SLAC-R-948

Phase I Final Report

Optimal Model-Based Fault Estimation and Correction for Particle Accelerators and Industrial Plants Using Combined Support Vector Machines and First Principles Models

DE-FG02-07ER828

Dr. Bijan Sayyar-Rodsari, PI
Dr. Carl Schweiger

Pavilion Technologies, Inc., Austin, TX 78758

June 30, 2008

Work supported in part by US Department of Energy contract DE-AC02-76SF00515.

SLAC National Accelerator Laboratory, Menlo Park, CA 94025

Summary

Timely estimation of deviations from optimal performance in complex systems and the ability to identify corrective measures in response to the estimated parameter deviations has been the subject of extensive research over the past four decades. The implications in terms of lost revenue from costly industrial processes, operation of large-scale public works projects and the volume of the published literature on this topic clearly indicates the significance of the problem. Applications range from manufacturing industries (integrated circuits, automotive, *etc.*), to large-scale chemical plants, pharmaceutical production, power distribution grids, and avionics.

In this project we investigated a new framework for building parsimonious models that are suited for diagnosis and fault estimation of complex technical systems. We used Support Vector Machines (SVMs) to model potentially time-varying parameters of a First-Principles (FP) description of the process. The combined SVM & FP model was built (*i.e.* model parameters were trained) using constrained optimization techniques. We used the trained models to estimate faults affecting simulated beam lifetime. In the case where a large number of process inputs are required for model-based fault estimation, the proposed framework performs an optimal nonlinear principal component analysis of the large-scale input space, and creates a lower dimension *feature space* in which fault estimation results can be effectively presented to the operation personnel.

To fulfill the main technical objectives of the Phase I research, our Phase I efforts have focused on:

1. *SVM Training in a Combined Model Structure*: We developed the software for the constrained training of the SVMs in a combined model structure, and successfully modeled the parameters of a first-principles model for beam lifetime with support vectors.
2. *Higher-order Fidelity of the Combined Model*: We used constrained training to ensure that the output of the SVM (*i.e.* the parameters of the beam lifetime model) are physically meaningful.
3. *Numerical Efficiency of the Training*: We investigated the numerical efficiency of the SVM training. More specifically, for the primal formulation of the training, we have developed a problem formulation that avoids the linear increase in the number of the constraints as a function of the number of data points.
4. *Flexibility of Software Architecture*: The software framework for the training of the support vector machines was designed to enable experimentation with different solvers. We experimented with two commonly used nonlinear solvers for our simulations.

The primary application of interest for this project has been the sustained optimal operation of particle accelerators at the Stanford Linear Accelerator Center (SLAC). Particle storage rings are used for a variety of applications ranging from 'colliding beam' systems for high-energy physics research to highly collimated x-ray generators for synchrotron radiation science. Linear accelerators are also used for collider research such as International Linear Collider (ILC), as well as for free electron lasers, such as the Linear Coherent Light Source (LCLS) at SLAC. One common theme in the operation of storage rings and linear accelerators is the need to precisely control the particle beams over long periods of time with minimum beam loss and stable, yet challenging, beam parameters.

We strongly believe that beyond applications in particle accelerators, the high fidelity and cost benefits of a combined model-based fault estimation/correction system will attract customers from a wide variety of commercial and scientific industries. Even though the acquisition of Pavilion Technologies, Inc. by Rockwell Automation Inc. in 2007 has altered the small business status of the Pavilion and it no longer qualifies for a Phase II funding, our findings in the course of the Phase I research have convinced us that further research will render a workable model-based fault estimation and correction for particle accelerators and industrial plants feasible.

<u>Table of Contents</u>

1 - Introduction

In this section we briefly introduce this project's hybrid framework for building parsimonious models that are suited for diagnosis and fault estimation of complex technical systems. Specific applications to the control, operation and diagnosis of large particle accelerators are described. This hybrid framework uses Support Vector Machines (SVMs) to model potentially time-varying parameters of a First-Principles (FP) description of the process (Fig. 1). Unlike prior approaches to modelling, the combined SVM & FP model may be built (*i.e.* model parameters trained) with extremely efficient optimization techniques. The trained models are suitable for online/offline diagnosis and fault estimation in realistic applications with large number of input and output parameters. The proposed framework allows for optimal nonlinear mapping of large input spaces into lower-dimension *feature spaces*, and is critically important to accurate estimate of the fault, specially with highly noisy measurements.

1.1 - Fault Estimation and Diagnosis for Fail-Critical Applications - Main Challenges

Timely estimation of deviations from optimal performance in complex systems and the ability to identify corrective measures in response to the estimated parameter deviations has been the subject of significant research effort over the past four decades. The implications in terms of lost revenue from costly industrial processes, operation of large-scale public works projects and the volume of the published literature on this topic clearly indicates the significance of the problem. Applications range from manufacturing industries (integrated circuits, automotive, *etc.*), to chemical plants, pharmaceutical production, power distribution grids, and avionics.

For this proposal, the primary application for demonstration of the technique is the sustained optimal operation of particle accelerators at the Stanford Linear Accelerator Center (SLAC). One common theme in the operation of particle storage rings[1] and linear accelerators[2] is the need to precisely control the particle beams over *long* periods of time with minimum beam loss and stable, yet challenging, beam parameters. In either case, however, effects caused by component drift/failure, and disturbances (such as temperature and ground motion) tend to detune operating conditions away from optimum values. Timely detection of the onset of this detuning (*i.e.* fault), accurate estimate of the probable sources of the fault, and hands-off determination of appropriate corrective measures (within operational constraints) via a combined SVM & FP model that is suitable for both local and global analysis of the fault is the key objective of this project[3].

Despite the well-recognized importance of early fault detection (*i.e.* detection of deviations from optimal tuning), and the need for timely identification and correction of faults for sustained plant operation, a hands-off solution to the problem has yet to emerge. Clearly there is and will be increased need for commercial software for online model-predictive fault estimation in critical applications where gradual deviations from optimal operating conditions lead to plant malfunction or even shutdown. In our view, the following challenges have hindered the development of such software:

1. Parsimonious models of the process that are suitable for real time optimization, prediction, and control have been difficult to build for complex large-scale nonlinear systems.
2. Fast optimization algorithms have not been available for online fault estimation in large-scale systems with fast dynamics where the solver must converge in a fraction of a second for online fault estimation to be feasible.

[1] Particle storage rings are used for a variety of applications ranging from "colliding beam" systems for high-energy physics research to highly collimated x-ray generators for synchrotron radiation science.

[2] Linear accelerators are used for collider research such as International Linear Collider (ILC), as well as for free electron lasers, such as the Linear Coherent Light Source (LCLS).

[3] We recognize that while abundant real-time data is available for *online fault estimation* in process industries, for large particle accelerators this may not always be the case.

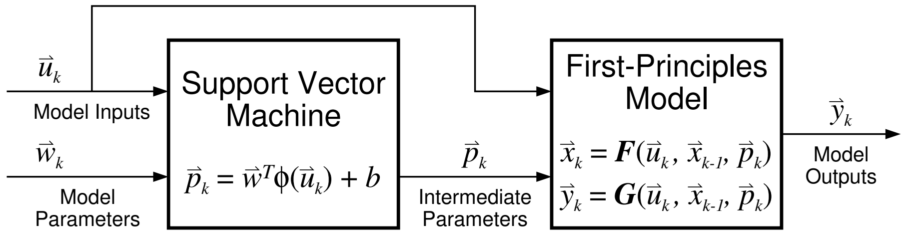

Fig. 1: Block Diagram of the Combined SVM and FP blocks that will be used to model nonlinear beam matrix and estimate the onset of a fault.

Pavilion Technologies proposed to investigate the systematic ways to overcome these challenges. The techniques for doing so and the mathematical foundation for these techniques are briefly described in this report. This report also provides the results from our simulation studies in Phase I that will clearly demonstrate the feasibility of the approach.

The primary target application was the modelling of beam-response measurements for electron storage rings at SLAC beyond linear range for model-based fault estimation and correction. More specifically we proposed the following:

1. Develop a series combination of SVM & FP models as a framework to accurately represent large scale nonlinear processes in particle accelerators and industry. The goal was to show that:
   (a) With a combined SVM & FP framework, process data and fundamental process knowledge are utilized in an *optimal* manner[4]. Objective functions and constraints allow the user to define the optimality criterion in a transparent way.
   (b) SVM formulation enables optimal nonlinear mapping to a lower-dimension feature space, optimally removing redundancy in measured data through a nonlinear principal component analysis of the process data. This feature will render our proposed approach applicable for online fault estimation in large-scale systems (in addition to offline applicability).
   (c) Expert's knowledge about fault conditions can be systematically utilized (*e.g.* as constraints for fault estimation problem). Transparent inclusion of this knowledge simplifies maintenance and modification of the knowledge over the lifetime of the application.
2. Implement efficient optimization algorithms to estimate changes in system model parameters given deviations from the normal/desired system response. A priori knowledge about the sources of deviations from the nominal condition (in the form of probability distribution functions) can be systematically included in the optimization problem and the optimization problem will remain convex.
3. Advance the state of the art in support vector machines by producing original insight into the requirements for efficient training of the support vectors when their output(s) are not directly measured.

In this project, SLAC worked with Pavilion to investigate the use of a combined SVM-nonlinear model for the beam life time so that deterioration of beam quality could be detected and automatically corrected (within operational constraints) more effectively than currently performed.

[4]In the case of application at SLAC, The combined SVM & FP model accurately captures (linear and nonlinear) variations in the beam response matrix over various operating regimes of the storage ring.

1.2 - Fault Estimation and Diagnosis for Particle Storage Rings and Linear Accelerators

The primary application of interest at host laboratory SLAC is sustained operation of particle storage rings and linear accelerators at optimum performance levels. Particle storage rings are used for a variety of applications ranging from "colliding beam" systems for high-energy physics research to generation of high power, collimated x-ray beams for synchrotron radiation science. Similarly, linear accelerators are used for collider research such as International Linear Collider (ILC) as well as for short pulse, high power free electron lasers (FEL), such as the Linear Coherent Light Source (LCLS). One common theme in the operation of storage rings and linear accelerators is the need for precision control of the particle beams over long periods of time with minimum beam loss and stable beam parameters.

Up to now, when the systems operate normally, the task of maintaining quality beam delivery can be achieved for relatively short periods of time in a linear accelerator and for longer periods of time in storage rings. In either case, effects caused by component drift and other factors (temperature, ground motion, *etc*) tend to detune operating conditions away from optimum values. As a consequence, the on-going correction of key parameters such as optical functions, beam emittance and coupling requires intervention of experts to resolve discrepancies from the optimum values.

As described below, Response Matrix Analysis (RMA) techniques are used to compare measured data with the accelerator model in order to diagnose drift or faults among a large array of system parameters. When the optical functions deviate from optimum in a storage ring light source, for instance, beam brightness and in some cases beam lifetime are compromised. The effect is more damaging in a free-electron laser where the beam emittance and energy spread must fall within stringent tolerances to fulfill the lasing criteria. Given the high costs associated with operating a multi-GeV linear accelerator and maintaining complicated experimental apparatus, time lost to "tuning" accelerator for requisite beam conditions can erode the laboratory operating budget and undermine scientific production at the research facility.

1.2.1 - The Current SLAC Solution and its Limitations: Typical techniques used to diagnose accelerator operation involve Response Matrix Analysis (RMA) and related error-finding software. The basic principle behind RMA is to acquire large data sets containing the *response* of the beam position as it is scanned through the spatially-varying magnetic field structure of the accelerator magnets. Simultaneously, a simulated response matrix is "computed" using the online accelerator model. The analysis proceeds by varying parameters within the model to make the computed "model" response matrix agree with the measured response matrix. As the agreement between model and measured data converges, the model yields a progressively more accurate representation of the physical accelerator hardware.

