Data-Driven Surrogate-Assisted Evolutionary Fluid Dynamic Optimisation

This research proposal aims to permit the application of evolutionary algorithms, a class of global search metaheuristics, to fluid dynamic optimisation of highly complex industrial systems by exploiting surrogate models and modern machine learning techniques.

Environment and Track Record

This project forms a collaboration between two of the UK's leading groups, both consisting of experts in evolutionary optimisation and computational fluid dynamics.

University of Exeter The AI and Machine Learning research group in the College of Engineering, Mathematics and Physical Sciences (CEMPS) has a strong publication record and an excellent RAE 2008 result in which over 83% of the research outputs were rated as 3* or above (ranking 3rd in the UK on this metric). The group, led by Prof. Everson, is interdisciplinary with both theoretical and applied research crossing the boundaries of Computer Science, Mathematics and Engineering. Their work encompasses statistical machine learning and optimisation of difficult industrial problems, involving work with industrial partners such as Motorola and the National Air Traffic Service (NATS).

Dr Tabor's CFD research group is part of the Centre for Water Systems in CEMPS, an internationally regarded grouping consisting of 4 professors and 7 lecturers working on Sustainable Urban Drainage and water resource problems. Core activities include high performance modelling involving CFD, 2D flood modelling, pipe network modelling and the application of GA optimisation techniques in these areas.

Faculty of Engineering and Physical Sciences, University of Surrey In RAE2008, two thirds of FEPS activity was rated internationally excellent, and 20% is world leading. FEPS has now 170 academic staff and approximately 200 postdoctoral research assistants and over 600 post-graduate researchers. The Nature Inspired Computing and Engineering Group, which Prof. Jin leads, has seven permanent academic staff, four research fellows, and approximately 25 PhD students working in cross-disciplinary areas including evolutionary computation, machine learning, computational neuroscience and computational biology. Research of the NICE group has been supported by funding agencies like EU, EPSRC, RAE, Leverhulme Trust, as well as many industries. Powerful and up-to-date computer clusters are available in FEPS.

Richard Everson is a Professor of Machine Learning. His research interests focus on statistical pattern recognition and multi-objective optimisation and the interactions between them. Work in statistical pattern recognition of particular relevance to this proposal is the introduction of a new framework for active learning for Bayesian probabilistic classifiers [1, 2], which we propose to extend here for active learning of surrogate models. Recent work has also shown how to use automatic relevance determination and variational Bayesian approximations for robust learning in non-stationary environments [3–5] which we plan to extend for the learning of surrogate models. In multi-objective optimisation he has developed state-of-the-art simulated annealing methods for [6], multi-objective optimisation of receiver operating characteristics [7], optimisation in the presence of uncertainty [8, 9], and the visualisation of multi and many-objective populations [10, 11]. Methods for the optimisation of short term conflict alert for air traffic control [12, 13] have been implemented for use across UK airspace via a KTP project.

Jonathan Fieldsend is well known for his theoretical and applied work on evolutionary multi-objective optimisation [6–13] and was named as one of the 33 primary researchers in the field in the Coello Coello, Lamont and Van Veldhuizen textbook on multi-objective optimisation. He has also published widely in the field of machine learning, and the interface between the two areas. He has over 50 refereed publications and a multi-objective optimisation patent with Motorola.

Fieldsend has been the University Supervisor on two successful KTP projects using multi-objective optimisation, working with NATS (on optimising short term conflict alert for air traffic control) and AI Corp (developing and optimising classifiers for fraud detection). Both KTPs resulted in software tools now in use by the companies involved. He is also the university supervisor for a third ongoing KTP using multi-objective optimisation algorithms for low-power radio networks.

Gavin Tabor is a Senior Lecturer in Mechanical Engineering the College of Engineering, Maths and Physical Science, University of Exeter. A specialist in CFD, he has worked with and contributed to the OpenFOAM open source CFD project which is one of the codes that will be used in this project and is an acknowledged expert in its use [14]. His research covers both the fundamental development of CFD methodologies (e.g. turbulence modelling [15]) and their application to various industrial flow problems [16, 17]. He has been involved in numerous EPSRC and other projects (4 as P.I.), including a current STREAM Eng.D project on optimisation of wastewater treatment systems using adjoint solutions and CFD.

Yaochu Jin is a Professor of Computational Intelligence, Head of the Nature Inspired Computing and Engineering Group, Department of Computing, University of Surrey. He has been engaged in the area of Computational Intelligence and its industrial applications for over 20 years. He has published over 160 peer-reviewed papers, which have received more than 6000 citations. He holds seven EU/US/Japan patents on evolutionary optimisation. He is an Associate Editor of eight international journals, including four IEEE journals. He is an IEEE Distinguished Lecturer and an AdCom member of the IEEE Computational Intelligence Society. He has delivered keynote speeches at 16 international conferences and was invited to give a course on surrogate assisted evolutionary optimisation at the 22nd Jyvaskyla Summer School. His research has been supported by EC FP7, EPSRC, Bosch UK, HR Wallingford, Aero Optimal and and Honda Europe.

