some solutions proposed by Larsson and Wang,\(^\text{25}\) but this is still an area of active development. Another issue is nonunique solutions for Euler\(^\text{13}\) and RANS.\(^\text{14}\) Hysteresis is seen in angle of attack for both the Common Research Model\(^\text{82}\) and the NASA Trapezoidal Wing.\(^\text{165}\) An error estimate may target the closest attractor and not the solution that best matches wind tunnel or flight test.

**IV.C. Flow Solver Interactions**

Adapted grid element shape must meet the requirements of the CFD solver. This implies that adapted element shape may need to meet a minimum criteria or the solver may need to be hardened to effectively use irregular, high gradation grids. This includes fast iterative convergence, solver robustness, and high quality derivative quantities (e.g., heating or skin friction). See Shewchuk\(^\text{48}\) for an investigation of linear elements. Sun\(^\text{166}\) examined the impact of element shape on a finite element discretization. Diskin and Thomas\(^\text{167}\) examined the impact of mesh regularity on finite-volume discretizations. Grid element shape and regularity can also impact components of the CFD scheme. For example, gradient reconstruction is impacted.\(^\text{168,169}\) Iterative convergence is noted as a barrier to output-based grid adaptation on a 3D high lift configuration.\(^\text{170}\) To resolve singularities (particularly to support high-order methods) a large gradation in element size is required.\(^\text{78}\) The adapted grid must be provided to the solver in a way that supports the scalability and parallel efficiency of the flow and adjoint solution. This includes any pre- and post-processing required by the flow solver.\(^\text{171,172}\)

**IV.D. System Performance**

The key attribute to any adaptive scheme is how well the entire process converges.\(^\text{173}\) The convergence of the adaptation process may be critical to realizing high-order accurate schemes\(^\text{8}\) and implies that each component in Fig.\(^5\) performs its task satisfactorily and in harmony with the other components. If a nonmetric compliant element is introduced by the grid mechanics, the adaptive process must correct this element in the next adaptive cycle without sabotaging the flow solver, error estimation, or new metric formation. Some error estimates show sensitivity to initial grids or appear to converge to the wrong answer.\(^\text{51,174}\)

One aspect where this harmony is important is in Hessian recovery for some error estimation techniques. While the accuracy of the Hessian recovered from a second-order accurate discrete solution can be suspect, a key question is whether a reconstructed Hessian can be used in place of the true Hessian. Labbé et al.\(^\text{175}\) provide a verification strategy to track a norm of the difference between the reconstructed Hessian and the true Hessian on a sequence of uniformly refined meshes. Agouzal and Vassilevski\(^\text{176}\) also question if the discrete Hessian locally converges toward the differential Hessian and show it does converge for adapted
grids with the boundary Hessian extrapolated from the interior. Lipnikov and Vassilevski indicate the choice of the recovery method must depend on the norm in which the estimated error is controlled. They show that the Hessian reconstruction methods used have a convergence rate predicted by theory on refined grids. Kamenski and Huang also state that the reconstructed Hessian is not convergent for a second-order scheme, i.e., linear finite element methods, but this nonconvergent Hessian can still be used for grid adaptation. The complete adaptive process should be analyzed to ensure it is free of this type of ambiguity.

Metric gradation control reduces noise in the metric field and smoothly increases the influence of small features to a local region. This metric noise can result from errors in Hessian reconstruction, noise in the error estimate, or rapid changes in element size or shape in the current grid. Adapting a grid to conform to a metric is easier for lower gradation metric fields, which increases grid adaptation robustness. Flow solvers generally exhibit better robustness and convergence rates of grids with smooth gradation. This limit on metric variation can have an impact on overall system performance by increasing the robustness of the entire procedure or reducing performance by increasing grid size. Smooth but large gradation in element size is required to resolve singularities in the solution.

The efficiency and scalability of the entire adaptation process can be limited by the poorest performing component. This rules out sequential execution for everything except for trivial processes. These required parallel steps include the initial grid generation, load balancing, error estimation, grid adaptation, and interpolation. The use of mass storage (hard disk) or nonvolatile memory should be avoided because it tends to be much slower than primary storage. The components should communicate via published software interfaces instead of these mass storage devices for maximum throughput and to facilitate interchange of different implementations for testing and comparison.

