Model based compensation of thermal disturbance in a precision linear electromagnetic actuator

Jonathan Hey, Choon Meng Kiew, Guilin Yang, Ricardo Martinez-Botas

***Abstract*—Thermal disturbance is a major source of positioning error in precision positioning systems. Conventional approach of using special materials for construction and sophisticated geometric design based on advanced computer simulation can be costly as well as time consuming to implement. Moreover, dynamic thermal disturbances cannot be effectively compensated for by such methods. The approach presented in this paper uses model estimated position error coupled with sensor measurement as a feedback compensator of the output shaft position. A state space model is used in a Modified Kalman Filter (MKF) to reduce the total number of temperature sensors needed for estimation of the thermally induced position error. A maximum temperature estimation error of 1.8% of the measurement range is recorded. An online parameter estimation method is implemented to 'fine tune' a transfer function model during a calibration stage before compensation. The model based compensation method resulted in a mean unidirectional positioning deviation of $-0.2\mu\mathrm{m}$ and repeatability of $\pm0.7\mu\mathrm{m}$ during a 5 hour continuous operation.**

***Keywords*—Compensation, Identification, Kalman Filter, Thermoelastic**

I. Introduction

THERMAL disturbance is a major source of positioning error in precision positioning systems. Positioning error arises due to the internal or external thermal loading which leads to thermal strain. Material deformation due to temperature changes causes deviation of the output position from the desired position [1]. Such thermal error have been minimized by using construction material with low coefficient of thermal expansion (CTE) and crafted geometric designs to ensure the thermal center is at the output shaft [2]. Some have cooling system which actively maintains a constant operating temperature. However, such approaches can be costly, difficult to implement and not all the dynamic thermal error can be eliminated by these methods [1].

Fig. 1 shows a linear electromagnetic actuator designed for submicron positioning accuracy. Such actuators are common in precision stages which require long range travel but they are susceptible to disturbances [3]. However, the test device uses an electromagnetic driving scheme coupled with flexure-based supporting bearings to improve the position accuracy [4]. Nevertheless, it still suffers from thermal disturbance due to the internal heat generation at the coil during operation. Thermal expansion of the output shaft leads to eventual positioning error. A model based estimation of the disturbance coupled with active output compensation is a cost effective method to minimize the effects of dynamic disturbance [2],[5]. The relationship between temperature increase and the resulting thermal disturbance is determined from experimental data through model identification techniques [6],[7].

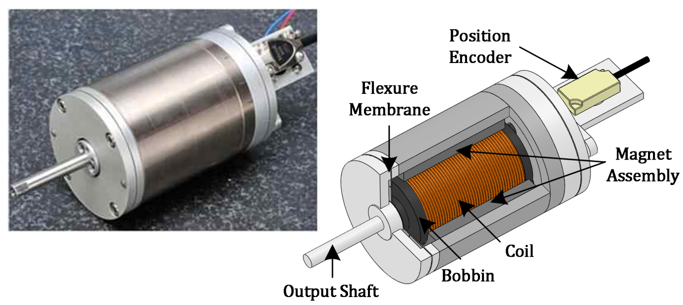


Fig. 1: Flexure based electromagnetic linear actuator

Recent advances in this area of research focuses on systematic reduction of temperature sensors required for modeling [8]. Parallel work on optimal temperature estimate of the rotor coil in an induction machine is shown using Kalman filtering technique [9]. Strategies for fine tuning model estimation accuracy is reported in [7], [10] which shows the trend of online parameter estimation using iterative algorithms. In this paper, a model based compensation method is presented which shows its effectiveness in minimizing thermal disturbance in the test device. The compensation model is experimentally determined from the measured input output data using the subspace and least square method for model parameter estimation.

Internal temperature distribution is estimated using a state space model coupled with measurement feedback from an embedded temperature sensor in the coil. A Modified Kalman Filter (MKF) is presented to show how sensor reduction can be achieved in such an application. The thermo elastic process is modeled using this temperature information as input to a transfer function model. A recursive parameter estimation method is used to tune the transfer function model before compensation is carried out in an extended 5 hour test to simulate continuous operations. The purpose of this work is to show how the position accuracy of the actuator can be improved in real time during continuous operation using a model based compensation method. The aim is to achieve a mean unidirectional position deviation of less than $\pm0.5\mu$m while using only a single embedded sensor in the coil for temperature monitoring.