He is one of the pioneering researchers on surrogate assisted evolutionary optimisation. He proposed a number of earlier ideas on surrogate assisted evolutionary optimisation [18–20], and developed several state of the art algorithms [21–25]. He has led research projects on surrogate assisted evolutionary aerodynamic optimisation of complex systems, such as turbine blades, F1 race cars and aircraft fuselages. He is a member of the EC GARTEUR Action Group on Surrogate Assisted Global Optimisation and maintains close connections with internatioanlly leading researchers on surrogate-assisted evolutionary optimization, such as Prof Y.-S. Ong, Nangyang Technological University.

John Doherty joined the University of Surrey in 2012 as Reader in the Aerodynamics and the Environment Research Group within Mechanical Engineering Sciences. He was previously Technical Fellow for Aerodynamic and Multidisciplinary Design Optimisation at QinetiQ (RAE, DERA). He is a leading expert in aerodynamic/CFD shape optimisation and his CODAS capability has been used widely in industry, including Airbus (A380, A340-600 and A350 wing design). A central research theme has been the integration of design knowledge, through novel parameterisation and objective/constraint formulations, for efficient exploration and optimisation. He was PI for Integration and Optimisation in the TSB Integrated Wing project and CI with Airbus for Integration in the TSB Multi-Disciplinary Optimised Wing project.

References

[1] R Fredlund, R Everson, and J Fieldsend. "A Bayesian Framework for Active Learning". In: IJCNN. 2010.

[2] R.M. Fredlund. "A Bayesian expected cost reduction approach to active learning". PhD thesis. University of Exeter, 2011.

[3] J. Christmas and R.M. Everson. "Robust autoregression: Student-t innovations using variational Bayes". In: IEEE Trans. Sig. Proc. 59.1 (2011), pp. 48–57.

[4] A. Clark and R.M. Everson. "Multi-objective learning of Relevance Vector Machine classifiers with multi-resolution kernels". In: Pattern Recognition 45 (2012), pp. 3535–3543.

[5] J. Christmas, R. M. Everson, Rodríguez-Muñoz, and T. Tregenza. "Variational Bayesian Tracking: Whole Track Convergence for Large Scale Ecological Video Monitoring". In: International Joint Conference on Neural Networks. 2013.

[6] K I Smith, R M Everson, J E Fieldsend, R Misra, and C Murphy. "Dominance-based multi-objective simulated annealing". In: IEEE T-EVC 12 (2008), pp. 323–343.

[7] R. Everson and J. Fieldsend. "Multi-class ROC analysis from a multi-objective optimisation perspective". In: Patt. Rec. Let. 27 (2006), pp. 531–556.

[8] J E Fieldsend and R M Everson. "Multi-objective optimisation in the presence of uncertainty". In: CEC'05. 2005, pp. 476–483.

[9] J. E. Fieldsend and R. M. Everson. "The Rolling Tide Evolutionary Algorithm: A Multi-Objective Optimiser for Noisy Environments". In: IEEE T-EVC (2013). (Submitted).

[10] D.J. Walker, R.M. Everson, and J.E. Fieldsend. "Ordering and Visualisation of Many-objective Populations". In: IEEE CEC. 2010.

[11] D.J. Walker, R.M. Everson, and J.E. Fieldsend. "Visualising Mutually Non-dominating Solution Sets in Many-objective Optimisation". In: IEEE T-EVC 17.2 (2013), pp. 165–184.

[12] R.M. Everson and J.E. Fieldsend. "Multi-objective optimisation of safety related systems: An application to Short Term Conflict Alert". In: IEEE T-EVC 10.2 (2006), pp. 187–198.

[13] W.J. Reckhouse, J.E. Fieldsend, and R.M. Everson. "Variable interactions and exploring parameter space in an expensive optimisation problem: Optimising Short Term Conflict Alert". In: IEEE CEC. 2010.

[14] H. G. Weller, G. Tabor, H. Jasak, and C. Fureby. "A Tensorial Approach to Computational Continuum Mechanics using Object Orientated Techniques". In: Computers in Physics 12.6 (1998), pp. 620 – 631.

[15] L. Jiang, G. Tabor, and G. Gao. "A new turbulence model for separated flows". In: Int.J.Comp.Fluid Dyn. 25.8 (2011), pp. 427 –438.

[16] M. J. Baker and G. R. Tabor. "Computational analysis of transitional airflow through packed columns of spheres using the finite volume technique". In: Computers and Chemical Engineering 34 (2010), pp. 878 –885.

[17] G. Queguigneur, D. Jarman, E. Paterson, and G. Tabor. "Computational investigation of vortex flow controls at low flow rates". In: I.C.E. Proc: Engineering and Computational Mechanics (2013).

[18] M. Huesken, Y. Jin, and B. Sendhoff. "Structure optimization of neural networks for aerodynamic optimization". In: Soft Computing 9.1 (2005), pp. 21–28.

[19] Y. Jin, M. Olhofer, and B. Sendhoff. "On evolutionary optimization with approximate fitness functions". In: GECCO. 2000, p. 786.

[20] Y. Jin, M. Olhofer, and B. Sendhoff. "A framework for evolutionary optimization with approximate fitness functions". In: IEEE T-EVC 6 (2002), p. 481.

[21] N.M. Le, Y.S. Ong, S. Menzel, Y. Jin, and B. Sendhoff. "Evolution by adapting surrogates". In: Evolutionary Computation (2012).