Multidisciplinary environments must be supported where the flow solver in Fig. 8 is actually an ensemble of different analysis tools. Design requires the computations of increments, which can become buried in the noise created by changes in discretization error between two different grids. This can be addressed formally by including error estimation in the design process or using an error estimate that is robust to small changes. The adaptive process may be used to build large databases for simulation or to form snapshots for reduced order modeling or uncertainty quantification. In these cases, full field (not just a single targeted function) may be preferred.

V. Conclusions

The CFD 2030 Vision Study roadmap anticipates that automated in-situ mesh with adaptive control could replace the typical practice of using a fixed grid by the 2030 timeframe. To accomplish this vision, adaptive unstructured grid computations based on error estimation and control must displace fixed grids as the default so that the CFD practitioner will rarely visualize the grid directly. The maturation of error estimation techniques and grid mechanics will effectively make the grid invisible to the CFD analysis process. This could lead to the certification of analysis or certification by analysis. The impacts of realizing this process change also include improvements to demonstrating asymptotic convergence rates sooner on smaller meshes, building aerodynamic databases, assessing and developing turbulence models, automation of discovery, design optimization, uncertainty quantification, and dramatic improvements to vehicle mission performance.

However, the CFD 2030 Vision Study identifies that mesh generation and adaptivity may continue to be significant bottlenecks in the CFD workflow preventing these impacts because very little government investment has been targeted in these areas. Another complication is that emerging High Performance Computing environments forecast to arrive in the next fifteen years will gain increases in performance through increased concurrency and not an increase in single core capability. This means that without new software investments to ensure efficient execution on these new architectures, existing software could be slower on new machines for the first time (after decades of increasing sequential performance).

Recommended investments are provided for the next fifteen years to address these concerns. These recommendations include cross-cutting technologies (e.g., flow solver robustness, CAD and CAD surrogate geometry, parallel mesh databases, viscous wall functions, software development and testing) and technologies specific to unstructured grid adaptation (e.g., semi-structured regions, high-aspect ratio elements, boundary conforming elements, high geometry curvature and complexity). All disciplines must also invest in software development to ensure efficient performance on emerging computer architectures. Recommendations are also made to expedite the adoption of unstructured grid adaptation technology into production workflows based on case studies and research into technology diffusion. This adoption is critical because impacting production
workflows is ultimately the measure of unstructured grid adaptation success. A detailed description of the elements of unstructured grid adaptation is provided with a bibliography of the last fifteen years to place these recommendations in context with the past velocity of development.

Participation of government, university, and industrial researchers is the most robust approach to accomplish the recommendations for improving unstructured grid adaptation, particularly with the observation that very little government investment has been targeted in these areas. This report is provided as a framework to accomplish the research required to realize the vision and impacts of truly automated geometry-to-solution with uncertainty quantification and discretization error control.

**Appendix A: Vision CFD 2030 Capabilities and Roadmap**

The proposed Vision 2030 capabilities and roadmap are presented to provide context for the status and forecast of unstructured grid method development. Without proper integration with other proposed capabilities, improved adaptation techniques will be of little utility. This includes improvements to increase the capability of attacking larger and more complex problems and increase the capacity to automatically predict an ensemble of conditions. Therefore, unstructured grid adaptation developments must dovetail with the roadmap of other technologies detailed in Fig. and the basic set of capabilities from Slotnick et al. Many of the issues that are expected to arise from using a greater degree of concurrency are also outlined by Löhner and Baum. The Fast Adaptive Aerospace Tools (FAAST) project identified many of the same bottlenecks addressed by these capabilities and roadmap.

1. Emphasis on physics-based, predictive modeling. In particular, transition, turbulence, separation, chemically reacting flows, radiation, heat transfer, and constitutive models must reflect the underlying physics more closely than ever before.

2. Management of errors and uncertainties resulting from all possible sources: (a) physical modeling errors and uncertainties addressed in item #1, (b) numerical errors arising from mesh and discretization inadequacies, and (c) aleatory uncertainties derived from natural variability, as well as epistemic uncertainties due to lack of knowledge in the parameters of a particular fluid flow problem.

3. A much higher degree of automation in all steps of the analysis process is needed including geometry creation, mesh generation and adaptation, the creation of large databases of simulation results, the extraction and understanding of the vast amounts of information generated, and the ability to computationally steer the process. Inherent to all these improvements is the requirement that every step of the solution chain executes high levels of reliability/robustness to minimize user intervention.