For storage ring systems, the response matrix data often comes in the form of "closed orbit" perturbation measurements whereas for linear accelerators and transport lines the response matrix data comes in the form of "betatron oscillation" measurements. Closed orbit analysis can be thought of as a "global" accelerator diagnostic, although the same data can be used to "locally" diagnose specific components in the storage ring such as the beam-collision area or an insertion device. Analysis of betatron oscillation data from a linear accelerator is often used to diagnose more localized regions of the system. In this case, the analysis can be complicated by variation of beam energy along the accelerating system, but precise knowledge and control of the beam energy profile is critical to producing the beam parameters required for well-controlled experiments.

It is important to note that each element of the response matrix data set is a non-linear function of the accelerator system parameters. As such, it is common to assume local linearity of the system about the present operating point, Taylor-expand the model to first order and arrange the resulting system of simultaneous equations into matrix format for analysis using linear algebra. Singular Value Decomposition (SVD) has been the method of choice for RMA because it provides a natural basis upon which the large-scale least-squares problems can be "filtered" to remove systematic effects or reduce the impact of noise. System non-linearity causes undesirable mode mixing and consequently modelling imperfections. Typically non-linearities (e.g. sextupole fields or BPM 'pincushion' effects in a storage ring) are accounted for by repeated application of the linear algorithm with fresh data and re-expansion of the first-order Taylor series model required at each step.

A more desirable approach would be direct application of modelling techniques taking non-linearities directly into account. The payoff clearly becomes increasingly important as the non-linear effects increase. A case in point is the photo-cathode gun and associated transport line used at the injector for the SLAC LCLS free-electron laser or the development of an accurate second-order model for example the SPEAR3 light source. Industrial applications in chemical processing plants or power distribution would clearly benefit from developments in the large-scale accelerator systems.

1.2.2 - New Technical Approach to Address the Current SLAC Problem: Our proposed software system seeks to use response-matrix data in combination with the SVM/First-Principles model approach to develop non-linear models that identify sources of accelerator detuning and suggest corrective action such as change in power supply currents and RF system parameters. The flexible model-based system will permit efficient non-linear control of global beam parameters (such as beam size, beam energy, tunes, chromaticity, *etc.*) by solving a carefully constructed optimization problem.

Specifically, using large modern computers available today, the operating conditions of the accelerator are systematically recorded with time-stamped "state of the machine" data. With the large amount of information stored it is feasible to identify and correct the cause of operational errors using data-mining techniques such as the Support Vector Machines (SVM) we propose to study. Direct application of non-linear techniques would expedite the presently linear RMA process to enhance machine control and consequently scientific production. Many cases can be resolved without direct intervention by experts.

We propose using both closed orbit (static) response matrix data and turn-turn (storage ring) or betatron oscillation (linac) data in future studies since it is often possible to gather data during normal operations whereas the nominal response matrix data takes dedicated time.

Data from optical diagnostics can also be used in the optimization process. For storage rings this includes on-line pinhole camera measurements of emittance and gated-iccd or streak camera measurements of transverse and longitudinal beam stability. One example of a real pay-off is in the linac application using OTR measurements, wire scanners and/or streak camera measurements to monitor 6-D phase space with optimization directed at delivering the optimum FEL or collider bunch structure.

2 - Background Information

Complex systems running under sophisticated closed loop control strategies are now common in all aspects of modern life from manufacturing and process industries, to aircraft and air travel control, to communication and power networks, to particle accelerators. While these controllers are designed to be robust to many types of disturbances, there are some changes in the system, known as faults, that the controller can not handle without changes to its structure, if at all. Such changes include system parameter changes (*e.g.* heat exchange fouling), disturbance parameter changes (*e.g.* extreme ambient temperature changes), actuator changes (*e.g.* sticking valve or damaged corrector magnet), and sensor changes (*e.g.* biased measurements).

To ensure that the system operations satisfy performance specifications, *monitoring systems* are used to detect, diagnose, and ideally remove the fault or at least recommend the remedial action to the operation personnel [1]. An effective monitoring system is expected to assist operation and maintenance personnel to make appropriate remedial actions to remove abnormal behavior resulting in reduced downtime, improved safety, and higher profitability of the process.

2.1 - Existing Measures for Process Monitoring

The goal of fault estimation is to develop measures that are maximally sensitive and robust to all possible faults [1]. Process monitoring measures may be classified in three categories.

1. *Data-driven Measures:* Derived directly from data, the strength of data-driven techniques is in their ability to transform the high-dimensional data into a lower dimension in which important information is captured. The main draw back of data-driven techniques is that their proficiency is highly dependent on the quality and quantity of data. data-driven techniques include Principal Component Analysis (PCA), a technique for optimal dimensionality reduction in terms of capturing the variance in data [2], Fisher Discriminant Analysis (FDA), dimensionality reduction using pattern classification [3], Partial Least Square (PLS), maximizing covariance between predictor and predicted blocks [4], and Canonical Variate Analysis (CVA), a generalized singular value decomposition for maximizing correlation measure between two sets of variables [5].
2. *Analytical Measures:* Analytical approach uses mathematical models of the process that are often derived from first-principles information for process monitoring. Based on the measured inputs and outputs of the process, the analytical methods use analytical model of the process to generate *features* for the process. Commonly used features include residuals, parameter estimates, and state estimates [6]. The main advantage of the analytical approach is the ability to incorporate physical understanding of the process into the process monitoring scheme. With a detailed analytical model, the analytical approach "can significantly outperform the data-driven measures [1]". The disadvantage of the analytical approach is that it is hard to apply this approach to large scale systems, because "detailed models for large scale systems are expensive to obtain given all the cross-coupling associated with a multivariable system [1]".
3. *Knowledge Based Measures:* Uses qualitative models to develop fault monitoring measures. Qualitative models are obtained through causal analysis [7], expert knowledge [8], and pattern recognition techniques (such as artificial neural networks and self-organizing maps) [9]. The advantage of this approach is that it can incorporate operator knowledge directly, and it could be simple despite the complexity of the underlying process. The disadvantage is in the fact that it does not make direct use of the detailed knowledge of the process and hence it is highly dependent of what expert-knowledge is incorporated in the model.

In general no single approach is always preferable to the others. It is therefore important to have flexible framework in which all measures are easily representable. Providing such an inclusive framework and demonstrating its applicability in modelling beam matrix in storage rings at SLAC is the main objective of our Phase I proposal.

2.2 - Fault Estimation in Particle Accelerators - Current Status

Currently, in particle accelerator facilities around the world "the principal fault detectors are the operators [10]", who register a fault upon a malfunction (through a log book or an email) and issue notifications to those responsible. Compared to the alarm monitoring systems in facilities with standardized control systems (such as those in nuclear power plants), the existing alarm systems in accelerator facilities are often in experimental stages and "new monitoring and analysis tools to reduce downtime and to restore safe operation [11]" are in strong demand. This is partly due to the fact that in most of these particle accelerators "the control parameters are frequently changing and evolving and are often themselves under study [10]".

The significance of a reliable fault monitoring system however has long been recognized and efforts to develop systematic fault monitoring systems, varying in scope and the level of success, have been reported [12, 13, 14]. Fault tolerance (especially in facilities with ambitious goals such as experiments in high-energy physics or highly collimated x-ray generators for synchrotron radiation science) is a major design criterion (*e.g.* [15]) and hence the findings of Phase I research is of significance in applications beyond beam matrix modelling that is initially targeted in this proposal.

2.3 - Support Vector Machines in Fault Estimation

Support Vector Machines (SVMs) are a new class of kernel-based techniques that are developed within the disciplines of statistical learning theory and structural risk minimization [16]. SVMs are primarily used for classification and nonlinear function estimation (even though preliminary extensions to recurrent models and optimal control are also reported [17]).

SVMs have attracted considerable attention due to a number of favorable properties. The kernel-based nature of the SVMs allows them to accommodate different types of data (*e.g.* vectors, strings, trees, graphs) that are common in applications ranging from system identification and control [18], to pattern recognition and classification [19], to fault estimation [20] in a straightforward fashion. Additionally, as will be shown in this brief overview, SVM solutions are characterized by *convex optimization problems* (typically quadratic programming) with a unique global minimum (as opposed to many local minima in other classification and learning methods). Furthermore, the model complexity (embodied in the identified kernel) also follows from solving this convex optimization problem. Once trained, kernels incorporate physical knowledge and unlabelled data into the learning algorithm and hence "summarize the relevant features of the primary data, encapsulate ... knowledge, and serve as input to a wide variety of subsequent data analysis [19]."

To provide a background for the problem formulation that is described in Section 4, we briefly describe a simple binary classification problem using SVMs[5]. Given a training set $\{\vec{x}_k, y_k\}_{k=1}^{N}$ with input data $\vec{x}_k \in \mathbb{R}^{n_i}$, and the corresponding binary class label $y_k \in \{-1, +1\}$, the SVM classifier seeks a nonlinear mapping $\phi(\cdot) : \mathbb{R}^{n_i} \to \mathbb{R}^{n_f}$, a corresponding weight vector $\vec{w} \in \mathbb{R}^{n_f}$, and a bias factor $b \in \mathbb{R}$ such that for all $k \in \{1, \cdots, N\}$ the following holds:

$$\begin{cases} \vec{w}^T \phi(\vec{x}_k) + b & \geq & +1, & \text{if } y_k = +1 \\ \vec{w}^T \phi(\vec{x}_k) + b & \leq & -1, & \text{if } y_k = -1 \end{cases} \tag{1}$$

or equivalently $y_k \left[\vec{w}^T \phi(\vec{x}_k) + b\right] \geq +1$, for $k = 1, \cdots, N$. Note that N is the number of data points, n_i is the dimension of the input space, and n_f is the dimension of a so called *feature space* to which the

[5] The exposition in this section follows an excellent description of SVMs in [18].

inputs are mapped using $\phi(\cdot)$. Note that the n_f does not have to be explicitly known. The following optimization problem is then defined to identify $\vec{w}$ and b:

$$\min_{\vec{w},\mathbf{b},\zeta} \mathcal{J}(\vec{w},b,\zeta) = \frac{1}{2}\vec{w}^{T}\vec{w} + c\sum_{k=1}^{N}\zeta_{k} \tag{2}$$

$$\text{subject to:} \quad y_{k}\left[\vec{w}^{T}\phi\left(\vec{x}_{k}\right) + b\right] \geq 1 - \zeta_{k}, \quad \zeta_{k} \geq 0, \quad \text{for} \ \ k = 1, \cdots, N \tag{3}$$

Here $c > 0$ is a tuning parameter, and ζ_k are slack variables allowing for the violation of the constraints, if necessary. It can be shown that the constrained optimization problem in (9) (also known as the *primal* problem) has a *dual* formulation that is often easier to solve:

$$\max_{\{\alpha_{k}\}} \mathcal{Q}\left(\{\alpha_{k}\}\right) = -\frac{1}{2}\sum_{k,l=1}^{N} y_{k}y_{l}K\left(\vec{x}_{k},\vec{x}_{l}\right)\alpha_{k}\alpha_{l} + \sum_{k=1}^{N}\alpha_{k} \tag{4}$$

$$\text{subject to:} \quad \sum_{k=1}^{N}\alpha_{k}y_{k} = 0, \quad , \quad 0 \leq \alpha_{k} \leq c \quad \text{for} \ \ k = 1, \cdots, N \tag{5}$$

where α_k are the dual variables, $K\left(\vec{x}_{k},\vec{x}_{l}\right) = \phi\left(\vec{x}_{k}\right)^{T}\phi\left(\vec{x}_{l}\right)$ is the SVM kernel. Several choices for kernel function are possible. Some of the most commonly used kernels are $K\left(\vec{x},\vec{x}_{k}\right) = \vec{x}_{k}^{T}\vec{x}$ (linear SVM), $K\left(\vec{x},\vec{x}_{k}\right) = \left(\vec{x}_{k}^{T}\vec{x}+1\right)^{d}$ (polynomial SVM of degree d), and $K\left(\vec{x},\vec{x}_{k}\right) = \tanh\left(\kappa\vec{x}_{k}^{T}\vec{x}+\theta\right)$ (multi-layer perceptron SVM). Note that $\vec{w}$ is related to $\vec{\alpha}$ as follows:

$$\vec{w} = \sum_{k=1}^{N}\alpha_{k}y_{k}\phi\left(\vec{x}_{k}\right) \tag{6}$$

To the best of our knowledge the reported applications of SVMs to fault estimation fall within the category of data-driven approaches [20] and our Phase I research has addressed a systematic methodology for its use in model-based fault estimation (especially in large scale systems) that has been lacking.