Manuscript received xxxxx, 2013. This work is supported by the Agency for Science, Technology and Research (A*STAR) of Singapore.

Jonathan Hey and Ricardo Martinez-Botas are with the Mechanical Engineering Department, Imperial College London, Exhibition Road, London, SW7 2AZ, UK (e-mail: hhh09@imperial.ac.uk, r.botas@imperial.ac.uk)

Choon Meng Kiew and Guilin Yang are with Mechatronics group, Singapore Institute of Manufacturing Technology, 71 Nanyang Drive, Singapore, 638075 (e-mail: cmkiew@simtech.a-star.edu.sg, glyang@simtech.a-star.edu.sg)

II. Model identification

A model identification methodology is used to determine the underlying mathematical models of the thermo elastic process. The method involves the selection of the appropriate excitation signal, filtering of measured response, selection of a suitable model structure, model parameter estimation and model validation. The selected model structures are the discrete time transfer function and state space model. The choice of model structure and parameter estimation technique will be further elaborated in the subsequent subsections.

A. Transfer function model

A discrete time Multiple Input Single Output (MISO) transfer function model is used to model the position error arising from temperature changes using the Output Error (OE) parameter estimation technique. The model is derived by combining several Single Input Single Output (SISO) models in parallel. A general SISO model is presented mathematically in (1) where y is the output, u is the input while b,f are the model parameters. n denotes the current time instance while n_b and n_f are the number of past input and output data used respectively for estimating the current output.

$$\begin{aligned} y(n) + f_1 y(n-1) + \cdots + f_n y(n-n_f) \\ = b_1 u(n-1) + \cdots + b_{n_b} u(n-n_b) \end{aligned} \tag{1}$$

A first order MISO model is chosen to model the thermal expansion ($\hat{y}_L$) using discrete temperature points (u_T) as inputs. The MISO model can be expressed mathematically by (2) where m refers to the m^{th} temperature point out of a total of M points.

$$\hat{y}_L(n+1) = f_1 \hat{y}_L(n) + \sum_{m=1}^{M} \sum_{r=0}^{1} b_{m,r} u_T^m (n-r) \tag{2}$$

A total of 12 discrete temperature points are used as input to the model as illustrated in Fig. 2 . Having more sensors and past temperature measurements will enhance the estimation of the thermal expansion. But the first order MISO model is selected for its simplicity as required in a real time application. Nevertheless, it is possible to incorporate a longer time horizon with such a model if extra computation power is available. The sensor positions are selected to capture the changes in thermal gradients in the coil assembly and the surrounding structure due to the outflow of heat from the coil over time.

Equation (2) can be expressed in a more compact form with matrix notations as shown by (3) where φ and θ are the data vector and model parameter respectively.

$$\begin{aligned} &\hat{y}_L(n+1) = \varphi(n)\theta \text{ where} \\ &\theta = [b_{1,1} \cdots b_{M,1} \; b_{1,2} \cdots b_{M,2} \; f_1], \\ &\varphi(n) = \left[u_T^1(n-1) \cdots u_T^M(n-1) \; u_T^1(n) \cdots u_T^M(n) \; \hat{y}_L(n)\right] \end{aligned} \tag{3}$$

$$e(n) = y_L(n) - \hat{y}_L(n) = y_L(n) - \varphi(n-1)\theta \tag{4}$$

Equation (4) defines the output error (e) as the difference between the measured output (y) and model output ($\hat{y}$). A least square algorithm is used to determine θ that minimizes e for a set of data vector - $\{\varphi(1), \cdots, \varphi(n), \cdots, \varphi(N)\}$ where