4. Ability to effectively utilize massively parallel, heterogeneous, and fault-tolerant HPC architectures that will be available in the 2030 time frame. For complex physical models with nonlocal interactions, the challenges of mapping the underlying algorithms onto computers with multiple memory hierarchies, latencies, and bandwidths must be overcome.

5. Flexibility to tackle capability- and capacity-computing tasks in both industrial and research environments so that both very large ensembles of reasonably-sized solutions (such as those required to populate full-flight envelopes, operating maps, or for parameter studies and design optimization) and small numbers of very-large-scale solutions (such as those needed for experiments of discovery and understanding of flow physics) can be readily accomplished.

6. Seamless integration with multidisciplinary analyses that will be the norm in 2030 without sacrificing accuracy or numerical stability of the resulting coupled simulation, and without requiring a large amount of effort such that only a handful of coupled simulations are possible.

**Appendix B: Partial Bibliography of Unstructured Grid Methods**

The status of unstructured grid adaptation is reviewed in the late 1990s by Baker, Mavriplis, and the proceedings of the ICASE/LaRC Workshop on Adaptive Grid Methods. Jansen, Shephard, and
Beall\cite{Beall2018} provide a summary of the issues associated with the construction, adaptation, and application of anisotropic meshes. Alauzet and Loseille\cite{Alauzet2009} survey the progress made in anisotropic mesh adaptation for CFD in the last decade.

A brief summary of the last fifteen years is provided here to anchor the predictions for the next fifteen years. Forecasting the developments and capabilities of unstructured grid adaptation presents similar challenges as cost estimation, which is hampered by psychological bias but often anchored successfully with historical data.\cite{Beall2018} Progress in most technology is exponential over long time scales, but appears to show linear or parabolic growth when viewed in shorter windows. Progress in CFD is often a series of steps where grand challenge sized problems become routine within a few years or a decade. Grid adaptation techniques typically evolve from 2D to 3D and potentially gain anisotropy, parallel execution, time accuracy, or output error controls.\cite{Alauzet2009}

V.A. Progress 2000–2005

The five year period starting in 2000 saw multiple implementations of mixed element subdivision. These were applied to time-accurate simulations and transient fixed point adaptation was introduced. Linkages to CAD systems were implemented. Parallelization efforts used shared memory and distributed memory paradigms. Anisotropy increased for 2D and 3D applications. Output-based error indicators were implemented for finite-volume discretization.

V.B. Progress 2006–2010

The five year period starting in 2006 saw a number of researchers implement $y^+$ adaptation for stacks of prismatic boundary layer elements. Transient fixed point unsteady adaptation was extended to a 3D anisotropic implementation. The use of cut cells for unstructured anisotropic finite-element and finite-volume output-based adaptation of triangles and tetrahedra was applied to Euler and laminar simulations. The entropy adjoint was identified as an inexpensive error estimate.
V.C.  Progress 2011–2015

The most recent five year period saw application to larger and more complex cases, particularly the AIAA drag, high lift, and sonic boom prediction workshops. Transient fixed point adaptation was extended to goal-based error estimates. Local edge based operators continued to mature and were unified by the cavity based operator. The insertion of regions of metric orthogonal grid was demonstrated. Adaptive grid simulations with large geometry displacements and overset grid systems were demonstrated. Design with grid adaptation was shown by a few researchers. The verification of adaptive schemes with turbulent benchmark cases was accomplished in 2D and 3D. A number of applications included eddy resolving unsteady adapted grids based on the average flow field or metric intersection. The convergence of the output-based adaptive process was examined with multiple metric formulations and multiple implementations of grid adaptation mechanics.

V.D.  Chronological Citations

Here is a chronological list of citations that begins in the year 2000. A brief description is provided for each reference.

2000  
Li, Shephard, and Beall develop methods for robustly recovering curved boundaries in 3D adapted grids.
Mavriplis performs subdivision on mixed-element grids.
Coupez, Digonnet, and Ducloux show parallel isotropic grid adaptation.
Minyard and Kallinderis investigate load balancing issues for parallel tetrahedral grid adaptation.
Habashi et al. discuss general principles related to anisotropic mesh adaptation.