3 - Outstanding Research Issues

The technical approach discussed in Section 4 is motivated by the need to address the following outstanding research issues:

3.1 - SVM Training in a Combined Model Structure

To the best of our knowledge the training of SVM models in a combined structure (such as Fig. 1) has not been addressed in the literature at all. In phase I, we investigated the training of the SVM models when FP block is a nonlinear *static mapping*. We attempted to derive the most general conditions under which dual problem can be derived. For the examples of interest in Phase I, the primal formulation proved to be easily solvable. Future research will focus on applying Taylor Series expansion to the nonlinear FP block to show that under mild conditions the optimization problem for the combined model remains convex. We also plan to study the training of the SVM model with a *dynamic* nonlinear FP block in the future. We anticipate that constraints from physical knowledge of the process will be critical to successful training of the SVM block in the combined model.

3.2 - Higher-order Fidelity of the Combined Model

Combined models may in general be expected to have superior accuracy compared to a pure FP model. However, for models to be useful for online optimization (*e.g.* in a fault estimation scenario), first/second order gains (*i.e.* derivatives of output w.r.t. input) must also be physically meaningful. To the best of our knowledge, such higher order accuracies are not discussed in the SVM literature. We investigated constrained training of the SVM models (in particular gain-constrained training of the SVMs). An interesting challenge for future research is to see if such constraints could be made data-independent and to evaluate the effect of this generalization on model quality.

3.3 - Optimizer Selection

One of the main advantages of the classical SVM formulation is that the optimization problem is convex. For the example of interest in this project, *i.e.* beam life time model, the training of the combined model of Fig. 1 did not render a convex optimization problem. We experimented with two commercially available nonlinear solvers, LSGRG and SNOPT, in our simulation studies of the combined model training. Future research will compare the performance (*i.e.* quality, speed, robustness of solutions) of a the above mentioned nonlinear solvers applied to the exact non-convex problem vs. that of a convex solver applied to the approximate problem after Taylor Series expansion to better understand the tradeoffs.

3.4 - Implicit First-Principles Models

The standard objective function for the training of the combined model is the sum-squared-error of the FP model outputs. Often, outputs of the FP block in Fig. 1 are only implicit functions of input/state/parameters, *i.e.* FP model is of the form $g(u_k, p_k, x_{k-1}, y_k) = 0$, where y_k is the output vector, u_k is the input vector, p_k is the parameter vector, and x_{k-1} is the state vector, and it is not possible to explicitly define the output as $y_k = G(u_k, p_k, x_{k-1})$. This situation (entirely ignored in the literature on combined modelling) warrants careful examination. At least two different formulations are possible: (a) Include y_k as decision variables for each datapoint. With this choice, the number of datapoints contributes to the problem size, and (b) Use an inner-loop nonlinear solver/optimizer to obtain the y_k values. If there are excess degrees of freedom, this approach leads to an explicit bi-level optimization problem [21] that could be computationally expensive. While the use of the implicit FP models in the context of combined SVM-FP training was beyond the scope of Phase I project, successful use of nonlinear solvers in the course of Phase I research has given new impetus to the need for in depth examination of the implicit FP models in a combined SVM-FP training.

3.5 - Numerical Efficiency of the Training

For NLP optimizers, the required gradients must be calculated analytically or estimated numerically. Analytic gradients may be expected to yield more accurate solutions. For large models, obtaining analytic gradient expressions can be a daunting task. Algorithms for symbolic differentiation (such as those in Maple and GAMS) are well known and are expected to be a part of the eventual commercial software for building combined models. Even with such algorithms, how the required gradients are to be obtained plays a significant role in the choice of problem formulation. For example, if an independent inner-loop solver is used for implicit outputs, obtaining gradients may require a further linear inversion routine.

3.6 - Flexibility of Software Architecture

The software framework for the eventual fault estimation and monitoring product must be flexible to accommodate: (a) easy interface to different solvers (*e.g.* interior point and Sequential Quadratic Programming (SQP)), (b) easy navigation of the problem formulation, constraint sets, SVM weights so that the trained model can be analyzed/debugged, (c) easy integration with procedural programs such as expert systems that might be used for some aspects of monitoring in the system, (d) easy interface with various real-time operating systems used in control rooms and plant floors. We planned to address software architecture issues in the Phase II of this project. Pavilion is investigating the requirements for such software architecture at the moment.

4 - Technical Approach[7]

This section provides a detailed description of the problem formulation for: (a) training of the combined Support Vector Machine (SVM) and First-Principles (FP) model (Section 4.1), (b) fault estimation using the combined model (Section 4.2), and (c) nonlinear mapping to lower dimension feature space (Section 4.3). To the best of our knowledge the training of the SVM block in the context of combined model of Fig. 1 is original. The key innovation in our presentation here is the use of Taylor Series expansion to restore convexity of the training and fault estimation problems.

4.1 - SVM training in a Combined Model Structure

We assume that the first-principles block, in the combined model of Fig. 1, is a parametric nonlinear dynamic mapping described as:

$$\left\{ \begin{array}{lcl} \vec{x}_k & = & F_k\left(\vec{u}_k, \vec{x}_{k-1}, \vec{p}_k\right) \\ \vec{y}_k & = & G_k\left(\vec{u}_k, \vec{x}_{k-1}, \vec{p}_k\right) \end{array} \right. \tag{7}$$

where $\vec{x}_k \in \mathcal{R}^{N_x \times 1}$ is the state vector, $\vec{u}_k \in \mathcal{R}^{N_u \times 1}$ is the input vector, $\vec{y}_k \in \mathcal{R}^{N_y \times 1}$ is the output vector, and $\vec{p}_k \in \mathcal{R}^{N_p \times 1}$ is the parameter vector at time k. Note that, for clarity of the derivation, $\vec{x}_k$ and $\vec{y}_k$ are defined as explicit functions of state/input/parameters. Each element of the parameter vector, $p_{i,k}$, is modelled as a function of process inputs, $\vec{u}_k$, using a SVM:

$$p_{i,k} = \vec{w}_i^T \Phi_i\left(\vec{u}_k\right) + b_i \tag{8}$$

To maintain focus on the central challenge in the training problem (*i.e.* lack of direct access to SVM output(s)), we only provide a detailed derivation for scalar p_k and y_k, noting that extensions to nonscalar case are straightforward (and eliminated in the interest of space):

[7] The material in this section is proprietary and confidential.

$$\min_{\{\vec{w},\mathbf{b},\vec{\zeta}^+,\vec{\zeta}^-\}} \mathcal{J}\left(\vec{w},b,\vec{\zeta}^+,\vec{\zeta}^-\right) = \frac{1}{2}\vec{w}^T\vec{w} + c\sum_{k=1}^{N}\{\zeta_k^+ + \zeta_k^-\} \tag{9}$$

subject to:

$$\begin{cases} y_k - G_k\left(\vec{u}_k,\vec{x}_{k-1},\left[\vec{w}^T\phi(\vec{u}_k)+b\right]\right) & \le & \zeta_k^+ + \epsilon, \ \ \zeta_k^+ \ge 0, \quad \text{for } k=1,\cdots,N \\ -y_k + G_k\left(\vec{u}_k,\vec{x}_{k-1},\left[\vec{w}^T\phi(\vec{u}_k)+b\right]\right) & \le & \zeta_k^- + \epsilon, \ \ \zeta_k^- \ge 0, \quad \text{for } k=1,\cdots,N \end{cases} \tag{10}$$

This constrained optimization problem, also known as the *primal problem*, differs from that studied in the SVM literature due the presence of $G_k(\cdot)$ in the constraint set (10). This constraint set is no longer necessarily convex in $\vec{w}$, and hence the primal problem is convex only if $G_k(\cdot)$ is convex in $\vec{w}$[8].

For a general nonlinear function $G_k(\cdot)$, we propose to restore the convexity of the constraint set, under mild conditions, using Taylor series expansion of $G_k(\cdot)$ around $(\vec{u}_0,\vec{x}_0,\vec{p}_0)$ as follows:

$$y_k - \frac{\partial G_k}{\partial \vec{u}_k}(\vec{u}_k - \vec{u}_0) - \frac{\partial G_k}{\partial \vec{x}_{k-1}}(\vec{x}_{k-1}-\vec{x}_0) - \frac{\partial G_k}{\partial p_k}\left(\overbrace{\left[\vec{w}^T\phi(\vec{u}_k)+b\right]}^{p_k} - p_0\right) \le \zeta_k^+ + \epsilon, \ \ \zeta_k^+ \ge 0, \ \text{for } k=1,\cdots,N \tag{11}$$

$$-y_k + \frac{\partial G_k}{\partial \vec{u}_k}(\vec{u}_k - \vec{u}_0) + \frac{\partial G_k}{\partial \vec{x}_{k-1}}(\vec{x}_{k-1}-\vec{x}_0) + \frac{\partial G_k}{\partial p_k}\left(\overbrace{\left[\vec{w}^T\phi(\vec{u}_k)+b\right]}^{p_k} - p_0\right) \le \zeta_k^- + \epsilon, \ \ \zeta_k^- \ge 0, \ \text{for } k=1,\cdots,N \tag{12}$$

The constraint set (12) is convex in $\vec{w}$ if $\frac{\partial G_k}{\partial \vec{u}_k}$ and $\frac{\partial G_k}{\partial \vec{x}_k}$ can be bounded[9]. To derive the dual problem, we will follow the standard derivation procedure in primal-dual optimization problems:

1. Define the Lagrangian as follows:

$$\begin{aligned} \mathcal{L}\left(\vec{w},b,\vec{\zeta}^+,\vec{\zeta}^-;\vec{\alpha}^+,\vec{\alpha}^-,\vec{\nu}^+,\vec{\nu}^-\right) &= \mathcal{J}(\vec{w},b,\vec{\zeta}^+,\vec{\zeta}^-) - \sum_{k=1}^{N}\nu_k^+\zeta_k^+ - \sum_{k=1}^{N}\nu_k^-\zeta_k^- \\ &- \sum_{k=1}^{N}\alpha_k^+\left(y_k - G_k\left(\vec{u}_k,\vec{x}_{k-1},\left[\vec{w}^T\phi(\vec{u}_k)+b\right]\right) - \zeta_k^+ - \epsilon\right) \\ &- \sum_{k=1}^{N}\alpha_k^-\left(-y_k + G_k\left(\vec{u}_k,\vec{x}_{k-1},\left[\vec{w}^T\phi(\vec{u}_k)+b\right]\right) - \zeta_k^- - \epsilon\right) \end{aligned} \tag{13}$$

2. Solve the following max-min optimization problem to obtain the saddle point of the Lagrangian that characterizes the solution to the primal problem:

$$\max_{\{\vec{\alpha}^+,\vec{\alpha}^-,\vec{\nu}^+,\vec{\nu}^-\}} \min_{\{\vec{w},b,\vec{\zeta}^+,\vec{\zeta}^-\}} \mathcal{L}\left(\vec{w},b,\vec{\zeta}^+,\vec{\zeta}^-;\vec{\alpha}^+,\vec{\alpha}^-,\vec{\nu}^+,\vec{\nu}^-\right) \tag{14}$$

for which $\frac{\partial\mathcal{L}}{\partial\vec{w}} = 0$, $\frac{\partial\mathcal{L}}{\partial b} = 0$, and $\frac{\partial\mathcal{L}}{\partial\zeta_k} = 0$ are the necessary conditions for optimality.

