CHEMICAL ENGINEERING TRANSACTIONS
VOL. 45, 2015
A publication of AIDIC
The Italian Association of Chemical Engineering
www.aidic.it/cet
Guest Editors: Petar Sabev Varbanov, Jiří Jaromír Klemeš, Sharifah Rafidah Wan Alwi, Jun Yow Yong, Xia Liu

ISBN 978-88-95608-36-5; **ISSN** 2283-9216
DOI: 10.3303/CET1545059

Methodology for "Surrogate-Assisted" Multi-Objective Optimisation (MOO) for Computationally Expensive Process Flowsheet Analysis

Ishan Sharma*[a], Andrew Hoadley[b], Sanjay M. Mahajani[c], Anuradda Ganesh[d]

[a] IITB-Monash Research Academy, Indian Institute of Technology Bombay, Powai, Mumbai, 400076, India
[b] Department of Chemical Engineering, Monash University, Clayton, VIC-3168, Australia
[c] Department of Chemical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai, 400076, India
[d] Department of Energy Science and Engineering, Indian Institute of Technology Bombay, Powai, Mumbai, 400076, India
114174001@iitb.ac.in

In recent years, Multi-Objective Optimisation (MOO) has increasingly been applied in chemical engineering problems where multiple, often conflicting, objectives need to be optimized simultaneously. Genetic Evolutionary Algorithms (EAs) are often used in order to solve MOO problems, especially when the user is concerned with the trade-offs involved between the multiple objectives. One particular form of the evolutionary algorithm is the Genetic Algorithm (GA), which is a population based algorithm that tries to mimic the process of biological natural selection. GA or any other EA is able to yield a set of non-dominated solutions in a single run. However, a downside associated with such population based algorithms is that a large fraction of the datasets evaluated in order to guide the search, are actually not represented in the final solution. This becomes an issue, especially for problems involving computationally expensive functional evaluations. For such problems, "surrogate" or "meta" models are often used to approximate the exact, but computationally expensive models. This results in a significant saving in terms of computation time. However, building a surrogate model, accurate enough in the entire decision variable space is a challenge in itself. In the present work, a methodology to perform "surrogate-assisted" MOO has been proposed. The efficiency of the proposed methodology is compared against another similar methodology for surrogate-assisted MOO. The two approaches are first tested against two mathematical test problems and the most suitable method is then applied to the chemical engineering flowsheet optimisation of a coal to ammonia process with Carbon Capture and Sequestration (CCS). The results show savings in computation time even in the most conservative case.

1. Introduction

MOO is often used to analyse the trade-offs involved between two or more, often conflicting objectives. There are numerous methods available to solve MOO problems including; objective weighing method, ε-constraint method and evolutionary algorithm based methods, among others. GAs are a class of EAs that use a population based approach to mimic the process of biological natural selection. GA or any other EA has the advantage that they are able to yield a set of non-dominated solutions in a single run. However, in order to do so, a lot of candidate solutions need to be evaluated and ranked. Hence, solving a MOO problem by using such a population based approach takes a lot of time if the functions being evaluated, are computationally intensive. Also, a large fraction of the candidate solutions evaluated during the MOO run are not included in the final solution. Hence, there exists an opportunity to save a lot of computation time if the functions could be approximated by a computationally simpler function.
A "surrogate" model refers to an approximation of the exact or "detailed" model. "Surrogate-assisted" MOO refers to the use of surrogate models, either partially or completely, during the course of a MOO run with an aim to speed up the search. However, building a surrogate model, accurate enough in the entire

Please cite this article as: Sharma I., Hoadley A., Mahajani S.M., Ganesh A., 2015, Methodology for "surrogate-assisted" multi-objective optimisation (moo) for computationally expensive process flowsheet analysis, Chemical Engineering Transactions, 45, 349-354 DOI:10.3303/CET1545059

decision variable space is a challenge in itself. Wilson et al. (2001) used such an approach of using a single global surrogate to carry out MOO. Jin (2005) discusses this approach in detail and concludes that this approach can be problematic in cases suffering from high dimensionality, ill distribution and limited number of training datasets. It is for this reason that Jin (2005) recommends the use of both the "detailed" and the "surrogate" model during the MOO run.