2001  
Becker and Rannacher detail the dual-weighted-residual method for optimal output error control.
Pain et al. perform anisotropic tetrahedral grid adaptation for a time accurate simulation.

2002  
Dompierre et al. show 2D anisotropic metric-based adaptation based on solution Hessian that is robust to initial grids based on the principles in Habashi et al.
Venditti and Darmofal performed 2D isotropic grid adaptation based on an embedded grid output-based error estimate.
Park extended the isotropic Venditti error estimate to 3D inviscid geometries with a linkage to CAD via CAPRI and the GridEx framework.
Lepage, Suerich-Gulick, and Habashi show 3D anisotropic grid adaptation with stacks of prismatic boundary layer elements.

2003  
Beall, Walsh, and Shephard detail the issues associated with accessing CAD geometry for grid generation and adaptation.
Alauzet et al. introduces transient fixed point unstructured mesh adaptation with isotropic grids.
Venditti and Darmofal extend their embedded grid output-based error estimate to an anisotropic metric with a scaled Mach Hessian and perform 2D turbulent calculations.

2004  
Lepage, St-Cyr, and Habashi parallelize the 3D anisotropic grid adaptation method with domain decomposition.
Waltz shows unsteady parallel adaptive isotropic grids for time accurate problems on a shared memory architecture.
Park implements the Venditti anisotropic metric and applies it to very mild 3D anisotropic Euler adaptation and regenerated turbulent grids.

2005  
Lee-Rausch et al. further extended the Venditti anisotropic metric to 3D turbulent flows with a parallel implementation and grid mechanics with the capability for higher aspect ratio tetrahedra.
Cavallo and Grismer\textsuperscript{133} present parallel 3D mixed element isotropic refinement.

Park and Kwon\textsuperscript{195} present parallel 3D mixed element isotropic refinement for time accurate simulation.

Kallinderis and Kavouklis\textsuperscript{196} present parallel 3D mixed element isotropic refinement and coarsening for steady and time accurate simulations.

2006 Jones, Nielsen, and Park\textsuperscript{197} apply 3D Euler anisotropic output-based adaptation to near-field sonic boom prediction of axisymmetric bodies.

Bibb et al.\textsuperscript{147} apply 3D parallel anisotropic feature-based grid adaptation to reentry vehicle.

Gorman\textsuperscript{43} document 3D parallel anisotropic adaptation for time accurate problems.

Alrutz and Orlt\textsuperscript{198} detail parallel hybrid refinement.

Kim, Takano, and Nakahashi\textsuperscript{199} combine an output-based indicator with element subdivision projected to a reconstructed surface. Kim and Nakahashi\textsuperscript{200} extend this approach to viscous flows and stack of prismatic elements.

Alauzet et al.\textsuperscript{42} develop parallel anisotropic adaptation methods for curved geometries.

2007 Loseille et al.\textsuperscript{4} demonstrate second-order spatial convergence with 3D anisotropic adaptation to control interpolation error for cases with discontinuities.

Alauzet et al.\textsuperscript{201} extend transient fixed point\textsuperscript{162} to 3D anisotropic metrics and adaptation.

Fidkowski and Darmofal\textsuperscript{202} develop a higher-order cut cell method with an anisotropic metric oriented in the next higher-order gradient direction.

2008 Sahni et al.\textsuperscript{54} combine first problem plane projection\textsuperscript{127} with local operator adaptation of prismatic element stacks to resolve turbulent flows in 3D.

2009 Alrutz and Vollmer\textsuperscript{203} show examples of $y^+$, hierarchical, parallel, and goal-based grid adaptation.

Loseille and Löhner\textsuperscript{41} apply local edge operators to adapt 3D boundary layers.

Allmaras et al.\textsuperscript{204} describe a production level 2D fully adaptive RANS solver.

Persson and Peraire\textsuperscript{71} use nonlinear mechanics to guarantee validity of curved linear mesh.

2010 Sahni et al.\textsuperscript{73} present adaptive procedures to generate curved semi-structured boundary layer meshes.

Park and Carlson\textsuperscript{23} apply parallel anisotropic grid adaptation outside of a frozen boundary to the First Shock Boundary-Layer Interaction Workshop and to resolve pressure signature of a supersonic boattail and plume.

Bartels et al.\textsuperscript{205} apply frozen boundary adaptation\textsuperscript{23} to a launch vehicle.