3. Replacing $\frac{\partial G_k}{\partial \vec{u}_k}$ and $\frac{\partial G_k}{\partial \vec{x}_k}$ with appropriate bounds, the dual problem is as follows:

[8]Convexity is a desired feature for the primal problem. Convex optimization problems are computationally well-behaved. They have unique optima and avoid complications with local solutions that frequently occur in other nonlinear modelling approaches such as NNs to the potential detriment of the model.

[9]The effect of the more conservative constraints obtained by such bounds on the quality of the trained model will be carefully studied.

$$\max_{\{\vec{\alpha}^+,\vec{\alpha}^-\}} \mathcal{J}^*\left(\vec{\alpha}^+,\vec{\alpha}^-\right) = -\frac{1}{2}\sum_{k,l=1}^{N}\left(\alpha_k^+ - \alpha_k^-\right)\left(\alpha_l^+ - \alpha_l^-\right)\left(\frac{\partial G_k}{\partial \vec{w}}\right)^T\left(\frac{\partial G_k}{\partial \vec{w}}\right)$$

$$-\epsilon\sum_{k=1}^{N}\left(\alpha_k^+ + \alpha_k^-\right) + \sum_{k=1}^{N} y_k\left(\alpha_k^+ - \alpha_k^-\right) \tag{15}$$

$$\text{subject to: } \sum_{k=1}^{N}\left(\alpha_k^+ - \alpha_k^-\right) = 0, \quad 0 \le \alpha_k^+, \alpha_k^- \le c, \quad \text{for} \quad k = 1, \cdots, N \tag{16}$$

where α_k^+, α_k^- are the dual variables. Note that $\vec{w}$ is related to $\vec{\alpha}$ as follows:

$$\vec{w} = \sum_{k=1}^{N}\left(\alpha_k^+ - \alpha_k^-\right)\left(\frac{\partial G_k}{\partial \vec{w}}\right) \tag{17}$$

4.2 - Optimization Problem for Fault Estimation

Once the combined SVM & FP model is trained to represent a process, the combined model can be used to estimate the onset of a fault given: (a) process measurements $\vec{y}$, (b) knowledge of disturbance sources in the form of their probability distribution function (pdf), and (c) any other operational data that might be available:

$$\min_{\vec{u}} \left(\vec{y} - \hat{\vec{y}}\right)^T \mathcal{Q}^{-1}\left(\vec{y} - \hat{\vec{y}}\right) + \sum_{\nu=1}^{V}\mathcal{J}_\nu \tag{18}$$

subject to:

$$\begin{cases} \hat{\vec{y}} &= G_k\left(\vec{u}_k, \vec{x}_{k-1}, \left[\vec{w}^T\phi\left(\vec{u}_k\right) + b\right]\right), \\ u_j &\in \left[u_{j,min}, u_{j,max}\right] \qquad \text{for jth input} \end{cases} \tag{19}$$

where $\mathcal{Q}$ is the covariance of output measurement noise (and also a tuning parameter), $\mathcal{J}_\nu = \log\left(p_{u_\nu|\vec{y}}\right)$ where $p_{u_\nu|\vec{y}}$ is the conditional probability density function for νth element of the input vector assuming u_ν, given the measurement of the output $\vec{y}$. In [22] it is shown that for common pdfs (such as Gaussian noise, Uniform noise, Laplacian noise, exponential noise) the log term in the cost function of Eq. (18) is convex in input vectors, and hence the fault diagnosis problem remains convex if $G_k(\cdot)$ is convex (or can be made convex via bounding techniques)[10].

4.3 - Optimization Problem for Dimension Reduction

One of the main features of the Support Vector Machine (SVM) component of the combined model we propose is its ability to efficiently handle large number of inputs and to effectively remove the effect of noise in measurements. As we will show in this section, the SVM block performs two roles in the combined model of Fig. 1, both of which are crucial to the successful use of the combined models in fault estimation problems in real world applications:

1. The SVM block is an efficient nonlinear function approximation that captures the potentially nonlinear variation of the system parameters as a function of process inputs.

2. The SVM block is capable of performing a nonlinear Principal Component Analysis (PCA) by mapping the large-dimension input spaces into selectably lower dimension *feature* spaces where redundancy in input data is removed. This mapping can be achieved via a convex optimization

[10] Today, an "off-the-shelf solver allows achieving 200-500 millisecond update on a PC in a realistic aircraft application. The computation time can be further reduced by 1-2 orders of magnitude by developing specialized solvers." [26] Such efficiency in our view can lead to ground breaking applications in online accelerator monitoring. Examination of the computational efficiency of the fault estimation is a key objective of the Phase I project.

(hence extremely efficiently for realtime applications). The lower dimension feature space can be displayed for operators.

The problem formulation for the nonlinear mapping from large-dimension input space to lower dimension feature space is straightforward [33]:

$$\min_{\vec{w},\vec{e}} \frac{1}{2}\vec{w}^T\vec{w} - \gamma\frac{1}{2}\sum_{k=1}^{N} e_k^2 \tag{20}$$

$$\text{subject to:} \quad e_k = \vec{w}^T \phi\left(\vec{u}_k\right) - \hat{\mu}_\phi \tag{21}$$

where $\hat{\mu}_\phi = \frac{1}{N}\sum_{k=1}^{N}\phi(\vec{u}_k)$ and the optimization problem attempts to force as many elements of the weight vector $\vec{w}$ to zero as possible. Of course, this objective must be balanced against the nonlinear function approximation property of the combined model as reflected in the optimization problem of Eqs. (9) and (10). This leads to a multi-objective optimization problem, the solution of which and the study of the potential trade-offs is a major topic for our future research.

5 - Phase I Research Results

The main technical objectives of the Phase I research were to demonstrate that: (a) constrained training of the support vector machines in the hybrid model structure of Fig. 1 is possible, and (b) the combined model can be used in a model-based fault detection scenario to identify the most likely source of an observed deviation from normal operation. We have achieved both these objectives by:

1. *Successful SVM Training in a Combined Model Structure*: We developed the software for the constrained training of the SVMs in a combined model structure, and successfully modeled the parameters of a first-principles model for beam lifetime with support vectors.

2. *Imposing Higher-order Fidelity on the Combined Model*: We used constrained training to ensure that the output of the SVM (*i.e.* the parameters of the beam lifetime model) are physically meaningful.

3. *Improving Numerical Efficiency of the Training*: We investigated the numerical efficiency of the SVM training. More specifically, for the primal formulation of the training, we have developed a problem formulation that avoids the linear increase in the number of the constraints as a function of the number of data points.

4. *Enabling Flexibility in Software Architecture*: The software framework for the training of the support vector machines was designed to enable experimentation with different solvers.

5.1 - Fault Detection for Beam Lifetime using Hybrid SVM Models

Synchrotron light is used for a wide variety of scientific disciplines ranging from physical chemistry to molecular biology and industrial applications. The synchrotron light is radiated from a relativistic electron beam circulating in a storage ring particle accelerator [23, 24, 25]. As the electron beam circulates, random single-particle collisional processes lead to decay of the beam current in time. As the electron beam decays, so does the intensity of the synchrotron light resulting in detuned optics in the photon transport lines (*e.g.* mirrors, gratings, slits), changes in material properties of the experimental sample, degradation of detector performance and uncertainties in data reduction. Hence, at all synchrotrons, a premium is placed on delivering constant photon beam intensity to the photon beam lines [27].

Efforts to systematically model the electron beam loss in synchrotron light sources have therefore been of primary interest. At SPEAR3, for example, every two weeks, up to 48 hrs of beam time is allocated for machine development studies, including programs to measure, characterize, and mitigate electron beam loss. In this section we introduce the hybrid SVM and first-principles framework for the systematic modeling of electron beam loss. This framework, also known as **P**arametric **U**niversal **N**onlinear **D**ynamics **A**pproximator (PUNDA), consists of a series connection of a Nonlinear Empirical Model (NEM) block (support vector machine in this case) and a Parametric First-principles Model (PFM) block (see Figure 1). The parameters, $\vec{P}_k$, in the PFM block may vary as a function of process inputs, $\vec{u}$. For the beam loss model of interest in this study, the instantaneous electron beam current, Eq. (22), constitutes the PFM block, where the characteristic beam decay time constant is the varying parameter. A support vector machine model forms the NEM block. This SVM model is trained to capture the variation in the characteristic beam decay time constant as a function of variation in vertical scraper position (Y_s), RF voltage (V_rf), initial beam current (I_0), and total number of bunches (M_b).

The ultimate goal of this study is to build accurate and computationally efficient models that quantify beam loss from electron-gas scattering (elastic and inelastic) and intrabeam scattering (electron-

electron). Once such models are constructed, an optimization-based approach may be adopted to identify potential sources of beam loss when beam loss is detected. Furthermore, having correctly identified *current* operating conditions of the beam, appropriate corrective actions within operation constraints for the beam, may be determined in a timely fashion to potentially prevent further deterioration of beam quality (hence extending beam life time).

5.1.1 - Physics of Electron Beam Loss: The physics behind electron beam loss is conceptually straight-forward but nevertheless a highly non-linear process. In principle, the electron beam current decays in time due to (a) elastic electron-gas collisions (Coulomb scattering) (b) inelastic electron-gas collisions (Bremsstrahlung scattering), and (c) intrabeam electron-electron collisions [28, 29, 30, 32]. In this section, global models for electron beam loss are briefly described.

At any given time t, the instantaneous electron beam current may be written as:

$$I_t(t) = I_0 e^{\frac{-t}{\tau}} \tag{22}$$

where I_0 is the initial beam current, and τ is the characteristic beam decay time constant. Due to the uncorrelated nature of the collisional processes, the characteristic decay time depends on the individual contributions from Coulomb, Bremsstrahlung, and Intrabeam sources for scattering. Given that we can only measure the net beam decay time, (τ), the gas and intrabeam components must be inferred from an array of measurements under different experimental conditions. For the simulation study in this paper, τ is assumed to be a non-linear function of vertical scraper position (y_s), RF voltage (V_{rf}), initial beam current (I_0), and total number of bunches (M_b).

Over longer time periods of time, electron beam loss deviates from pure exponential and is governed by a more general rate equation:

$$\frac{dN_e}{dt} = \sum_i A_i \; N_e \; N_i \; \sigma_i \tag{23}$$

where N_e is the number of electrons, N_i is the number of scattering centers, σ_i is the scattering cross section for each type of collision, and A_i are characteristic proportionality constants. The cross sections σ_i quantify the probability of particle loss for each collision process. Note that integration of the rate equation, $\frac{dN}{dt} = -\alpha N^2$, yields a beam decay profile in time:

$$N(t) = \frac{N_0}{1 + N_0 \; \alpha \; t} \sim N_0 e^{-\frac{t}{\tau}} \tag{24}$$

for short times t. The long-term decay curve is more complicated than the exponential expression for instantaneous decay because the density of scattering centers is reduced roughly in proportion to circulating beam current.

The objective of the current study is to build PUNDA models that accurately predict the electron beam decay as a function of electron beam parameters and synchrotron operating parameters. The training of the SVM block in the PUNDA model will be constrained by first-principles models (*i.e.* particle collision physics) for each scattering mechanism, and hence the trained model will be physically meaningful.