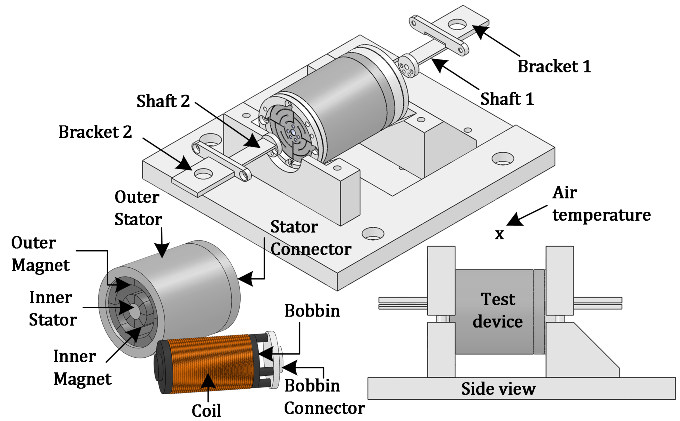


Fig. 2: Temperature points of interests on test device

N is the total number of sampled instances. The parameter estimation method is known as the Output Error (OE) method simply because it uses the output error as a measure of the residual remaining from the parameter estimation process.

A block least square algorithm makes use of the whole set of data to estimate the model parameter in a single mathematical operation. The appended data matrix ($\mathbf{\Phi}$) is formed by stacking φ row wise for each sampled instance. To ensure a good estimate of the model parameters, the experiment duration should be sufficiently long (in the order of hours) while the sampling rate is maximized (in order of ms) to reduce the loss of information and improve tracking accuracy. The dimension of the data matrix is such that $\mathbf{\Phi} \in \mathbb{R}^{(2M+1)\times(N)}$. The solution of a block least square problem requires the inversion of $\mathbf{\Phi}$ which is a costly mathematical operation.

A recursive algorithm is chosen because the solution is calculated iteratively in small steps and uses an efficient mathematical substitution to eliminate the need for matrix inversion. The recursive least square algorithm is summarized here by (5) and it is is derived from the iterative solution of the block least square algorithm. The solution is proven to be equivalent to that of the block least square [11].

$$\begin{aligned} &\mathbf{Q}(n) = \frac{1}{\lambda(n)} \left[\mathbf{Q}(n-1) - \frac{\mathbf{Q}(n-1)\varphi^T(n)\varphi(n)\mathbf{Q}(n-1)}{\lambda(n) + \varphi(n)\mathbf{Q}(n-1)\varphi^T(n)} \right] \\ &\mathbf{L}(n) = \frac{\mathbf{Q}(n-1)\varphi^T(n)}{\lambda(n) + \varphi(n)\mathbf{Q}(n-1)\varphi^T(n)} \\ &\theta(n) = \theta(n-1) + \mathbf{L}(n)\left[y_L(n) - \varphi(n)\theta(n-1)\right] \end{aligned} \tag{5}$$

It should be noted that an iterative solution would require an initial estimate of the variables. $\mathbf{Q}$ is a measure of the confidence in the estimated model parameter θ. A sufficiently large $\mathbf{Q}(0)$ should be chosen to move the initial estimate $\theta(0)$ towards its final solution [11]. λ that appears in (5) is a weight factor for adjustment of the relative importance of the time data used for model parameter estimation. However, such fine tuning strategies are not exploited in the current work.

A recursive least square method have some practical benefits such as offering the flexibility of online model tuning which is suitable for applications which require on-site set up calibration of the system. The user also has the flexibility of choosing the calibration duration based on the maximum time or level of residual error allowed before terminating the process. However, the method does require that computation be completed within one sampling interval. This is possible with simple models like the first order MISO transfer function model or when computation power is not a limitation.

B. State space model

Temperature changes in the system needs to be monitored to estimate the thermal expansion as discussed in the preceding section. The device is constructed with very high mechanical tolerance to enhance electromagnetic performance. There is little access to the internal components such as the internal magnet assembly and bobbin (refer to Fig. 1). A state space model is selected to model the dynamic temperature response of the device at 12 points (refer to Fig. 2). Measurement from a single embedded sensor in the coil is used as feedback in a Modified Kalman Filter for output adjustment. The details of the method will be presented in the next subsection.