Alkandry et al.\textsuperscript{148} apply output- and feature-based frozen boundary adaptation\textsuperscript{23} to hypersonic reentry vehicle with decelerator jet.

Loseille and Löhner\textsuperscript{206} adapt an Euler solution to propagate a supersonic business jet signature to the ground where new points are projected to CAD surface.

Park and Darmofal\textsuperscript{207} apply an Euler cut cell method with output-adapted tetrahedral background grids to near-field sonic boom prediction of complex configurations.

Kavouklis and Kallinderis\textsuperscript{62} present parallel 3D mixed element isotropic refinement and coarsening for steady and time accurate simulations.

Fidkowski and Roe\textsuperscript{161} present an error estimation technique based on the entropy adjoint.

Alauzet and Mehrenberger\textsuperscript{85} present conservative interpolation and show it reduces mass variation.
2011 Farrell and Maddison\textsuperscript{104} present a conservative interpolation scheme for discontinuous fields.

Yano, Modisette, and Darmofal\textsuperscript{6} illustrate the importance of grid adaptation for higher-order schemes.

Park, Lee-Rausch, and Rumsey\textsuperscript{170} apply a single grid turbulent output-based error estimate to the First High Lift Prediction Workshop.

Loseille and Alauzet\textsuperscript{79, 208} publish detailed references for the Continuous Mesh Framework.

2012 Belme, Dervieux, and Alauzet\textsuperscript{163} extend the transient fixed point approach to anisotropic goal-based error estimate.

Michal and Krakos\textsuperscript{44} demonstrate anisotropic tetrahedral adaptation using only edge split and collapse operations.

Yano and Darmofal\textsuperscript{77, 78} use surrogate error-metric models with a cost constraint to create an optimized metric from an implied metric for arbitrary order schemes.

Krakos and Darmofal\textsuperscript{164} extend the Yano and Darmofal\textsuperscript{77, 78} approach to unsteady problems.

Loseille and Löhner\textsuperscript{209} introduce the cavity operator for boundary layer meshing and boundary recovery.

2013 Loseille and Menier\textsuperscript{45} unify local grid adaptation operations with the cavity primitive.

Ovcharenko et al.\textsuperscript{210} parallelize anisotropic grid adaptation with semi-structured meshes of stacks of layered elements for boundary layer meshes.

Hoffman et al.\textsuperscript{83} combine parallel isotropic refinement with dual-weighted residual error estimates of time-accurate simulations.

2014 Salah El Din, Dagrau, and Loseille\textsuperscript{211} apply anisotropic grid adaptation to the First Sonic Boom Prediction Workshop.

Park et al.\textsuperscript{16} combine prismatic boundary layer subdivision with the Michal and Krakos\textsuperscript{44} tetrahedral anisotropic adaptation scheme for the Fifth Drag Prediction Workshop.

Vatsa et al.\textsuperscript{141} use a time intersected metric and parallel anisotropic adaptation to predict landing gear noise.

Shenoy, Smith, and Park\textsuperscript{141} extend parallel anisotropic adaptation with a time intersected metric to component grids of an overset grid system.

Rasquin et al.\textsuperscript{212} perform eddy resolving turbulent simulation on 11 and 92 billion element adapted meshes on 768K cores.

Elmiligui at al.\textsuperscript{213} perform steady turbulent grid adaptation to resolve vorticies of a fighter configuration.

Lee-Rausch, Park, and Rumsey\textsuperscript{214} and Chitale et al.\textsuperscript{215} perform steady-state calculations on adapted grids for the Second High-Lift Prediction Workshop test cases.

Rasquin et al.\textsuperscript{84} perform eddy resolving turbulent calculations adapted to vorticity of time-averaged flow field with 1.3 billion tetrahedra for the Second High-Lift Prediction Workshop test cases.

Dolejši\textsuperscript{80} describes an anisotropic metric and variable order elements to control interpolation error.

Loseille\textsuperscript{46} extend cavity-based operators to create adaptive metric-orthogonal anisotropic grids.

Alauzet\textsuperscript{216} adapt grids to follow large displacements of the geometry.

2015 Barral, Alauzet, and Loseille\textsuperscript{86} extend Alauzet\textsuperscript{216} to metric-based anisotropic mesh adaptation.
Loseille, Dervieux, and Alauzet\textsuperscript{158} publish a norm-based error estimate that simultaneously controls multiple functionals of interest.