The next issue to be dealt with is how to decide whether to use the surrogate or the detailed model for an evaluation. Jin (2005) described this as "model management" or "Evolution Control" (EC). There are two basic strategies used in EC namely, fixed EC and adaptive EC. The salient feature of fixed EC is that the frequency of detailed model evaluations is fixed. Fixed EC includes two sub-categories viz., individual and generation based EC. Individual based EC involves only some individuals in every generation using the detailed models. Generation based EC, on the other hand, involves the candidate solutions being evaluated by detailed model only after a fixed interval of evolutionary generations. Adaptive EC involves a variable, rather than a fixed, frequency of detailed model evaluations. This frequency depends on the accuracy of the surrogate models. Ray et al. (2009) proposed a Surrogate Assisted Evolutionary Algorithm (SAEA) which used a mixture of the basic EC strategies discussed above. In SAEA, in addition to generation based fixed EC, a form of adaptive individual based EC is also used. In SAEA, the Euclidean distance of every normalized candidate solution is measured from each and every normalized training dataset. If the minimum distance is below a particular threshold value (confidence region), then only the surrogate model is used for function evaluation. SAEA involves periodic retraining of the surrogates by making use of all the detailed model evaluations stored in an "external archive". The periodic re-training of the surrogate model is advantageous especially in the cases where the dataset, used to initially fit the surrogate, is significantly different from the final Pareto-optimal solution.

There is a wide variety of model approximations available to choose from, ranging from polynomial models, multi-layered feed-forward neural networks, kriging approximations, radial basis functions, and support vector machines. Santos et al. (2013), for example, compared the performance of multi-layered feed-forward neural networks and radial basis functions in detecting and locating leaks in gas transportation pipelines. Jin (2005) reports that though there are several studies comparing the performance of different approximation methods, there is a lack of clarity as far as the advantages and disadvantages of different methods is concerned. This is mainly because performance of individual methods is also dependent on the problem being solved. Also, there may be several criteria on which the comparison could be made.

Isaacs et al. (2009) recently proposed another set of algorithms for surrogate-assisted MOO. The two algorithms proposed by Isaacs et al. (2009) are Multiple Spatially Distributed Surrogates (MSDS) and Multiple Adaptive Spatially Distributed Surrogates (MASDS). There are a lot of similarities between SAEA and these two algorithms. For example, like SAEA, MSDS and MASDS also use EC. The basic difference between the two approaches is that while SAEA uses single surrogate models in order to predict the functions in the entire decision variables space, MSDS and MASDS use multiple, locally accurate surrogate models in order to predict functions in different regions. The basic genetic algorithm used in MSDS and MASDS is the Non-dominated Sorting Genetic Algorithm (NSGA)-II proposed by Deb et al. (2002). Isaacs et al. (2009) demonstrated these algorithms for six mathematical MOO problems along with a welded beam design optimisation problem. Isaacs et al. (2009) concluded that MASDS performed by far the best for most of the problems. Recently, Zhu et al. (2014) proposed a multi-objective, variable-fidelity optimization method, i.e. a surrogate-assisted method, for genetic algorithms. They noted that the optimal decision variable space is typically a small subset of the original search. Hence, in their approach new points are only included for surrogate model fitting if they are expected to impact the domination decisions made by the genetic algorithm and on the basis of error estimates, provided by the kriging models. However, the computation complexity in fitting the surrogate model increases significantly with an increase in the size of the dataset used to fit the model. This problem is especially acute for problems involving three or more objectives. Liu and Collette (2014) proposed an improved version of the method proposed by Zhu et al. (2014) to incorporate multiple surrogate fitting, just like what Isaacs et al. (2009) had proposed in MSDS and MASDS, to deal with the increase in complexity of surrogate fitting. A complimentary way to deal with this issue could have been to limit the search space periodically and fitting the surrogate in just the promising subset of the original decision variable space. This is particularly important for many real life problems where the initial search space could be drastically different than the final solution. The present work aims to address this research gap, not considered in the above mentioned works, by proposing an appropriate surrogate-assisted MOO framework. The proposed framework has been compared against MASDS for two mathematical test problems. This is then followed by applying the proposed algorithm to a chemical engineering flowsheet optimisation of coal to ammonia process with Carbon Capture and Sequestration (CCS) by a non-selective Rectisol^{TM} process.

