Chapter 35
A Proposal for Parameter-Free Surrogate Building Algorithm Using Artificial Neural Networks

Srinivas Soumitri Miriyala
Indian Institute of Technology Hyderabad, India

Kishalay Mitra
Indian Institute of Technology Hyderabad, India

ABSTRACT

Surrogate models, capable of emulating the robust first principle based models, facilitate the online implementation of computationally expensive industrial process optimization. However, the heuristic estimation of parameters governing the surrogate building often renders them erroneous or under-trained. Current work aims at presenting a novel parameter free surrogate building approach, specifically focusing on Artificial Neural Networks. The proposed algorithm implements Sobol sampling plan and intelligently designs the configuration of network with simultaneous estimation of optimal transfer function and training sample size to prevent overfitting and enabling maximum prediction accuracy. A novel Sample Size Determination algorithm based on a potential concept of hypercube sampling technique adds to the speed of surrogate building algorithm, thereby assuring faster convergence. Surrogates models for a highly nonlinear industrial sintering process constructed using the novel algorithm resulted in 7 times faster optimization.

INTRODUCTION

Optimization techniques are frequently applied in chemical process and manufacturing industries, business, economics, health-care, finance and energy management (Ries, Beullens & Wang, 2012; Vasant, 2014; Dostal, 2014; Yuce & Mastrocinque, 2016; Lechuga, Martinez & Ramirez, 2016; Vasant, 2011; Pombo, Garcia, Bousson, & Felizardo 2017). Although the advent of high performance computers has empowered the industries by tremendously increasing the speed of processing, control and optimization

DOI: 10.4018/978-1-6684-2408-7.ch035

of industrial problems still remain to be computationally intensive (Mitra, Majumdar & Raha, 2004). The genesis of large computational times lies with the simulation of complex first principle based model for generating the candidates for optimization (Miriyala, Mittal, Majumdar & Mitra, 2016). The first principle based models, such as, those trying to capture the dynamics of reaction networks in a polymer industry or a model handling the wake effects or turbulence in fluid flow, etc. usually involve several highly nonlinear coupled Ordinary and Partial Differential Equations (ODEs & PDEs) (Khan, Hussain, Mansourpour, Mostoufi, Ghasem & Abdullah, 2014; Yousefi & Karimi, 2013; Mogilicharla, Chugh, Majumdar, & Mitra, 2014). This necessitates the involvement of robust simulation packages, such as, ASPEN, Computational Fluid Dynamics (CFD), or Differential Algebraic solvers etc., to solve the system of ODEs and PDEs to facilitate their implementation at industrial scale (Azargoshasb, Mousavi, Amani, Jafari, & Nosrati, 2015; Douguet 2010; Espinet, Shoemaker & Doughty, 2013; Jin & Sendhoff, 2009). The intrinsic complexity of these models forms the genesis for the large computational time consumed by the optimizer, compelling the entire process to run over several weeks or even months (Mogilicharla, Mittal, Majumdar, & Mitra, 2014). The problem grows by multiple folds when the considered system is multi-dimensional in nature (say m dimensions) with optimization formulation involving multiple conflicting objective functions instead of one (Miriyala et al., 2016). The conflicting nature of the objective functions results in a set of non-dominating solutions called Pareto Optimal (PO) solutions (Deb, Sindhya & Hakanen, 2016). The selection of single solution from the PO set is through some higher order information, often provided by the decision maker (Deb, 2001). The solution obtained in this way aims at enabling a decision support system to program and simulate the given process in an optimum fashion. This concept of online optimization is practically imbibed in industry when the combined functioning of optimizer and controller is realized in real time of the live process.

The tremendous industrial growth and ever-expanding demand over the last decade have created strong need for the solutions, which could cater multiple objectives at the same time. This requires solving the underlying multi-objective optimization problem (MOOP) (Deb, Agrawal, Pratap & Meyarivan, 2000). Until date, owing to the advent of fast computing machines, the ability of modern evolutionary methods for solving the MOOP has remained unparalleled (Deb, 2002). On the other hand, due to the predominant condition, wherein absence or expensive computation of gradient information of the complex models has become a common scenario, the modern evolutionary optimization techniques have gained enormous prominence over their classical counterparts, which provide every future course of movements depending on the current gradient information (Deb, 2001). The procedure of solving the MOOP by the robust evolutionary techniques, which primarily work with population of candidate solutions, necessitates multiple function evaluations in order to generate those solutions required in optimization process (Nain & Deb, 2002). These aspects together make the concept of online optimization a far-fetched impractical concept confined to theory, which cannot be realized practically unless the optimization happens in real time (Miriyala et al. 2016). The key to this problem lies with fast and accurate surrogate models, which essentially are data based models trying to emulate the given complex first principle or physics based models (Jin 2011; Tabatabaei, Hakanen, Hartikainen, Miettinen, & Sindhya, 2015; Assefi, Ghaedi, Ansari, Habibi, & Momeni, 2014). These surrogates then replace the original physics based models in the optimization algorithm thereby shielding them from the optimizer while generating the candidate solutions. With surrogates in place, the entire optimization algorithm may proceed in a fast manner thus enabling a step towards online optimization (Miryala, Pantula, Majumdar & Mitra, 2016).

However, the surrogate building approaches in general contain certain parameters such as the optimal number of training sample points, which needs to be estimated a priori. The existing methodologies (as