5.2 - Simulation Results

For the simulation study in this paper the following first-principles model is used to describe the electron beam loss:

$$I_t(t) = \frac{I_0 e^{-bt}}{1 + \left(\frac{a}{b}\right) I_0 \left(1 - e^{-bt}\right)} \tag{25}$$

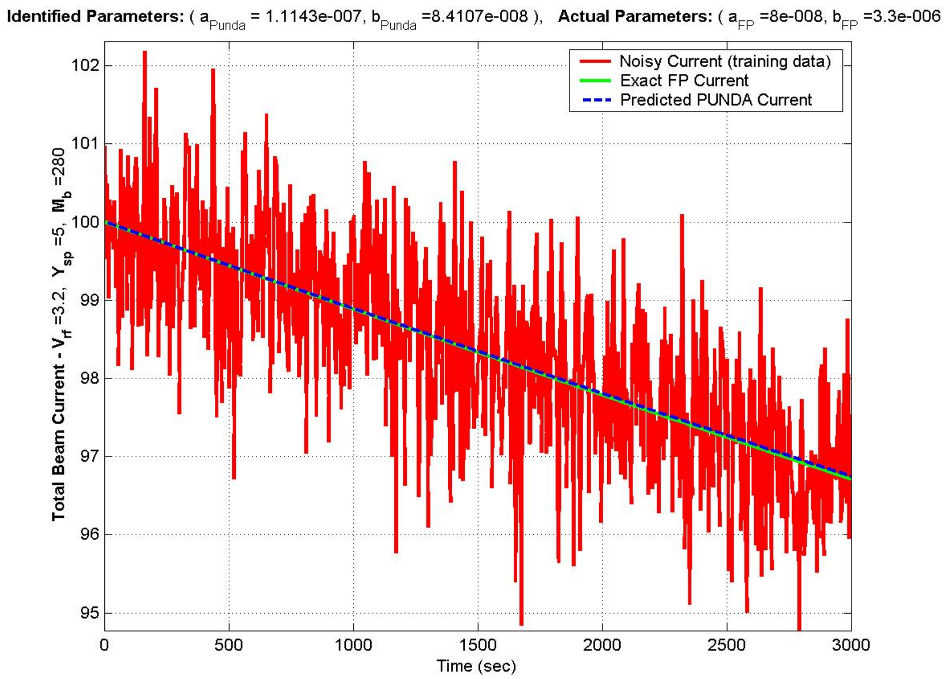

Figure 2: Prediction of electron beam decay using PUNDA model. The beam decay is accurately predicted while acceptable estimates of the decay model parameters are obtained. The actual parameter values used in the FP model of Eq. (25) are reflected in the title of the plot. Under the given beam operating conditions, indicated along the y axis, beam current decays *linearly* over the course of approximately one hour of operation.

where I_0 is the initial beam current, and a and b are parameters of the parametric model for electron beam loss that are functions of the beam operating conditions.

The parameter $a = a_T + a_B + a_C$, is affected by Touschek (a_T), Bremsstrahlung (a_B), and Coulomb (a_C) effects on beam loss [31] which in turn depend on gap voltage (V_{rf}), vertical scraper position (Y_{s}), dynamic vacuum pressure (P_{dyn}) and number of bunches (M_{b}). The parameter $b = b_B + b_C$ is affected by Bremsstrahlung (b_B) and Coulomb (b_C) collisions due to the base pressure.

For the simulation results shown in this section, four major parameters, *i.e.* Y_{s}, V_{rf}, M_{b}, and I_0 are varied over an operation range consistent with that at SPEAR3. Electron beam loss is simulated using Eq. (25). Random noise is added to the simulated electron beam current to reflect imperfect current measurements. The noisy electron beam current is then used to construct a PUNDA model where Eq. (25) constitutes the *Parametric First Principles Model* block, and an SVM model constitutes the *Nonlinear Empirical Model* block. The combined PUNDA model is trained via constrained optimization, identifying appropriate parameter values for the FP model (*i.e.* a and b), at the same time that the decay in electron beam current is modeled. Figure 2-7 capture typical simulation results.

PUNDA structure offers a framework in which both beam data and first principles models may be used to complement one another. The nonlinear empirical model block may be used to capture the less known aspects of the beam decay that is reflected in the operation data but is not fully explained by first-principles information. Once a PUNDA model is verified to capture the beam loss in a particle accelerator, the model can be used to identify potential sources of beam loss in real-time.

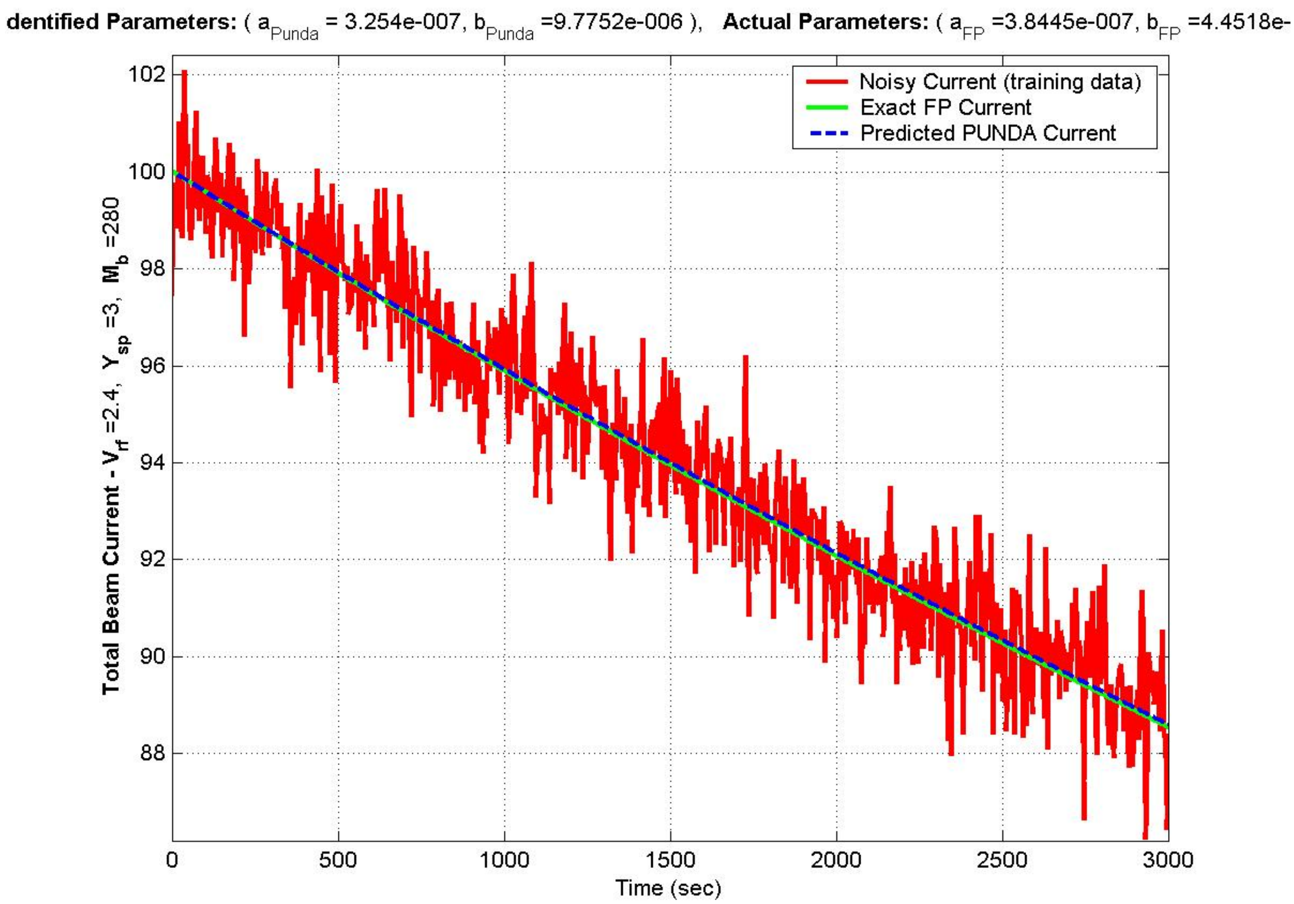

Figure 3: Prediction of electron beam decay using PUNDA model where gap voltage (V_{rf}) is reduced from 3.2 in Fig. 2 to 2.4, resulting in a larger loss in beam current.

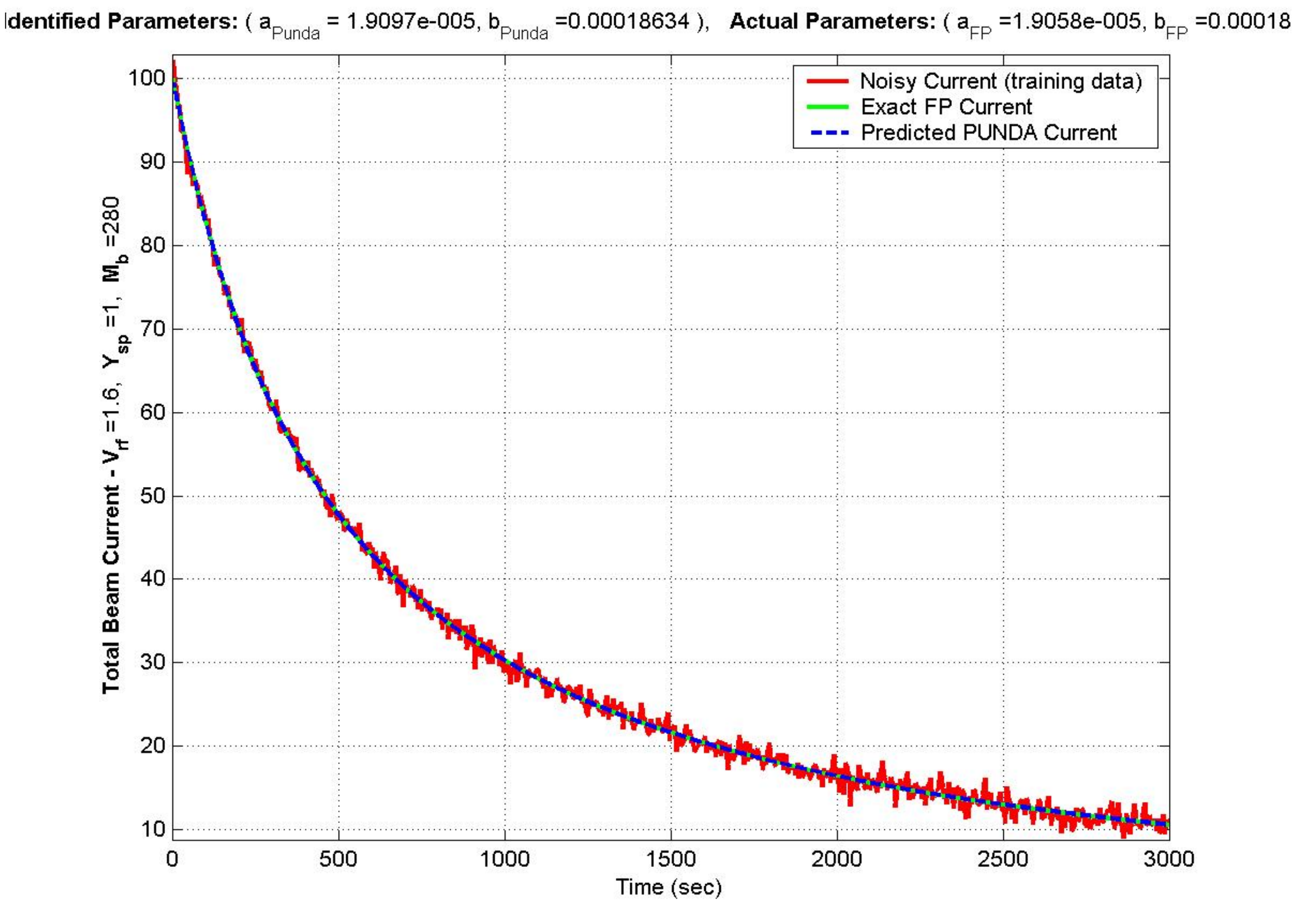

Figure 4: Prediction of electron beam decay using PUNDA model where gap voltage (V_{rf}) is reduced to 1.6 and vertical scraper position (Y_{s}) is set to 1. The beam decay in this case demonstrates an exponential decay (to significantly lower values of beam current) over the hour long simulation horizon.