A state space model is determined from the experimentally measured input-output data based on a combined stochastic-deterministic subspace method. A brief overview of the method from an end users perspective is outlined in the next section for completeness in presentation of the overall methodology. It is not within the scope of this paper to discuss the formulation details which has been well documented in the pioneering work [12],[13]. The key steps outline in the next section are based on work reported in [14].

It should be pointed out that the temperature used for model identification refers to the temperature change from the ambient. Substituting the absolute temperature with this relative temperature will improve the modeling accuracy since ambient temperature affects the overall heat loss of the device. The choice of input signal is important to ensure the proper identification of the model parameters based on the underlying causal relationship between the input and temperature rise. Eddy current, magnet hysteresis and mechanical losses which are common to electromagnetic actuators are significant only when there is large or high speed motion [15],[16].

For this test device the conversion losses come mainly from the resistive heating of the copper wires in the coil. The resistive heating is derived from the electrical power supplied which is given by the product of the supplied voltage (V) and current drawn (I) from the source and it is used as input to the state space model. The choice of input waveform and its frequency content is also important for the proper identification of the model which would be discussed in section V.

$$\begin{aligned}\hat{\mathbf{x}}(n+1) &= \mathbf{A}\hat{\mathbf{x}}(n) + \mathbf{B}u(n) \\ &\quad + \mathbf{K}(n)\left[\mathbf{y}(n) - \mathbf{C}\hat{\mathbf{x}}(n) - \mathbf{D}u(n)\right] \\ \hat{\mathbf{y}}(n) &= \mathbf{C}\hat{\mathbf{x}}(n) + \mathbf{D}u(n) + \left[\mathbf{y}(n) - \mathbf{C}\hat{\mathbf{x}}(n) - \mathbf{D}u(n)\right]\end{aligned} \tag{6}$$

A standard state space representation is shown in (6) where $\hat{\mathbf{y}}$ is the output and u is the input while $\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$ and $\mathbf{D}$ are the state matrices. The innovation term, $(\mathbf{y} - \mathbf{C}\hat{\mathbf{x}} - \mathbf{D}u)$, accounts for the stochastic or noise component of the measured output (y). $\mathbf{K}$ is the Kalman filter gain whose value is dependent on the noise content of the output. The state matrices and Kalman gain are determined using the combined stochastic-deterministic subspace method.

1) Combined stochastic-deterministic state space model identification using the subspace method: The subspace method makes use of the geometric projection methods such as the orthogonal and oblique projections of future and past input-output data sequence to extract the state matrices ($\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$,$\mathbf{D}$) and Kalman gain ($\mathbf{K}$). The method is based on these fundamental geometric operations and they are defined in (7)-(9) for completeness. The orthogonal projection of any matrix $\mathbf{Y}$ onto another matrix $\mathbf{U}$ is defined by (7). The projection onto its orthogonal complement, denoted with a superscript '$\perp$', is defined by (8). The oblique projection of a matrix $\mathbf{Y}$ onto matrix $\mathbf{W}$ along the row space of matrix $\mathbf{U}$ is given by (9) where the superscript '$\dagger$' denotes a pseudo inverse.

$$\mathbf{Y}/\mathbf{U} = \mathbf{Y}\mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1}\mathbf{U} \tag{7}$$

$$\mathbf{Y}/\mathbf{U}^{\perp} = \mathbf{Y}[\mathbf{I} - \mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1}\mathbf{U}] \tag{8}$$

$$\mathbf{Y}_U/\mathbf{W} = [\mathbf{Y}/\mathbf{U}^{\perp}] \cdot [\mathbf{W}/\mathbf{U}^{\perp}]^{\dagger} \cdot \mathbf{W} \tag{9}$$

$$\left[\frac{\mathbf{U}_p}{\mathbf{U}_f}\right] = \left[\begin{array}{ccc} u(0) & \cdots & u(N-1) \\ u(1) & \cdots & u(N) \\ \vdots & \vdots & \vdots \\ u(i-1) & \cdots & u(i+N-2) \\ \hline u(i) & \cdots & u(i+N-1) \\ u(i+1) & \cdots & u(i+N) \\ \vdots & \vdots & \vdots \\ u(2i-1) & \cdots & u(2i+N-2) \end{array}\right] \tag{10}$$