2. Proposed methodology

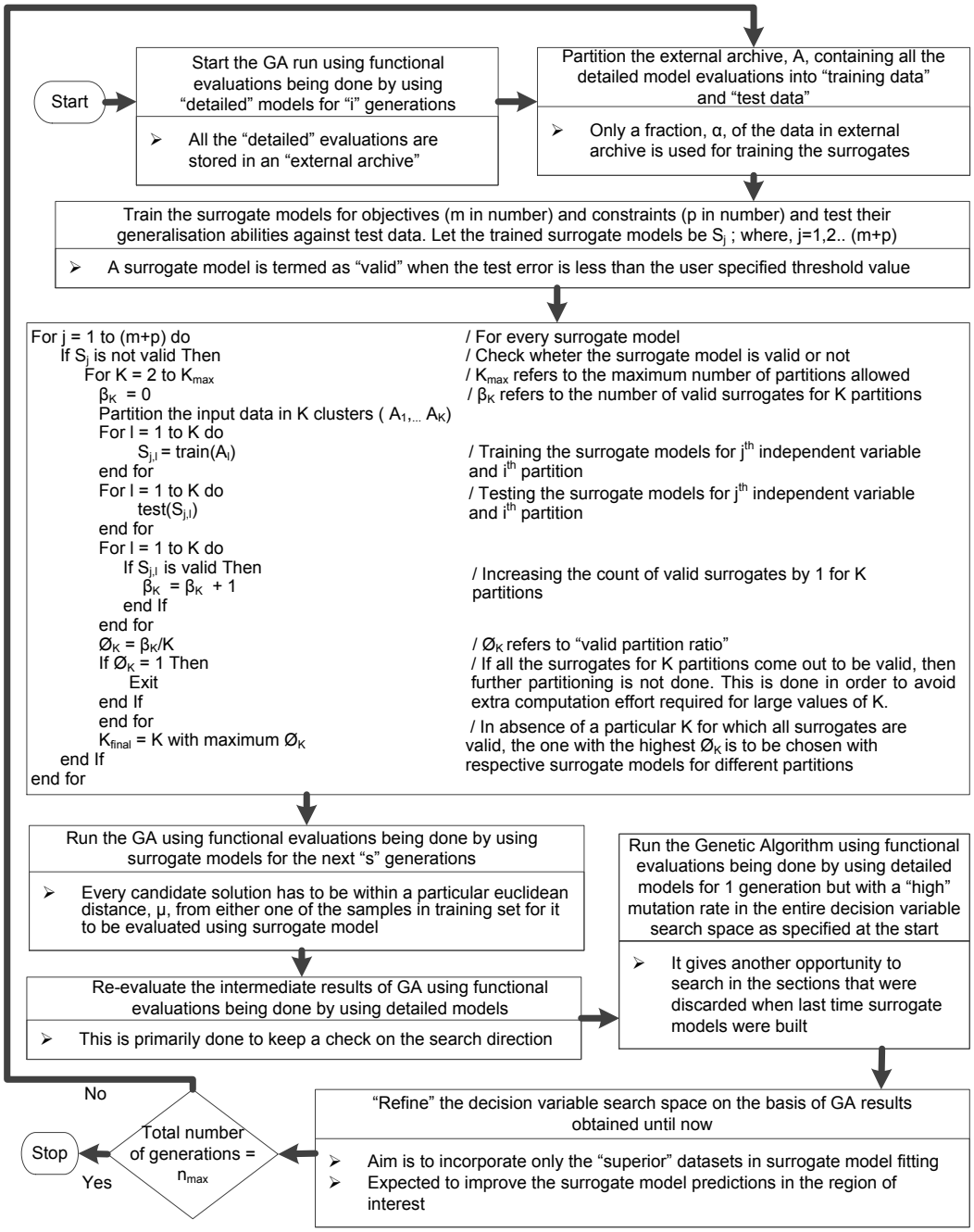

Figure 1: The proposed methodology for "surrogate assisted" MOO

The most significant differences between MASDS and the proposed methodology is that whenever the surrogate models are retrained after the initial training, the decision variable search space is "truncated" according to the final results (DNA) obtained until that point. The idea of periodic truncation helps to discard "inferior" datasets in the external archive and prevent them from contributing to the surrogate models fitting. It is expected that the accuracy of the surrogate models thus fitted, would be better in the "promising" regions.

3. Test Problems

3.1 Problem definition

Table 1 lists the two numerical test problems against which the performance of the proposed methodology and MASDS are compared. The SCH problem represents a case where the initial search space is drastically different from the optimal decision variable space. To facilitate comparison between the two algorithms crossover probability (0.85), default mutation probability (0.01), random number generator and the seed value for random number generation were kept the same for the two runs. The resolution parameter, that is μ, has been kept as constant, that is, 5 % of the length of the solid diagonal of the decision variable space. However, it could also have been variable as the search space, for the algorithm proposed in this work, keeps on changing. The value of μ should largely depend upon the accuracy of the surrogate models being fitted. The surrogate models in this work are fitted by using SUMO toolbox developed by Gorissen et al. (2010). The model being fitted are feed-forward neural networks where MATLAB's gads toolbox selects network parameters using a GA.