Park et al.\textsuperscript{173} study the convergence of the entire adaptive process for analytic functions and CFD solutions with multiple metric-based error estimation techniques and grid mechanics.

Menier\textsuperscript{217} accelerates parallel anisotropic grid adaptation with multigrid and verifies the process with a suite of published turbulent test cases.

Loseille and Menier\textsuperscript{218} perform parallel anisotropic grid adaptation to 1 billion tetrahedra on 120 cores.

Heath et al.\textsuperscript{219} perform design on anisotropically adapted grids for turbulent nozzle flows.

Rokos et al.\textsuperscript{220} exploit thread parallelism during anisotropic mesh adaptation.

Ceze and Fidkowski\textsuperscript{113} implement Yano and Darmofal\textsuperscript{77} to compare to hanging node subdivision of curvilinear grids.

Hu et al.\textsuperscript{221} apply Yano and Darmofal\textsuperscript{77} to a number of turbulent verification tests and show the graduation required to resolve singularities.

Alauzet and Loseille\textsuperscript{36} survey the progress made in anisotropic mesh adaptation for CFD during the last decade.

\textbf{2016} Alauzet\textsuperscript{103} presents a parallel conservative interpolation scheme and show it reduces space-time error, which grows for linear interpolation.

Sahni et al.\textsuperscript{222} extends parallel anisotropic grid adaptation for semi-structured meshes with stacks of layered elements for billions of elements.

\section*{Appendix C: Strategy for Diffusion of Unstructured Grid Adaptation Technology}

Beyond the capabilities demonstrated in publications, impacting production workflows is ultimately the measure of unstructured grid adaptation success. The stages of CFD penetration into the design process, current status, and issues preventing adoption are compiled by Malik and Bushnell.\textsuperscript{223} Johnson, Tinoco, and Yu\textsuperscript{29} illustrate how CFD developments are adopted by application engineers at Boeing Commercial Airplanes. This evidence for the diffusion of CFD innovation has been shown to be present in diverse fields and could serve as a model for the adoption of unstructured grid adaptation techniques. A pivotal study of the adoption of hybrid seed corn by Griliches\textsuperscript{224} was generalized with studies of the adoption of other technologies by Rodgers.\textsuperscript{225} This diffusion model has been studied and refined by many researchers. For example, Bass\textsuperscript{226} modeled the likelihood that a consumer would purchase a durable good and integrated this likelihood to predict total sales of those goods. The cumulative distribution function of these sales depicts an S-shaped growth curve.

Johnson, Tinoco, and Yu\textsuperscript{29} describe the adoption of each new generation of CFD technology as having five phases. The first three phases result in a minor amount of technology adoption by the product engineers, but the last three phases show an S-shaped adoption curve that ends in saturation. Ortt and Schoormans\textsuperscript{227} discuss the processes that come before the classic bell-shaped S-curve of the diffusion of innovation for a number of technologies. These phases are labeled the Innovation Phase, Market Adaptation Phase, and Market Stabilization Phase (the S-shaped pattern). They suggest the existence of these phases “indicate that it is important to establish the position of the technology in the pattern of development and diffusion and that strategies should be tailored to this position.”\textsuperscript{227}

A particularly good insight is that invention results from basic research, which is often monodisciplinary. The later technology phases usually require multidisciplinary cooperation among various actors outside the research department that produced the invention. For example, the tasks of geometry preparation, grid generation, and flow solution are typically performed by groups of individuals that specialize in a subset of tasks. An environment where unstructured grid adaptation is ubiquitous may encourage a different skill mix than these specialized practitioners. The Market Adaptation Phase often shows multiple false starts and is
hampered by the chicken-and-egg problem until a critical mass enables the Market Stabilization Phase to begin. Moore\textsuperscript{228} describes the gap between these phases as a Chasm.

Agarwal and Bayus\textsuperscript{229} provide historical evidence that the lag time between invention and diffusion is 4–30 years and the technology’s “take-off is caused by outward shifting supply and demand curves.” Where the supply of a technology is marked by the number of institutions that make a firm entry of a product into the market and (to a lesser degree) a price decrease. Translating this observation to unstructured grid adaptation would suggest that multiple research groups and commercial vendors providing grid adaptation technology is a more important factor than the efficiency of a particular implementation to trigger rapid adoption.