[22] D. Lim, Y. Jin, Y.S. Ong, and B. Sendhoff. "Generalizing surrogate-assisted evolutionary computation". In: IEEE T-EVC 14.3 (2010), pp. 329–355.

[23] X. Sun, D. Gong, Y. Jin, and S. Chen. "A new surrogate-assisted interactive genetic algorithm with weighted semi-supervised learning". In: IEEE Trans. Sys., Man & Cybernetics, Part B 43.2 (2013), pp. 685–698.

[24] C. Smith, J. Doherty, and Y. Jin. "Recurrent neural network ensembles for convergence prediction in surrogate-assisted evolutionary optimisation". In: IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments. IEEE, 2013.

[25] J. Lu, B. Li, and Y. Jin. "An evolution strategy assisted by an ensemble of local Gaussian process models." In: Genetic and Evolutionary Computation Conference. 2013.

Part 2: Proposed Research

Background

Computational optimisation is crucial across the breadth of science, commerce and industry: it ensures the safety of aircraft, reveals new solutions, new products and drugs, and reduces energy consumption. Optimisation has been part of engineering pratice for centuries, carried out in a heuristic manner. Frequently, however, there are a great many (hundreds) of design variables (e.g., the topology, shape and deployment of wing high-lift systems) that may be optimised to find the optimum value of the objective. In addition, there is often more than one objective (e.g., aerodynamic performance, fuel economy and emission reduction) and it is important to find a set of Pareto-optimal solutions describing the best trade-off between the objectives. The large number of parameters that may be optimised and the often highly nonlinear and sometimes discontinuous relation between the parameters and the objective function(s), renders traditional mathematical programming techniques ineffective in enhancing performance and revealing new solutions.

Evolutionary algorithms (EAs) are very attractive with their global search ability and high efficiency for achieving a set of Pareto-optimal solutions. In addition, EAs do not require exact mathematical models of the system to be optimised. However, as a population based search approach, EAs need a large number of fitness evaluations to achieve acceptable solutions.

There are many optimisation problems where the evaluations of the objectives for each solution can be prohibitively expensive. For example, fluid dynamic optimisation that involves the application of computational fluid dynamics (CFD) simulations is a typical class of expensive optimisation problems found in a wide range of important industrial problems, for instance in design of aircraft, cars and internal combustion engines such as diesel engines or gas turbines. For example, a 3-D CFD simulation will often take hours or even days even on high-performance computers.

To deal with computationally expensive optimisation problems, surrogate-assisted evolutionary optimisation has attracted attention over the past decade [26, 27]. For surrogate-assisted evolutionary optimistion to be successful, the surrogate must be properly combined with the precise, yet expensive, fitness evaluations, which is often known as model management [26]. Various model management strategies have been developed, which can be largely divided into four categories, namely, population-based [28], generation-based [e.g., 20], individual-based [e.g., 20, 29], and local search embedded in an EA [22]. It is also extremely important to choose a proper model as the surrogate. Various machine learning models, such as polynomials (response surface methodology), artificial neural networks [e.g., 20], and Gaussian processes [30] have been used. Multiple models or ensembles [e.g., 31, 32] have shown to outperform a single model. Recently we showed that a combination of local models and global models that smoothen the fitness landscape is beneficial [22]. Meanwhile, surrogates of multi-fidelity [33–35] and on-line selection of surrogates have demonstrated to be effective [21].

Despite the fact that fluid dynamic optimisation represents a large class of expensive optimisation problems and is of paramount importance to industry, not much work has been dedicated to surrogate-assisted fluid dynamic optimisation. One reason is that most surrogate-assisted EAs have been applied to relatively small-scale problems (less than 10 dimensional) [20, 30], as the approximation quality of surrogates will become very poor when the dimension of the decision space becomes high. Unfortunately, most fluid dynamic optimisation such as wing high-lift systems often involves shape optimisation, where a large number of decision variables are to be optimised. In addition, most fluid dynamic optimisation has more than one objective to optimise; consequently, surrogate techniques for multiple objectives need to be developed. Apart from these, fluid dynamic optimisation often needs to take into account the extremes of the operating envelope or at least the off-design conditions, which ultimately drives selection of practical or novel designs. This makes it important to search for robust optimal solutions [36]. Search for robust optimal solutions often incurs additional fitness evaluations and our previous work showed that surrogate models can effectively reduce computational cost [37].

Evolutionary fluid dynamic optimisation poses several challenges where surrogate models can play a key role. Breakthroughs in surrogate assisted evolutionary fluid dynamics optimisation can not only create significant impact on industry, but also deepen understanding of how to seamlessly integrate model selection strategies and advanced learning techniques with evolutionary optimisation, and more generally provide new insights into the interactions between optimisation and learning.

Research Hypothesis and Objectives

This proposal focuses on developing efficient and effective surrogate-assisted evolutionary algorithms for fluid dynamic optimisation. Compared to other expensive optimisation problems, evolutioanry fluid dynamic optimisation poses serious challenges yet also creates uniques opportunities.