Table 1: Numerical test problems used in this study

Problem	Objectives	Initial search space	Final Solution	Parameters
SCH problem (Convex Pareto)	Minimise: $f_1(x) = x^2$ Minimise: $f_2(x) = (x-2)^2$	$x \in [-1000, 1000]$	$x \in [0,2]$	$i = 1$ $s = 3$ $n_{max} = 13$
ZDT2 Problem (Concave Pareto)	Minimise: $f_1(X) = x_1$ Minimise: $f_2(X) = g(X) * h(f_1(X), g(X))$ $h(f_1, g) = \left[1 - (f_1/g)^2\right]$, $g(X) = \left(1 + \sum_{i=2}^{10} x_i\right)$	$x_i \in [0,1]$, $i = 1,2,3,...10$	$x_1 \in [0,1]$ $x_i = 0$, $i = 2,3,4,....10$	$i = 1$ $s = 7$ $n_{max} = 56$

3.2 Results

Figure 2(a) and 2(b) show the intermediate results obtained for SCH and ZDT2 test problems after 13 and 56 generations. For both the test problems, it can be seen that the proposed framework gave better results in terms of both nearness to actual Pareto and the spread of points on the Pareto. However, for ZDT2 problem, the proposed framework required a greater number of detailed model evaluations. The resolution parameter, that is μ, was kept constant as 5 % of the normalised solid diagonal length. Hence, the resolution (in absolute values) became smaller, as the search space became narrower. This meant that the minimum distance criterion for evaluation via surrogate models was not met in most of the cases. It is hence suggested that in such cases, μ could also be treated as variable, depending upon the accuracy of the fitted surrogate models.

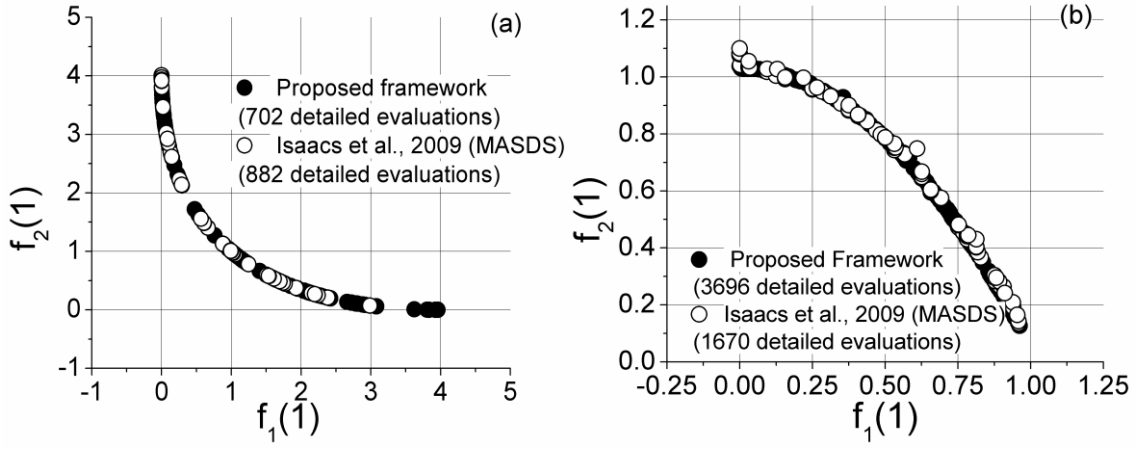

Figure 2: (a) SCH results after 13 generations; (b) ZDT2 results after 56 generations

Since, the proposed framework showed promising results, it was then applied to "global" optimisation of coal to ammonia process.

4. Rectisol™ Problem

Figure 3 depicts the coal to ammonia process (with CO_2 capture) under investigation. The gasifier is an entrained flow gasifier and chilled methanol is being used to capture the CO_2. The aim of this exercise is to optimise the "global" coal to ammonia process while using a computationally cheaper surrogate model for the CO_2 capture process (namely, Rectisol™). The detailed model for Rectisol™ involves a number of recycle streams, thereby making the Aspen Plus™ simulation, computationally expensive.