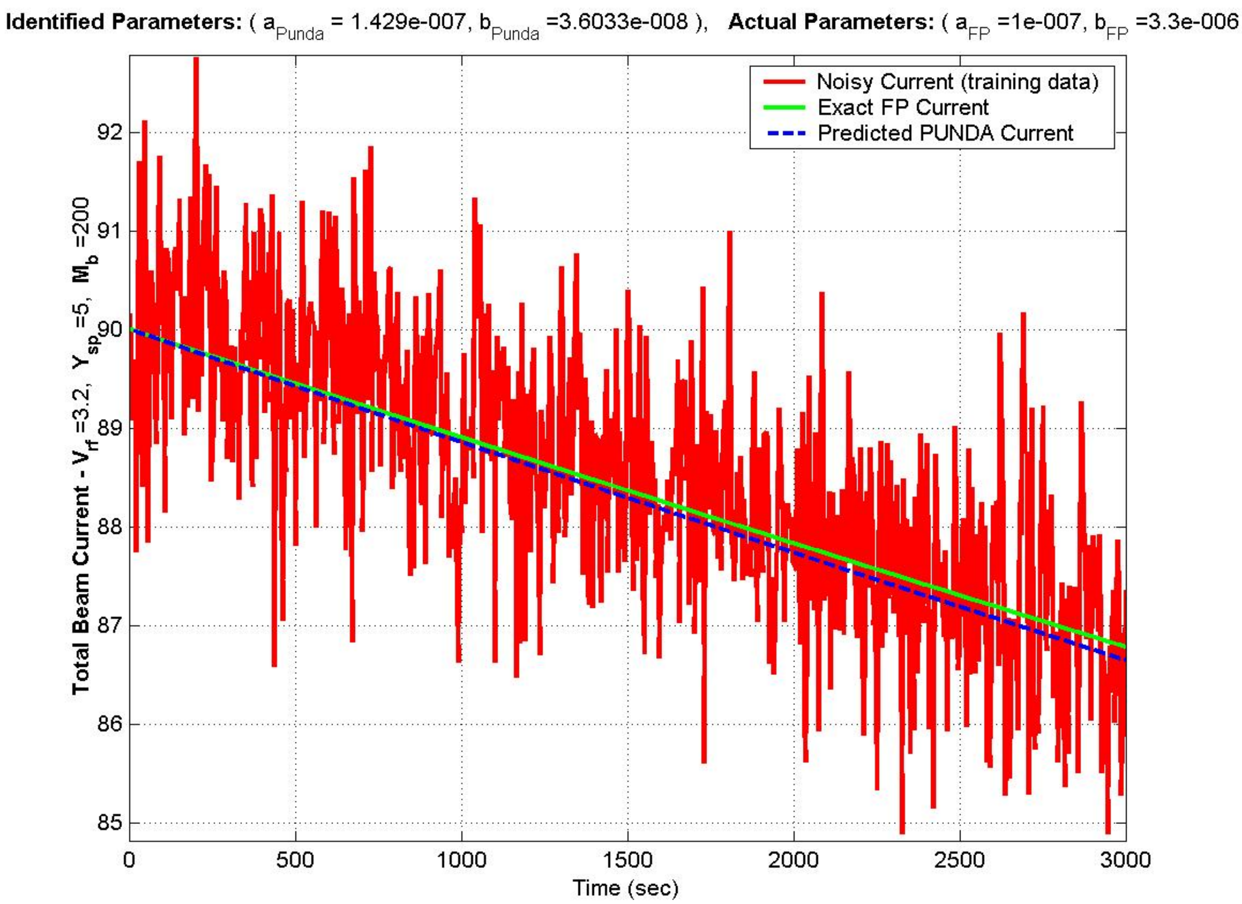

Figure 5: Prediction of electron beam decay using PUNDA model where gap voltage ($V_{\rm rf}$) and vertical scraper position ($Y_{\rm s}$) are the same as in Fig. 2, while number of bunches ($M_{\rm b}$) is reduced 200. The beam displays a linear decay (from a lower value of initial current compared to that in Fig. 2).

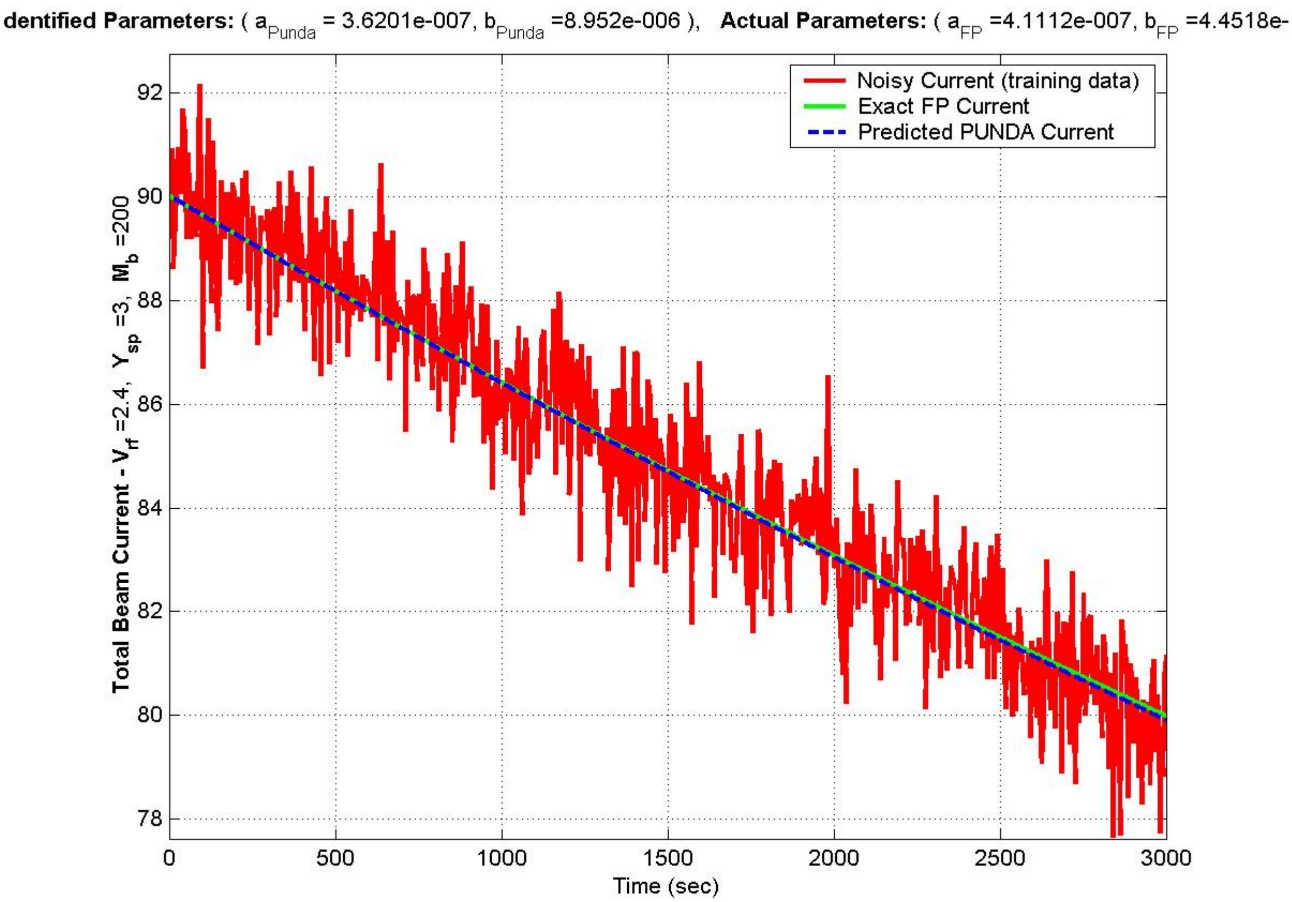

Figure 6: Prediction of electron beam decay using PUNDA model where gap voltage (V_{rf}) is reduced from 3.2 in Fig. 5 to 2.4, resulting in a larger loss in beam current.

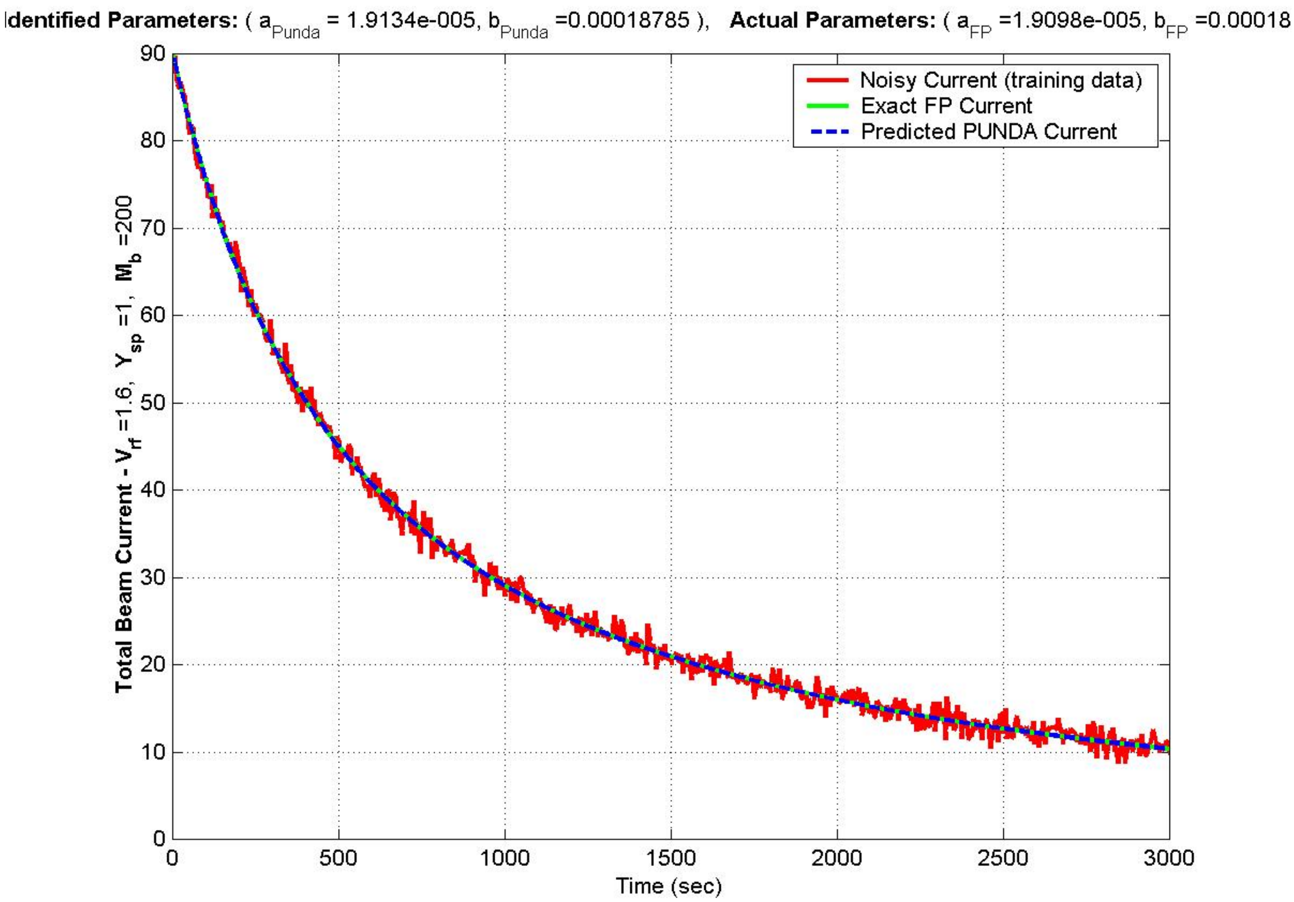

Figure 7: Prediction of electron beam decay using PUNDA model where gap voltage (V_{rf}) is reduced to 1.6 and vertical scraper position (Y_{s}) is set to 1, while number of bunches (M_{b}) is kept at the level of Fig. 5. The beam decay in this case demonstrates an exponential decay (to significantly lower values of beam current) over the hour long simulation horizon.