This adventurous project has the following aims:

1. To develop new surrogate models that can work effectively for high-dimensional problems by taking advantage of advanced learning techniques and the iterative nature of the CFD simulations in fluid dynamic optimisation.
2. To understand the influence of learning techniques and training sample selection on the surrogate quality as well as on the balance between exploration and exploitation in evolutionary search.
3. To apply the developed surrogate-assisted evolutionary optimisation algorithms to solve important industrial fluid dynamic optimisation problems including wing high-lift systems, cylone separation, and diesel engines.

To achieve these aims, the project has the following objectives:

1. Constructing a suite of artificial high-dimensional test problems that can closely simulate the properties and challenges in fluid dynamic optimisation, including off-design conditions, multi-objective problems and mix-integer problems. These test problems will reflect challenges in real-world fluid dynamic optimisation problems yet make the following investigations computationally manageable.
2. Identifying the most efficient modelling techniques for single and multi-objective optimisation of large design spaces. To address the curse of dimensionality, new modelling approaches, such as learning to determine rank and relevance, to classify convergence patterns and predict fluid dynamic performance after convergence will be investigated. In addition, advanced learning techniques such as active learning [e.g., 1] that can help choose the most informative training data, semi-supervised learning that can exploit unlabelled data as we shown recently [23], transfer learning [38] or multi-task learning [39] that can enhance learning from similar tasks [18] and ensemble technique will be employed.
3. Recognising the need for real-world solutions to operate over a range of conditions and provide at least adequate performance "off-design", we will develop new surrogate-based methods for assessing and optimising robust solutions.
4. New model management that can seamlessly integrate various modelling techniques including fitness, rank and convergence predictions will be developed.
5. Applying surrogate-assisted EAs developed in this project to real-world CFD problems, covering a diverse range of applications and gaining cross-fertilisation between the different applications.

Programme and Methodology

To achieve the ambitious aims and objectives of the proposal, the following interlinked work packages are planned:

WP1 Constructing Test Problems. To allow comprehensive investigation of the performance of surrogate-assisted methods we plan to develop a suite of extensible test problems with a wide variety of characteristics closely mimicking real-world single and multi-objective problems, which integrate prior optimisation experience from our industrial partners. Synthetic test problems have the advantages of being easy to describe, understand and visualise and they are generally fast to evaluate. Importantly, the problem characteristics can be preserved as the dimension of the design and objective spaces is manipulated from the small and easy to understand and visualise to realistic sizes.

We will build three categories of test problems. The first class of test suite will be optimisation functions with high-dimensional design spaces, multi-modal, linear and non-linear constraints and with mixtures of integer and continuous design variables. The second category of test problems to be constructed are functions involving the solution of ordinary differential equations that can closely simulate the iterative nature of computational fluid dynamics solvers. This suite will enable us to address the challenges that are particular to fluid dynamic optimisation. The final set of test problems will be very close to real-world fluid dynamic optimisation. We will use CFD solvers made available by our industrial partners and construct a number of simulators having grid resolutions ranging from coarse to fine. This suite of test problems will provide us close to real-world test problems with an easily controllable computational complexity and adjustable scale of the problem to be optimised.

WP2 Surrogate modelling for CFD convergence classification and prediction. Quality evaluations in fluid dynamic optimisation is computationally intensive mainly because they need to numerically solve a large number of the partial differential equations involving thousands of iterations. Traditionally, a surrogate is constructed to map the design space (e.g., geometry of a wing) to the fitness of the design, e.g., the lift/drag efficiency. One major difficulty of building surrogates in such examples is the high-dimensional design space, where tens or hundreds of decision variables are needed to represent the geometry of the wing. Due to lack of training data, the model quality is often very low.

This WP aims to develop novel data-driven surrogate techniques by taking advantage of the iterative nature in numerical solutions of the fluid dynamics. The basic new idea here is treat the intermediate CFD simulation data as time series and fully exploit the information embedded in these data. This WP is divided into the following three subtasks.

WP2.1 Early classification of CFD convergence pattern. In many complex aerodynamic design tasks such as high-lift wing design, CFD analysis is often a challenging task in that CFD simulations are less robust and take extremely long time to converge. Experienced aerodynamic engineers do not need to wait until the last iteration to judge that the candidate design may be uninteresting. This can be ascribed to the fact that CFD solutions generally comprise a transient stage (initiation to global approximation) followed by resolution stage (local features refined). Frequently this transient stage is perhaps only 5-10% of a total CFD convergence, but defines 90% of the flow features. Halting a CFD run that will not produce good results can spare a lot of computational time. According to personal communications with our industrial partners, in high-lift wing design, results of CFD simulations can largely be divided into several classes, among which only one or two is worth further analysis and are expected to produce good aerodynamic performance. Thus, huge amounts of computational time can be saved if we are able to predict the final pattern of the CFD simulation according to the first hundreds of CFD iterations. To this end, time series pattern classification techniques [40], in particular methods developed for early classification techniques of time series [41, 42] can be adopted. Note that here, we are interested in knowing whether the given design is worth further analysis or not, rather than predicting the exact convergence value of the CFD simulation. Both feature based and model based techniques for time series pattern classification techniques [43] will be investigated.