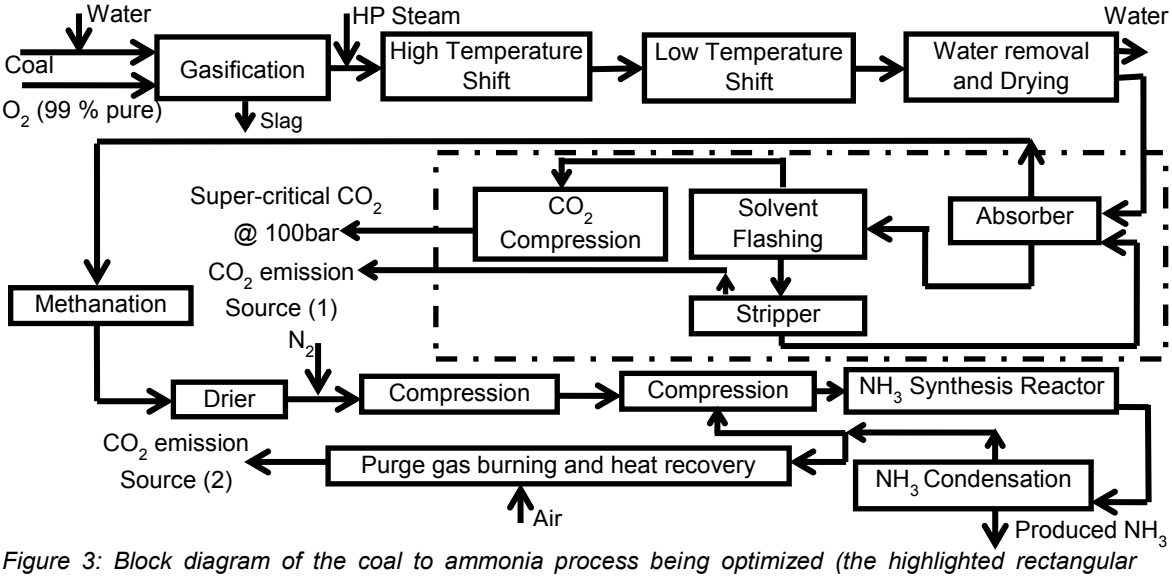

Figure 3: Block diagram of the coal to ammonia process being optimized (the highlighted rectangular portion represents the surrogate model's boundary)

4.1 Problem definition

Objectives: Minimise net electricity consumption (kW) and moles of CO_2 emitted per unit moles of NH_3 produced (for a minimum NH_3 production of 1,200 kmol/h) and a fixed coal feed rate.

Decision variables (seven in total): Oxygen flow to gasifier (kg/h), HP steam to shift reactor (kg/h), Purge fraction for ammonia synthesis loop, along with the four parameters specific to Rectisol™ unit- same as specified in (Sharma et al., 2014).

Constraints: H_2 and CO recovery across Rectisol™ unit should both be greater than 99 % and CO_2 mole fraction in the treated gas from Rectisol™ unit should be less than 10 PPM.

4.2 Results

Figure 4(a) shows the results obtained, for similar number of detailed model evaluations, with "business-as-usual approach" (BAU) (that is, where only detailed Aspen Plus™ evaluations were done) and the surrogate-assisted approach, proposed in the present work.

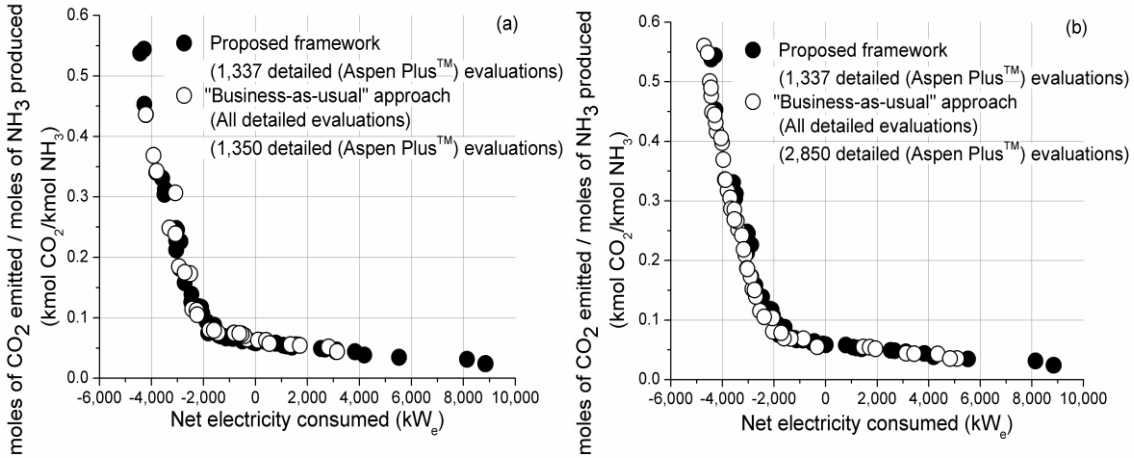

Figure 4: (a) Proposed framework and BAU approach (for similar number of detailed evaluations); (b) Proposed framework and BAU approach after 1,337 and 2,850 detailed evaluations