6 - Anticipated Benefits of Phase I Research

As mentioned in the background section of this proposal (Section 2), analytical and data-driven approaches to fault estimation are highly complementary paradigms. Our proposed Combined Support Vector Machine (SVM) and First-Principles (FP) approach offers a framework in which measured process data and first-principles/operational knowledge are *optimally* utilized for fault estimation/correction. In our view, the combined approach will emerge as one of the (if not *the*) most logical and useful frameworks for fault monitoring in complex system applications both online and off-line for the following reasons:

1. Effective handling of large-scale input spaces, and
2. Computational efficiency of the combined models for online fault monitoring and diagnosis.

Even though, the recent acquisition of Pavilion by Rockwell Automation alters the small business status of Pavilion, disqualifying us from pursuing further SBIR funding, we are convinced that further investment will enable the development of a commercial, integrated software designed specifically for easy construction, deployment, and maintenance of the relevant models for fault estimation/correction.

6.1 - Commercial Potential, Markets, Customers, and Competition

We believe that the combined SVM & FP paradigm for fault estimation/diagnosis targeted in this research represents such a powerful and enabling technology that it could lead to a major upsurge in the use of online *model-predictive* fault estimation/diagnosis in a variety of industries. Currently, there is no commercial product that offers such capability, and Pavilion plans to pursue further funding for to make sure that it will be the first to offer this technology to the market.

The successful completion of the Phase I project has provided SLAC with an unprecedented opportunity to define a research project with an ultimate goal to construct nonlinear beam matrix models that are particularly suitable for online fault estimation and correction. The new combined SVM and FP framework is also a versatile analysis tool for post-mortem fault analysis at SLAC and other accelerator/storage ring facilities.

While the results of this research will be very broadly applicable, Pavilion's existing customer base in process industry[11] will provide the initial target market for the innovations in this proposal. Of particular interest to the DOE mission, is Pavilion's biofuel manufacturing customer base[12]. The market size for model-based control/optimization/fault estimation is approximately \$450M[13]. Some of the customers have expressed willingness for collaboration to pursue additional funding for the commercialization of the fault estimation product once a prototype becomes available.

In addition to process industry in general, we believe that the high fidelity and cost benefits of a combined model-based fault estimation/correction system will attract customers from a wide variety of commercial and scientific industries. Applications with a need for high-accuracy and high speed fault estimation systems (such as avionics and underwater submarines) could find significant benefits in the findings of this project.

[11] Process Industry includes, for instance, polymer manufacturers such as Montel, BP-Amoco, Dow-Carbide, and Chevron; and power generation companies such as Sithe Energies, Virginia Power, South Carolina Electric and Gas, and Duke Power; Bio-fuel production companies such as Broin; food processing such as Kellogg's and Nestle. The above, as well as many dozens of other process industry companies, are already customers of Pavilion Technologies, Inc.

[12] Currently 15 biofuel companies, including Golden Grain Energy, Michigan Ethanol, Glacier Lakes Energy, Great Plain Ethanol, and Broin Companies are Pavilion's customers.

[13] There are 518 target plants for which model-based automation will be economically feasible. With an approximate 3 process lines per plant, and an approximate cost of \$300K per line for an automation solution, the market size is approximately \$450M.

Appendix A - XML File for Constrained Optimization of SVMs

We developed an appropriate language to formulate the constrained optimization for the training of the Support Vector Machines (SVMs). The following is a sample xml file that is used for the training of the SVM models for the two parameters of the beam lifetime model.

```
<formulation>
  <problem classname=’com.pav.model.hybrid.optimization.common.NLPProblem’ direction=’minimize’>
    <variable label=’weights’ modelVariable=’weights’ randomLower=’-0.05’ randomUpper=’0.05’
              useRandomInitial=’true’ randomSeed=’111’>
      <lower>
        [ -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        -1.0E30, -1.0E30, -1.0E30, -1.0E30, -1.0E30,
        0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0,
        0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]
      </lower>
      <upper>
        [ 1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        1.0E30, 1.0E30, 1.0E30, 1.0E30, 1.0E30,
        10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0,
        10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0, 10.0]
      </upper>
    </variable>
    <objective label=’obj’ modelVariable=’obj’/>
  </problem>

  <mapping classname=’com.pav.model.hybrid.mapping.expression.ExpressionMapping’>
    <mapping classname=’com.pav.model.hybrid.mapping.mlp.MlpMapping’ name=’mlp’>
      <input name=’x1’ layer=’layer0’ node=’0’ scale=’4’ bias=’1’/>
      <input name=’x2’ layer=’layer0’ node=’1’ scale=’2’ bias=’1.2’/>
      <input name=’x3’ layer=’layer0’ node=’2’ scale=’240’ bias=’40’/>
      <input name=’x4’ layer=’layer0’ node=’3’ scale=’80’ bias=’20’/>
      <output name=’a’ layer=’layer2’ node=’0’ scale=’0.0001’/>
      <output name=’b’ layer=’layer2’ node=’1’ scale=’0.001’/>
      <layer name=’layer0’ size=’4’ type=’identity’/>
      <layer name=’layer1’ size=’16’ type=’sigmoid’/>
      <layer name=’layer2’ size=’2’ type=’linear’/>
      <connection srcName=’layer0’ dstName=’layer1’/>
      <connection srcName=’layer1’ dstName=’layer2’/>
```

```
</mapping>

<set name='W'>
  <elements type='size' size='114'/>
</set>
<set name='J'>
  <elements type='range' start='1' stop='18'/>
</set>
<set name='K'>
  <elements type='size' size='601'/>
</set>
<set name='C'>
  <elements type='range' start='1' stop='3'/>
</set>
<set name='P'>
  <elements type='range' start='1' stop='5'/>
</set>
<input name='weights' isScalar='false'>
  <dimension type='ID' set='W'/>
</input>
<output name='a' isScalar='false'>
  <dimension type='ID' set='J'/>
</output>
<output name='b' isScalar='false'>
  <dimension type='ID' set='J'/>
</output>
<output name='I_c_model' isScalar='false'>
  <dimension type='ID' set='J'/>
  <dimension type='ID' set='K'/>
</output>
<output name='obj'/>

<parameter name='batchData' isScalar='false' file='SLAC_BeamLifetime/BatchDataFile.ssv'>
  <dimension type='ID' set='J'/>
  <dimension type='ID' set='P'/>
</parameter>
<parameter name='dynamicData' isScalar='false' file='SLAC_BeamLifetime/CurrDataFile.ssv'>
  <dimension type='ID' set='J'/>
  <dimension type='ID' set='K'/>
  <dimension type='ID' set='C'/>
</parameter>

<expressions>
  real y_scraper[J];
  real V_RF[J];
  real M_b[J];
  real I_t[J];

  y_scraper[j] = batchData[j,2] foreach j in J;
  V_RF[j] = batchData[j,3] foreach j in J;
  M_b[j] = batchData[j,4] foreach j in J;
  I_t[j] = batchData[j,5] foreach j in J;

  real t[J,K];
  real I_c[J,K];
  t[j,k] = dynamicData[j,k,2] foreach j in J, k in K;
  I_c[j,k] = dynamicData[j,k,3] foreach j in J, k in K;

  //real a[J], b[J];
  //real I_c_model[J,K];

  mlp( x1:y_scraper[j], x2:V_RF[j], x3:M_b[j], x4:I_t[j], weights:weights[*];
  a:a[j], b:b[j] ) foreach j in J;

  I_c_model[j,k] = I_t[j] * exp(-b[j]*t[j,k])/(1 + (1-exp(-b[j]*t[j,k]))*I_t[j]*a[j]/b[j] ) foreach j in J, k in K;

  obj = sum( j in J : sum( k in K : (I_c_model[j,k] - I_c[j,k])^2 ) );
```

```
    </expressions>
  </mapping>

  <solver classname='com.pav.model.hybrid.optimization.snopt.Snopt'>
    <parameter name='major feasibility tolerance' value='1e-4'/>
    <parameter name='major optimality tolerance' value='1e-4'/>
    <!--    <parameter name='hessian full memory' value='true' />-->
    <!--    <parameter name='hessian frequency' value='10' />-->
    <parameter name='major print level' value='1'/>
    <parameter name='print file' value='6'/>
    <parameter name='solution' value='true'/>
    <parameter name='solution file' value='6'/>
    <parameter name='print parameters' value='true'/>
  </solver>

</formulation>
```