WP2.2 Ensemble-based convergence prediction. If a candidate design is found to be promising for further analysis, additional CFD iterations will run until it converges to accurately evaluate the quality of the design. To reduce computational time, we will use recurrent neural networks to predict the converged fitness value using the data collected from the CFD simulations. The most attractive property of this approach is that the surrogate modelling task is seen as time series prediction that is no longer subject to the high-dimensionality of the design space. Our preliminary work using a single recurrent neural network [44] and ensembles of recurrent networks [24] have shown this convergence prediction is a very promising surrogate technique for reducing computation time in fluid dynamic optimisation. In this task, we will further investigate the use of heterogeneous ensemble of recurrent neural networks, where each base model uses different inputs, for convergence prediction. Techniques for determining when the time-consuming iterative process can be stopped based on the criterion that sufficient iterative data has been collected for predicting the converged value using the ensemble will be investigated. Little is known about how to guarantee the stability of the recurrent networks for on-line prediction and how to combine base models having different model types and different inputs. We will explore optimal combinations of various recurrent models for stable and efficient prediction of the convergence process.

WP3 Active Learning. Surrogate models are often used to approximate the objective(s) to evaluate candidate solutions from crossover and mutation; alternatively, design variables optimising the surrogate itself are found (cheaply) and are used as candidates for the next expensive full evaluation. These approaches suffer from a number of problems which we address in this following work package.

WP3.1 Optimisation as sequential Markov decision process. As an optimisation proceeds early evaluations of the objectives become stale, biasing a surrogate model towards early generations, and necessitating heuristic approaches [e.g. 45] designed to "forget" old evaluations. To bring a unified, principled approach to surrogate learning in the non-stationary environment encountered during the search we plan to use active learning. We propose to cast the search as a sequential Markov decision process [46], in which the search algorithm incorporates information learned from the most recent expensive evaluation into a Bayesian model of the objective landscape on which the decision about where to next evaluate is based. Since the learner's own (possibly stochastic) decision algorithm about where next to evaluate is known, it can be accounted for in learning from that evaluation.

Our recent work [1, 2] has developed an effective framework for active learning with Bayesian learners which explicitly models the selection of the next point to be evaluated. The posterior distribution provided by the Bayesian learner allows control of diversity and the exploration/exploitation bias, because the active learner may be set to evaluate points about which it is very confident (high exploitation) or which are potentially fruitful but with low confidence (high exploration). Bayesian learning also provides a natural method for incorporating the partial, but valuable, information available from CFD simulations which have been stopped early because the predictions made in WP2 indicate that they are suboptimal. Averaging over (surrogate) model parameters to reduce model uncertainty is a further advantage of the Bayesian learning paradigm. The challenge here will be to adapt this to the active learning of surrogate objectives in the non-stationary environment generated by the evolving optimisation and, in multi-objective optimisation, to choose points that provide useful information about more than one objective.

WP3.2 Sparse Bayesian learning. The curse of dimensionality arising from the high-dimensional design space and small number of true evaluations means that the response surface(s) are prone to over-fitting. To prioritise relevant design variables we plan to use Automatic Relevance Determination (ARD) priors [e.g. 3], together with repeated resampling that has proved effective for very high-dimensional gene expression data [47, 48]. Information on the relevant variables to the optimisation is a side benefit of the ARD methodology. For both this task and the sequential Markov decision process active learning (WP3.1), we will use sparse Gaussian processes, which can be cheaply learned [49].

WP3.3 Alternative loss functions: learning to rank. The common mean-squared loss function may not be appropriate for surrogate learning because models trained using it may poorly predict good candidates [50]. We plan to investigate the use of robust loss functions based on Student-t distributions which we have shown to be effective in signal processing and tracking applications [3, 5].

A natural loss function, particularly for dominance-based multi-objective methods, is a rank-based loss function by which the surrogate learns the relative quality of the objective evaluations [45] using a pair-wise loss. We plan to augment the regression learning of objectives with simultaneous learning of rankings which has been reported to enhance regression [51], and results using ranks in uni-objective EAs as surrogate model have been promising [52].

The University of Exeter will add value to the project by fully funding a PhD student who will work on WP3.3.

WP3.4 Gradient information. Full or partial gradient of the objective function(s) with respect to decision variables can often be available in CFD calculations, perhaps by use of an adjoint formulation [e.g. 53] or via automatic differentiation.

Gradient-based optimisation has been widely shown to be highly efficient, compared to a GA, for simpler aerodynamic design problems [54]. However it is limited when solutions are close to being discontinuous, as occurs in complex designs and flows. We plan to investigate the use of gradient information for guiding EAs and active learning for efficient surrogate-based search, particularly for multi-objective optimisation.

If successful, the outcome of this WP will be a principled algorithm, not requiring the tuning of parameters, that will build surrogate objective functions and optimally choose new candidates for evaluation with control of the exploration–exploitation bias.

WP4 Finding Robust Solutions Good solutions for practical problems should also be robust in the sense that the objective values are insensitive to small perturbations in the parameter values. In addition, it is often crucial that the optimised design displays acceptable performance in "off-design" and critical conditions; for example, high aircraft cruising efficiency is useless if the aircraft behaves poorly in a stall.