Hall and Khan\textsuperscript{230} observe that “at any point in time the choice being made is not a choice between adopting and not adopting but a choice between adopting now or deferring the decision until later.”

The primary implication of this way of looking at the problem is that there is “option value” to waiting; that is, adoption should not take place the instant that benefits equal costs, but should be delayed until benefits are somewhat above costs (that is, one invests when the option is “deep in the money”), thus providing yet another reason why diffusion may be rather slow.\textsuperscript{230}

Adoption is further slowed by the perceptions that the opportunity to adopt will persist, the successful implementation of the technology requires complex new skills, or acquiring the required level of worker competence is time consuming. This is also observed by Johnson, Tinoco, and Yu.\textsuperscript{29} “Every new or proposed CFD capability was initially viewed as too difficult to use, too costly to run, not able to produce timely results, not needed, and so on.” The technology must ultimately bring value to the customer via the product.

In industry, CFD has no value of its own. The only way CFD can deliver value is for it to affect the product. To affect the product, it must become an integral part of the engineering process for the design, manufacture, and support of the product. Otherwise, CFD is just an add-on; it may have some value but its effect is limited.\textsuperscript{29}

This technology diffusion model is appropriate for large aerospace companies, government agencies, and commercial vendors that are inherently conservative. There is a current trend to apply lean manufacturing techniques to innovation.\textsuperscript{231} These approaches can be complementary, where a minimum viable product begins rapid iteration with feedback from customers. These rapid iterations can hone candidate technologies to provide a polished product, which gains a foothold into the rapid adoption phase of the diffusion model.

Often [developers] spotted “niche” needs that could be satisfied by the introduction of their new technology. It was felt that when the users were satisfied with the usability and utility of the technology in these areas they would then be willing to consider whether or not replacing their old tools in other areas might offer distinct advantages. Once the users accepted a new capability, they often became very innovative and applied the codes in unanticipated ways, perpetually keeping the developers and validation experts in an anxious state. Most of the new applications were, in fact, legitimate, and the developers had to run fast to understand the implications involved as well as to try and anticipate future application directions.\textsuperscript{29}

There are a number of case studies where vertically integrated processes have provided value to production engineers. Their success stems from addressing the range of issues from geometry to solution. TetrUSS\textsuperscript{232} provides vertical integration of many technologies including geometry preparation, grid generation, flow solution, visualization, grid adaptation, and knowledge-based design. Wick et al.\textsuperscript{233} list critical enhancements made to this system that enable the evaluation of integrated distributed propulsion technologies. Cart3D\textsuperscript{34} combines discrete geometry Boolean operations with output-adaptive Euler cut cells to robustly populate aerodynamic databases. Allmaras et al.\textsuperscript{204} describe the key ingredients that go into a production 2D aerodynamics prediction tool. This work was motivated by the success of TRANAIR.\textsuperscript{234}

From our experience with TRANAIR grid adaption seems essential. The accuracy, nonlinear convergence and grid generation benefits are too valuable to dismiss in a production environment where thousands of simulations are carried out. With respect to grid generation benefits, the initial grid need only fill space and fit the vehicle geometry or at least the vehicle topology.\textsuperscript{204}
These approaches share a number of attributes. Output-based grid adaptation is implemented in a number of these tools. Complex 3D geometry is addressed by either the use of cut cells or translating geometry into a form that can be natively modified in the system. To streamline production aerodynamic prediction capability, complex and rapidly accessed geometry (ideally from a CAD) must be accommodated. Controlling discretization and iterative errors is a necessity for routine exploration of thousands of simulations.

Acknowledgments

The authors would like to thank Chris Rumsey (NASA Langley), Philippe Spalart (Boeing Commercial Airplanes), Daniel Ibanez (Rensselaer Polytechnic Institute), Onkar Sahni (Rensselaer Polytechnic Institute), Juliette Pardue (National Institute of Aerospace/Old Dominion University), and Bil Kleb (NASA Langley) for their constructive comments. Michael Park would like to thank the Transformational Tools and Technologies (TTT) Project of the Transformative Aeronautics Concepts Program (TACP) for supporting this work.

References


American Institute of Aeronautics and Astronautics