The proposed framework's performance is somewhat similar (but better in terms of spread of solutions) to the BAU. Pairwise comparison of the two Pareto datasets yielded that every point on the proposed framework plot was dominated by 0.50 points on the BAU plot. While every point on the BAU plot was dominated by 1.86 points on the proposed framework plot. Figure 4(b) shows the comparison of the results obtained from the proposed framework and those obtained from the BAU approach for a significantly higher number of detailed evaluations. The comparison depicts how close these results are to global optimum values, assumed to be depicted by the BAU plot in Figure 4(b). Such an analysis clearly shows the relative advantage offered by the proposed approach over the BAU approach, even for a conservative value of the resolution parameter (μ), kept as 5 % of the normalised solid diagonal length.

5. Conclusions

A surrogate-assisted MOO framework has been proposed in this work. The aim has been to improve upon a similar existing framework (MASDS). The proposed framework has been tested on two mathematical problems. This analysis has yielded some useful insights related to how to best apply this framework. The resolution parameter, μ, in particular could be varied depending upon the accuracy of the fitted surrogate models. This framework is then used for the global optimisation of coal to ammonia process, with CCS. The results of the optimisation run have been reported along with those obtained with the BAU approach. Each point on the non-dominated set generated with the proposed framework dominated, on an average, 1.86 points in the set generated by BAU approach. While the same metric for BAU approach was just 0.50. The results show huge potential of using such a technique for MOO problems in chemical engineering which would otherwise be computationally prohibitive The proposed framework, hence, could allow for better integration of computationally complex units into a large-scale plant simulation (i.e. global optimisation), rather than treat them as stand-alone units which could result in sub-optimal operation. Unlike other optimisation methods, where the solution is a set of optimised decision variables, this method provides both the optimised decision variables and an array of surrogate models accurate in different regions of the decision variable space, for future predictions.

Acknowledgement

The authors would like to thank Orica Ltd for funding the project through the IITB-Monash research academy.

References

Deb K., Pratap A., Agarwal S., Meyarivan T.A.M.T., 2002, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Transactions on Evolutionary Computation, 6(2), 182–197.

Gorissen D., Couckuyt I., Demeester P., Dhaene T., 2010, A Surrogate Modeling and Adaptive Sampling Toolbox for Computer Design, Journal of Machine Learning Research, 11, 2051-2055.

Isaacs A., Ray T., Smith W., 2009, Multi-objective design optimisation using multiple adaptive spatially distributed surrogates, Int. J. Product Development, 9 (1-2-3), 188-217.

Jin Y., 2005, A comprehensive survey of fitness approximation in evolutionary computation, Soft Computing, 9(1), 3-12.

Liu Y., Collette M., 2014, Improving surrogate-assisted variable fidelity multi-objective optimisation using a clustering algorithm, Applied Soft Computing, 24, 482-493.

Ray T., Isaacs A., Smith W., 2009, Surrogate Assisted Evolutionary Algorithm for Multi-Objective Optimization, G.P. Rangaiah (Ed.), Multi-objective Optimization: Techniques and Applications in Chemical Engineering, Singapore: World Scientific, 131-151.

Santos R.B., Rupp M., Bonzi S. J., Fileti A.M.F., 2013, Comparison between Multilayer Feedforward Neural Networks and a Radial Basis Function Network to Detect and Locate Leaks in Pipelines Transporting Gases, Chemical Engineering Transactions, 32, 1375-1380.

Sharma I., Hoadley A., Mahajani S.M., Ganesh A., 2014, Automated optimisation of multi stage refrigeration systems within a multi-objective optimisation framework, Chemical Engineering Transactions, 39, 25-30.

Wilson B., Cappelleri D., Simpson T.W., Frecker M., 2001, Efficient Pareto frontier exploration using surrogate approximations, Optimization and Engineering, 2(1), 31-50.

Zhu J., Wang Y.J., Collette M., 2014, A multi-objective variable-fidelity optimisation method for genetic algorithms, Engineering Optimisation, 46(4), 521-542.