WP4.1 Evolving off-design robustness. Off-design and critical operating conditions can handled by regarding performance under these conditions as one or more additional objectives. In order to avoid the computational expense of evaluating a large number of additional objectives, we will investigate the use of a single additional objective, which is the worst performance under all these conditions. Bayesian surrogate models will be used to model performance under each of the off-design conditions. This will allow confident prediction of which off-design condition is likely to be worst and is therefore the condition which should be optimised further, possibly warranting additional objective evaluations to gauge its performance. We will investigate this type of strategy on aerodynamic design, cyclone separation and in-cylinder diesel problems, all of which require robust off-design performance.

WP4.2 Evaluating robust solutions. Warm starts of the CFD code from an existing converged solution to evaluate a small perturbation of the solution are likely to converge rapidly. To characterise robustness we will model the objective function(s) in a solution's neighbourhood by predicting converged objective values from a small number of CFD iterations following warm starts. Here, unlike WP2.1, we are interested in the converged values, not just whether the solution is worth investigating further (see also WP5). Prediction of the objectives from warm starts at a number of locations will allow a complete picture of the Jacobian matrix relating changes in objectives to small changes in parameters to be cheaply constructed. Model consistency will be ensured by checking that the predicted objectives conform to a linear model; those that do not can be handled re-evaluating the convergence prediction with additional CFD iterations. We will also investigate the feasibility of this method to further characterise the robustness of a solution by cheaply constructing approximations to the relevant Hessian matrices.

WP5 Surrogate management strategies that systematically integrate various surrogates This WP will focus on developing new model management techniques that systematically integrate various surrogates models. We will seek answers to three closely related questions, i.e., how to improve convergence classification and prediction of the current design using CFD results from previous similar designs; how to enhance the quality of the surrogates using partially converged CFD data obtained in convergence prediction, and finally how to determine whether fitness prediction, rank prediction, convergence prediction, or full CFD simulations should be used for a given candidate design?

To answer the first question, i.e., to use convergence data from similar designs in previous or current generation, advanced machine learning techniques such as multi-task learning [39] will be considered. Our previous work has already shown that multi-task learning can significantly enhance learning performance when training data is lacking [18]. Recent developments in transfer learning [38], which extend the ideas in multi-task learning to semi-supervised learning, can also be considered.

Each partial convergence simulation can be seen as a low fidelity fitness evaluation of the expensive fitness evaluation (full convergence simulation). While partially converged CFD simulations can save computational time using convergence prediction, they are also potentially helpful for improving fitness prediction by introducing active semi-supervised learning [55] into co-kriging [34]. In this task, we will further explore our previous idea for using dynamic fidelity in evolutionary optimisation [35] and correlation-based performance indicators that we have recently been investigating [56].

The third question calls for a novel model management framework that can seamlessly integrate surrogate techniques by finding the best trade-off between correct convergence and minimisation of the computation cost, and between exploitation and exploration [22, 57]. Here, concerns in both learning and optimisation will be taken into account.

WP6 Solving Real-World Optimisation Problems The industrial cases on which we will prove the surrogate assisted EAs display a range of challenging characteristics typical of real-world CFD optimisation problems. All of them are time-consuming and all employ iterative solvers to which the convergence prediction and early classification proposals (WP2) can be applied. They are all high-dimensional, typically having 10-20 decision variables, although the aerodynamic design problem (WP6.1) may have tens to hundreds. Robust performance over a range of operating conditions (WP4) is required in the aerodynamic design (WP6.1), cyclone separation (WP6.2) and in-cylinder diesel (WP6.3) optimisation problems. All but the cyclone separation problem can be naturally cast in a multi-objective framework. The active learning (WP3) and surrogate management (WP5) proposals are applicable to each of them.

There will be a close collaboration with our industrial partners, including the relevant PDRAs spending several weeks over the course of the project to understand the major challenges of real-world problems and to interact with research and development engineers in industry.

Efficient and effective fluid dynamic optimisation depends also on problem representation, such as geometry representation in aerodynamic design optimisation [58]. Appropriate representations, including non-parameterised representations, e.g., nonuniform B-splines or free form deformation, will be chosen in conjunction with our industrial partners who already have operational codes for these problems. To facilitate automated evolutionary optimisation, CFD solvers, (e.g., OpenFOAM), that can automatically regenerate meshes will be adopted.

WP6.1 Aerodynamic design optimisation. Aerodynamic design optimisation is characterised by highly non-linear physics, multiple conflicting operating conditions, numerous constraints and use of computationally intensive CFD simulations. We have substantial experience in aircraft aerodynamic and multidisciplinary design optimisation with QinetiQ and Airbus [59] and for design of a turbine engine using surrogate-assisted evolutionary optimisation [20]. In this WP, we will apply and adapt the algorithms developed in WP2 for multi-element high-lift aerodynamics design. Our industrial partners have identified this highly relevant design challenge, for which existing surrogate modelling and direct optimisation approaches significantly struggle. Non-linear behaviour results from complex boundary layer/boundary layer interaction, which also leads to numerical instabilities and intermittent CFD failures.

WP6.2 Optimisation of cyclone separation. A cyclone or vortex flow separator uses the centrifugal force exerted on particles in a vortex flow to separate them from the flow (air or water). They are widely used in a variety of industrial applications from separation of particulates in drainage water to vacuum cleaners. The device geometry has a crucial effect on the properties of the system, with the geometry of the lower conical section determining the minimum size of particle which will be efficiently removed from the flow. In this WP we will develop surrogate models for cyclone separation of particles in water in conjunction with our partner Hydro International. Computation of these flows is complex, involving high resolution, turbulence modelling and particle tracking, areas in which we already have experience. This task represents a geometric variation problem with issues of modelling veracity and convergence to deal with. Hydro International has a number of vortex separators in their product line and will provide experimental data on their Swirl-Flo system for validation of the initial CFD simulations.