Appendix B - XML File for Optimization Problem for Fault Detection

The same language used for the training of the SVMs enabled constrained optimization for fault detection in a synthetic example for beam life time. The following is a sample xml file that is used for this purpose.


```
<config
classname="com.pav.model.hybrid.optimization.common.OptimizationFormulation">
  <mapping classname="com.pav.model.hybrid.mapping.expression.ExpressionMapping" name="">
    <set name="W">
      <elements type="size" size="98" />
    </set>
    <set name="J">
      <elements type="range" start="1" stop="4"/>
    </set>
    <set name="K">
      <elements type="range" start="1" stop="1202" />
    </set>
    <set name="C">
      <elements type="range" start="1" stop="9" step="1" />
    </set>
    <parameter name="data" isScalar="false" file="FaultDetection/FD_data.ssv">
      <dimension type="ID" set="K" />
      <dimension type="ID" set="C" />
    </parameter>
    <input name="y_scraper" isScalar="false">
      <dimension type="ID" set="J" />
    </input>
    <input name="V_RF" isScalar="false">
      <dimension type="ID" set="J" />
    </input>
    <input name="M_b" isScalar="false">
      <dimension type="ID" set="J" />
    </input>
    <input name="I_t" isScalar="false">
      <dimension type="ID" set="J" />
    </input>
    <input name="weights" isScalar="false">
      <dimension type="ID" set="W" />
      <values>[-8.658118395412577, -1.6880783200218543, -1.0331689092183722, -0.7024246586677668, -8.544753017372386,
               -1.6980953425768277, -1.58726110328377, -2.409046805447259, -1.769507244627045, -0.6382638712294072,
               -0.7122577225604569, 0.23012355738953164, -1.6990281004832828, -0.4615598968565889, -0.8417877875672115,
                0.2249875231377209, -2.9860919834858137, -4.424203974106881, -2.0188440467554947, -2.078737970742437,
               -1.2321615647339885, -0.5353655017178139, -0.4789805807036303, -0.1319946157343967, 1.6155136281849811,
               -4.987221660214916, -0.18525719280291567, 3.4173863815314673, -12.040383467728239, -3.361361820470712,
               -3.1882532273439823, -4.164846085776751, -7.206391829946805, 1.343766810648909, 1.5829049997121807,
                1.1511909706065158, -11.609370029036622, -3.4443598512574565, -3.330785237676132, -4.242652690106668,
               -12.778098337694201, -3.1963840364140643, -3.079874177987541, -4.0275812558277675, -1.8296727273338973,
               -0.6083429958072337, -0.5590938141071454, 0.280158437519067, -12.905145098084468, -0.08718923831641946,
                0.029473385308586772, -0.9845811661702158, 1.7948186872929583, -3.6793817121727934, -2.3917243550201572,
                1.5221886259468704, -1.4525204382591435, -0.46878513297868224, -0.4307827075639602, 0.6237528675530384,
                0.5869313213149651, -6.3950742655252455, -8.041619829046484, 1.95958406176998, 2.672250704968546,
                4.445482378336362, 2.479261027514004, 2.561103210533547, 0.7340469219814676, 1.3909877095245216,
                2.1583869856611666, 3.7362448805753936, 5.1407259134887715, 3.543161866846364, 3.8366903582353933,
                2.4885426350188156, 3.3083018004920435, 3.1063615069631347, 2.4801527875520004, 2.1252509977242093,
                0.10936682877917034, 0.06549096815859268, 1.3426303269851776, -3.910467611362013E-7,
                7.325055065908214E-4, 0.05601919537696628, 7.535793283863449E-4, 0.0011355953798053073,
                0.0022096513932162274, 0.0016352811921856814, 2.1092343715805898E-6, 0.3131246190914335,
                0.00534596330771736, 0.6830934228500399, 0.0024981279312064103, -3.890745734568386E-7,
                0.1699301531603487, 0.0020966297697849992]</values>
    </input>
    <output name="a" isScalar="false">
      <dimension type="ID" set="J" />
    </output>
    <output name="b" isScalar="false">
      <dimension type="ID" set="J" />
    </output>
```

```
  <output name="I_c_model" isScalar="false">
    <dimension type="ID" set="K" />
  </output>
  <output name="obj" value="0.185428340965721" />
  <expressions>
    set Ks[J] = [ {1:200}, {201:601}, {602:801}, {802:1202} ];
    real t[K];
    real I_c[K];

    t[k] = data[k,1] foreach k in K;
    I_c[k] = data[k,2] foreach k in K;

    //real a[J], b[J];
    //real I_c_model[J,K];

    mlp( x1:y_scraper[j], x2:V_RF[j], x3:M_b[j];  a:a[j], b:b[j] )  forall j in J;

    I_c_model[k] = I_t[j] * exp(-b[j]*t[k])/(1 + (1-exp(-b[j]*t[k]))*I_t[j]*a[j] ) foreach j in J, k in Ks[j];

    obj = sum( k in K : (I_c_model[k] - I_c[k])^2 );
  </expressions>
  <mapping classname="com.pav.model.hybrid.mapping.mlp.MlpMapping" name="mlp">
    <input name="x1" value="5.0" scale="4.0" bias="1.0" layer="layer0" />
    <input name="x2" value="3.2" scale="2.0" bias="1.2" layer="layer0" node="1" />
    <input name="x3" value="280.0" scale="240.0" bias="40.0" layer="layer0" node="2" />
    <output name="a" value="0.02441735087912554" scale="0.01" layer="layer2" />
    <output name="b" value="3.28464620721107E-6" scale="0.0010" layer="layer2" node="1" />
    <layer name="layer0" type="identity" size="3" />
    <layer name="layer1" type="sigmoid" size="16">-0.7024246586677668, -2.409046805447259, 0.23012355738953164,
                                                 0.2249875231377209, -2.078737970742437, -0.1319946157343967,
                                                 3.4173863815314673, -4.164846085776751, 1.1511909706065158,
                                                 -4.242652690106668, -4.0275812558277675, 0.280158437519067,
                                                 -0.9845811661702158, 1.5221886259468704, 0.6237528675530384,
                                                 1.95958406176998</layer>
    <layer name="layer2" type="linear" size="2">0.10936682877917034, 0.0020966297697849992</layer>
    <connection srcName="layer0" dstName="layer1">-8.658118395412577, -8.544753017372386, -1.769507244627045,
                                                 -1.6990281004832828, -2.9860919834858137, -1.2321615647339885,
                                                 1.6155136281849811, -12.040383467728239, -7.206391829946805,
                                                 -11.609370029036622, -12.778098337694201, -1.8296727273338973,
                                                 -12.905145098084468, 1.7948186872929583, -1.4525204382591435,
                                                  0.5869313213149651, -1.6880783200218543, -1.6980953425768277,
                                                 -0.6382638712294072, -0.4615598968565889, -4.424203974106881,
                                                 -0.5353655017178139, -4.987221660214916, -3.361361820470712,
                                                  1.343766810648909, -3.4443598512574565, -3.1963840364140643,
                                                 -0.6083429958072337, -0.08718923831641946, -3.6793817121727934,
                                                 -0.46878513297868224, -6.3950742655252455, -1.0331689092183722,
                                                 -1.58726110328377, -0.7122577225604569, -0.8417877875672115,
                                                 -2.0188440467554947, -0.4789805807036303, -0.18525719280291567,
                                                 -3.1882532273439823, 1.5829049997121807, -3.330785237676132,
                                                 -3.079874177987541, -0.5590938141071454, 0.02947338530858677 2,
                                                 -2.3917243550201572, -0.4307827075639602,
                                                 -8.041619829046484</connection>
    <connection srcName="layer1" dstName="layer2">2.672250704968546, 0.06549096815859268, 4.445482378336362,
                                                 1.3426303269851776, 2.479261027514004, -3.910467611362013E-7,
                                                 2.561103210533547, 7.325055065908214E-4, 0.7340469219814676,
                                                 0.05601919537696628, 1.3909877095245216, 7.535793283863449E-4,
                                                 2.1583869856611666, 0.0011355953798053073, 3.7362448805753936,
                                                 0.0022096513932162274, 5.1407259134887715, 0.0016352811921856814,
                                                 3.543161866846364, 2.1092343715805898E-6, 3.8366903582353933,
                                                 0.3131246190914335, 2.4885426350188156, 0.00534596330771736,
                                                 3.3083018004920435, 0.6830934228500399, 3.1063615069631347,
                                                 0.002498127931206410 3, 2.4801527875520004, -3.890745734568386E-7,
                                                 2.1252509977242093, 0.1699301531603487</connection>
  </mapping>
</mapping>
<problem direction="minimize">
  <variable label="y_scraper" useRandomInitial="true" randomSeed="100" randomLower="1.0" randomUpper="5.0"
```

```
                modelVariable="y_scraper">
      <lower>1.0</lower>
      <upper>5.0</upper>
      <value>[5.0, 3, 5, 1]</value>
    </variable>
    <variable label="V_RF" useRandomInitial="true" randomSeed="101" randomLower="1.2" randomUpper="3.2"
                modelVariable="V_RF">
      <lower>1.2</lower>
      <upper>3.2</upper>
      <value>3.2</value>
    </variable>
    <variable label="M_b" useRandomInitial="true" randomSeed="110" randomLower="80" randomUpper="280"
                modelVariable="M_b">
      <lower>80.0</lower>
      <upper>280.0</upper>
      <value>280.0</value>
    </variable>
    <variable label="I_t" useRandomInitial="true" randomSeed="111" randomLower="90" randomUpper="100"
                modelVariable="I_t">
      <lower>90.0</lower>
      <upper>100.0</upper>
      <value>100</value>
    </variable>
    <objective label="obj" modelVariable="obj">
    </objective>
  </problem>
  <solver classname="com.pav.model.hybrid.optimization.snopt.Snopt" gradientMode="1">
    <parameter name='major feasibility tolerance' value='1e-4'/>
    <parameter name='major optimality tolerance' value='1e-4'/>
    <parameter name='hessian full memory' value='true'/>
    <parameter name='major print level' value='1'/>
    <parameter name='print file' value='6'/>
    <parameter name='solution' value='true'/>
    <parameter name='solution file' value='6'/>
    <parameter name='print parameters' value='true'/>
  </solver>
</config>
```

References

[1] L. Chiang, E. Russell, and R. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. Springer, 2001.

[2] J. Jackson, *A User's Guide to Principal Components*. John Wiley & Sons, 1991.

[3] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973.

[4] J. MacGregor, *Statistical Process Control of Multivariate Processes*. Proc. of the IFAC Int. Symp. on Advanced Control of Chemical Processes, pp. 427-535, 1994.

[5] W. Larimore, *Canonical Variate Analysis in Control and Signal Processing*. In Statistical Methods in Control and Signal Processing, pp. 83-120, 1997.

[6] J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, 1998.

[7] G. Lee, B. Lee, E. Yoon, and C. Han, "Multiple-Fault Diagnosis under Uncertain Conditions by the Quantification of Qualitative Relations," *Ind. Eng. Chem. Res.*, vol. 38, pp. 988–998, 1999.

[8] J. Zhao, B.Chen, and J. Shen, "A Hybrid ANN-ES System for Dynamic Fault Diagnosis of Hydrocracking Process," *Computers and Chemical Engineering*, vol. 21, pp. 929–933, 1997.

[9] E. Micheli-Tzanakou, *Supervised and Unsupervised Pattern Recognition*. CRC Press, 1999.

[10] P. Duval, U. Laustroer, and R. Schmitz, "Fault Identification in Accelerator Control," *http://adweb.desy.de/mst/Mst_content/fault_identification.pdf*, 2004.

[11] R. Lauckner, "What Do We See in the Control Room?," *LHC Performance Workshop - Chamonix XII*, pp. 214–217.

[12] E. Malandain and P. Sharek, "An Expert System for Accelerator Fault Diagnosis," *PAC*, 1987.

[13] A. Fisher, "Instrumentation and Diagnostics for PEP-II," *Proc. Beam Instrumentation Workshop*, pp. 95–109, Stanford, CA 1998.

[14] J. Ribes and G. Delaunay, "RBF Neural Net-Based Classifier for the AIRIX Accelerator Fault Diagnosis," *Linac 2000, Proceedings of the XXth International Linac Conference*, August 2000.

[15] J. Biarrotte, M. Novati, P. Pierini, H. Safa, and D. Uriot, "Beam Dynamics Studies for the Fault Tolerance Assessment of the PDS-XADS Linac Design," *Proc. of EPAC 2004, Lucerne, Switzerland*, pp. 1282–1284.

[16] V. Vapnik, *Statistical Learning Theory*. New York: John Wiley, 1998.

[17] J. Suykens and J. Vandewalle, "Recurrent Least Squares Support Vector Machines," *IEEE Trans. on Circuits and Systems*, vol. 47, pp. 1109–1114, July 2000.

[18] J. Suykens, "Nonlinear Modelling and Support Vector Machines," *IEEE Instrumentation and Measurement Technology Conference*, pp. 287–294, May 2001.

[19] W. Noble, "Support Vector Machine Applications in Computational Biology," *Department of Genome Sciences, U. of Washington*, vol. 1, August 2003.

[20] G. Antonelli, F. Caccavale, C. Sansone, and L. Villani, "Fault Diagnosis for AUVs Using Support Vector Machines," *IEEE Int. Conf. on Robotics & Automation*, pp. 4486–4491, 2004.

[21] J. Bard, *Practical Bilevel Optimization: Algorithms and Applications*. Kluwer Academic Publishers, 1998.

[22] S. Samar, D. Gorinevsky, and S. Boyd, "Model Predictive Estimation of Evolving Faults," *Proc. of American Control Conference*, June 2006.

[23] J. Jackson, *Classical Electrodynamics.* New York: John Wiley and Sons, 1975.

[24] K. Kim, "Characteristics of Synchrotron Radiation," *Physics of Particle Accelerators, Eds. M. Month and M. Dienes, AIP Conf. Proc.*, p. 184, 1989.

[25] A. Hofmann, *The Physics of Synchrotron Radiation.* Cambridge University Press, 2004.

[26] S. Samar, D. Gorinevsky, and S. Boyd, "Embedded Estimation of Fault Parameters in an Unmanned Aerial Vehicle," June 2006.

[27] M. Cornacchia, "Requirements and Limitations on Beam Quality in Synchrotron Radiation Sources," *CERN 90-03*, 1990.

[28] C. Bocchetta, "Lifetime and Beam Quality," *Invited Lectures at the CERN Accelerator School on Synchrotron Radiation and Free Electron Lasers*, June 1997.

[29] H. Bruck, *Accelateurs Circulaires de Particules.* Paris: Presses Universitaires de France, 1969.

[30] H. Wiedemann, *Particle Accelerator Physics II.* Berlin: Springer Verlag, 1996.

[31] J. Corbett and et al, "Electron Beam Lifetime in SPEAR3: Measurement and Simulation," these proceedings.

[32] A. Sorensen, "Introduction to Intrabeam Scattering," *CERN Accelerator School, CERN 87-10*, 1987.

[33] J. Suykens, T. V. Gestel, J. Vandewalle, and B. D. Moor, "A Support Vector Machine Formulation to PCA Analysis and Its Kernel Version," *IEEE Trans. on Neural Networks*, vol. 14, pp. 447–450, March 2003.

[34] S. Nash and A. Sofer, *Linear and Nonlinear Prgoramming.* McGraw-Hill, 1996.