WP6.3 Diesel Particle Tracking. CFD modelling of diesel sprays is an important factor in developing diesel engines to satisfy environmental regulations on NOx and soot, both of which depend on the spray and fuel vapour distribution in the engine. This is an area of particular interest to our industrial partner Ricardo. Spray propagation depends on the injector design at a level of geometric detail which would be implausible to simulate, and which is included through multiparameter models, and also on the physical details of the droplet propagation, which is also described by multi-parameter models. Traditionally the values for these two parameter sets have been determined by hand by comparison with geometrically simplified off-line experiment and then included into full-cylinder simulations; the tuning process is therefore less rigorous than it could be if automated, and in particular the objective criteria are poorly defined. This therefore represents a model parameter variation problem with multiobjective criteria. Ricardo will supply experimental data and assistance with initial setup and validation of the CFD, together with assistance with development of suitable target functions for the model evaluation, using their Vectis code.

WP6.4 In-cylinder diesel optimisation. The fundamental aim of diesel engine design is to maximise power output Indicated Mean Effective Pressure whilst minimising pollutant emissions (NOx and soot). CFD simulation has provided a valuable tool in this. A subset of design parameters, related to WP6.3, is the variation of geometric parameters for the spray injector, together with timing and compression ratio, for injection and ignition in a standard diesel engine. This therefore represents a complex multidimensional, multiobjective optimisation problem. This WP and WP6.3 will be undertaken in collaboration with the IC Engines group at the Politecnico di Milano, who have world-renowned expertise in simulating diesel engines using OpenFOAM, full mesh motion, and characteristic time scale models for the combustion. The PDRA will spend 6 months working with the group in Milan on the simulation of the diesel engines.

Project management. A Steering Group, meeting every 4 months, comprising the investigators together with representatives from each of the industrial partners will be set up to oversee the direction of the work (see workplan). Everson will have overall academic investigative responsibility. To ensure coherence of the project, academic staff and PDRAs will, using Skype or similar, meet together on a fortnightly basis. We will also establish a code repository and project wiki to act as a common store for code, results and reports. There will be weekly supervisory meetings with PDRAs and the PhD student, with additional meetings as necessary. In addition to the times PDRAs spend with the industrial partners, annual joint workshops including all the industrial partners will be organised.

National Importance. The UK has the second largest aerospace industry in the world with significant capabilities in the key areas, including engines and airframe structures. To maintain UK's world-leading position in this sector, optimisation techniques like evolutionary algorithms will provide a unique opportunity to address the main challenges and create innovative designs. This research will remove the obstacles in applying evolutionary techniques to the aerospace industry. Similarly, the UK-based automotive industry is world-leading and also relies heavily on CFD.

In addition to CFD-related activities, research in optimisation has high impact across a wide range of UK scientific, commercial and public sector enterprises, which is highlighted by its underpinning of the EPSRC Challenge Themes. It is important in tackling challenges in Manufacturing the Future such as value chain optimisation, and design of efficient products, and managing sustainability across manufacturing systems. Optimisation of complex systems is important to the Energy theme in, e.g., renewable energy devices and transport networks. In Healthcare Technologies it is essential to drug design and the optimisation of diagnostic tools. EAs are applied to water distribution optimisation, addressing challenges in the Living with Environmental Change theme.

The most recent International Review of Mathematical Sciences (2010) stated that "UK researchers represent excellence in continuous, stochastic and combinatorial optimisation... in particular, UK work on hyper-heuristic methods is world-leading and has had a significant impact." This work will contribute to and sustain the UK's position in this area of the global research landscape.

References

[26] Y. Jin. "A Comprehensive Survey of Fitness Approximation in Evolutionary Computation". In: Soft Computing 9.1 (2005), pp. 3–12.

[27] Y. Jin. "Surrogate-assisted evolutionary computation: Recent advances and future challenges". In: Swarm and Evolutionary Computation 1.2 (2011), pp. 61–70.

[28] M. Sefrioui and J. Periaux. "A hierarchical genetic algorithm using multiple models for optimization". In: PPSN. 2000, pp. 879–888.

[29] J. Branke and C. Schmidt. "Faster convergence by means of fitness estimation". In: Soft Comp. 9 (2005), pp. 13–20.

[30] J Knowles. "ParEGO: A Hybrid Algorithm With On-Line Landscape Approximation for Expensive Multiobjective Optimization Problems". In: IEEE T-EVC 10 (2006), pp. 50–66.

[31] Y. Jin and B. Sendhoff. "Reducing fitness evaluations using clustering techniques and neural network ensembles". In: GECCO. 2004, p. 688.

[32] Y. Tenne and S. W. Armfield. "A framework for memetic optimization using variable global and local surrogate models". In: Soft Comp. 13 (2009), p. 81.

[33] M. Berci, P.H. Gaskell, R.W. Hewson, and V.V. Toropov. "Multifidelity metamodel building as a route to aeroelastic optimization of flexible wings". In: Proc. Inst. of Mech. Engineers, Part C: J. Mech. Eng. Science, 225 (2011), pp. 2115–2137.

[34] A.I.J. Forrester, A. Sobester, and A.J. Keane. "Multi-fidelity optimization via surrogate modelling". In: Proc. R. Soc. A 463 (2007), pp. 3251–3269.

[35] D. Lim, Y.-S. Ong, Y. Jin, and B. Sendhoff. "Evolutionary optimization with dynamic fidelity computational models". In: International Conference on Intelligent Computing. 2008, pp. 235–242.

[36] Y. Jin and J. Branke. "Evolutionary optimization in uncertain environments - A survey". In: IEEE Transactions on Evolutionary Computation 9.3 (2005), pp. 303–317.

[37] I. Paenke, J. Branke, and Y. Jin. "Efficient search for robust solutions by means of evolutionary algorithms and fitness approximation". In: IEEE Transactions on Evolutionary Computation 10.4 (2006), pp. 405–420.

[38] S. J. Pan and Q. Yang. "A survey on transfer learning". In: IEEE TKDE 22.10 (2010), pp. 1345–1359.

[39] E. Bonilla, K. Chai, and C. Williams. "Multi-task Gaussian process prediction". In: NIPS (2008), p. 153.

[40] Z. Xing, J. Pei, and E.J. Keogh. "A brief survey on sequence classification." In: SIGKDD Explorations 12.1 (2010), pp. 40–48.

[41] J.J.R. Diez and C.A. Gonzalez. "Learning Classification RBF Networks by Boosting". In: Multiple Classifier Systems (2001), pp. 43–52.

[42] Z. Xing, J. Pei, and P.S.Yu. "Early classification on time series". In: Knowl. Inf. Syst. 31.1 (2012), pp. 105–127.

[43] T.W. Liao. "Clustering of time series data - a survey". In: Pattern Recognition 38.11 (2005), pp. 1857–1874.

[44] Y. Cao, Y. Jin, M. Kowalczykiewicz, and B. Sendhoff. "Prediction of convergence dynamics of design performance using differential recurrent neural networks". In: IJCNN. 2008, pp. 529–534.

[45] I. Loshchilov, M. Schoenauer, and M. Sebag. "Self-adaptive surrogate-assisted covariance matrix adaption evolution strategy". In: GECCO'12. 2012.

[46] M. A. Osborne, R. Garnett, and S. J. Roberts. "Gaussian processes for global optimisation". In: 3rd International Conference on Learning and Intelligent Optimization (LION3). 2010, pp. 1–15.

[47] A.R.J. Clark. "Multi-objective ROC learning for classification". PhD thesis. Exeter University, 2012.

[48] Y. Li, C.Campbell, and M. Tipping. "Bayesian automatic relevance determination algorithms for classifying gene expression data". In: Bioinformatics 18 (2002), pp. 1332–1339.

[49] N. Lawrence, M. Seeger, and R. Herbrich. "Fast sparse Gaussian process methods: the Informative Vector Machine". In: NIPS. 2002, pp. 609–616.

[50] Y. Jin. "Quality measures for approximate models in evolutionary computation". In: GECCO Workshop. AAAI, 2003, pp. 170–173.

[51] D. Sculley. "Combined Regression and Ranking". In: ACM SIGKDD. 2010.

[52] I. Loshchilov, M. Schoenauer, and M. Sebag. "Comparison-Based Optimizers Need Comparison-Based Surrogates". In: PPSN. Vol. 1. 2010, pp. 364–373.

[53] M. B. Giles and N. A. Pierce. "An Introduction to the Adjoint Approximation". In: Flow, Turbulence and Combustion 65 (2000), pp. 393–415.

[54] D. Zingg, M. Nemec, and T. Pulliam. "A comparative evaluation of genetic and gradient-based algorithms applied to aerodynamic optimization". In: Revue européenne de mécanique numérique 17 (2008), pp. 103–126.

[55] A. B. Goldberg, X. Zhu, A. Furger, and J.-M. Xu. "OASIS: Online active semi-supervised leanring". In: AAAI Conf. AI. 2011, p. 362.

[56] L. Zhuang, K. Tang, and Y. Jin. "Metamodel assisted mixed-integer evolution strategies based on Kendall rank correlation coefficient". In: The 14th International Conference on Intelligent Data Engineering and Automated Learning. 2013.

[57] M. N. Le, Y. S. Ong, S. Menzel, Y. Jin, and B. Sendhoff. "Evolution by adapting surrogates". In: Evolutionary Computation 21.2 (2013), pp. 313–340.

[58] Y. Jin and B. Sendhoff. "A systems approach to evolutionary multi-objective structural optimization and beyond". In: IEEE Computational Intelligence Magazine 4.3 (2009), pp. 62–76.

[59] J.J. Doherty, S.R.H. Dean, P. Ellsmore, and A. Eldridge. "A Multi-Fidelity Approach for Supporting Early Aircraft Design Decisions". In: Collaborative Product and Service Life Cycle Management for a Sustainable World - Advanced Concurrent Engineering. Vol. 7. 2008, pp. 267–279.