The input and output data are organized into block matrices as shown by an example of the input block Hankel matrix in (10). i is the number of block rows of the data used for analysis and it is user selectable but should be greater than the system order. It is composed of two parts which are denoted by the subscript 'p' and 'f' to represent past and future data respectively. Past data includes data up to the instance $(i-1)$ before while future data include data up to the instance $(2i-1)$ after as illustrated in (10). The column width of the block matrix is determined by the total length of the data sequence denoted by N. A superscript of '+' and '-' represents a shift into the future or past respectively by one block row as illustrated in (11).

$$\left[\frac{\mathbf{U}_p^+}{\mathbf{U}_f^-}\right] = \left[\begin{array}{ccc} u(0) & \cdots & u(N-1) \\ u(1) & \cdots & u(N) \\ \vdots & \vdots & \vdots \\ u(i-1) & \cdots & u(i+N-2) \\ u(i) & \cdots & u(i+N-1) \\ \hline u(i+1) & \cdots & u(i+N) \\ \vdots & \vdots & \vdots \\ u(2i-1) & \cdots & u(2i+N-2) \end{array}\right] \tag{11}$$

The input block matrices, $\mathbf{Y}_f$,$\mathbf{Y}_p$,$\mathbf{Y}_p^+$,$\mathbf{Y}_f^-$ are defined in a similar manner. A block Hankel matrix is a combination of the input and output data sequence in one matrix. It also have the same notations as the input and output block matrices as illustrated by the forward shifted block Hankel matrix in (12).

$$\mathbf{W}_p^+ = \begin{bmatrix} \mathbf{U}_p^+ \\ \mathbf{Y}_p^+ \end{bmatrix} \tag{12}$$

The system matrices are arranged into block matrices. They are defined as the extended observability ($\mathbf{\Gamma}_i$), lower block triangular Toeplitz ($\mathbf{H}_i$) and the reversed extended controllability (Δ_i) matrix which are given by (13), (14) and (15) respectively.

$$\mathbf{\Gamma}_i = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} \tag{13}$$

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{D} & 0 & \cdots & 0 & 0 \\ \mathbf{CB} & \mathbf{D} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{CA}^{i-2} & \mathbf{CA}^{i-3}\mathbf{B} & \cdots & \mathbf{CB} & \mathbf{D} \end{bmatrix} \tag{14}$$

$$\Delta_i = \begin{bmatrix} \mathbf{A}^{i-1}\mathbf{B} & \mathbf{A}^{i-2}\mathbf{B} & \cdots & \mathbf{B} \end{bmatrix} \tag{15}$$

With the reorganized data and system matrices, the input-state-output relations can be expressed in compact form given by (16)-(17) where the superscript 'd' denotes the deterministic part of the data while 's' represents the stochastic component. $\mathbf{A}_i$ is the appended state matrix which appears in (17).

$$\mathbf{Y}_f = \mathbf{\Gamma}_i \mathbf{X}_f^d + \mathbf{H}_i^d \mathbf{U}_f + \mathbf{Y}_f^s \tag{16}$$

$$\mathbf{X}_f^d = \mathbf{A}_i \mathbf{X}_p^d + \mathbf{\Delta}_i^d \mathbf{U}_p \tag{17}$$

The identification process starts by taking the oblique projection of the output block matrix onto the block Hankel matrix along the row space of the input block matrix as defined in (18). The resulting expression, $\mathcal{O}_\mathbf{i}$, contains the extended observability matrix which is extracted using a Singular Value Decomposition (SVD) of the weighted expression in (19). The decomposed data is represented in the standard notation where $\mathbf{S_1}$ is the first n non zero singular value which serves as a guide for selecting the model order. The selection of the weighting matrices ($\mathbf{W_1}$,$\mathbf{W_2}$) is dependent on the algorithm selected by the user. The relative merits of the different algorithm is discussed in detail in [10]. For the current application and purpose, it is selected that $\mathbf{W_1} = \mathbf{W_2} = \mathbf{I}$ based on the N4SID algorithm [14]. The extended observability matrix is derived from the SVD process based on the definition given by (20).

$$\mathcal{O}_i = \mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p \tag{18}$$

$$\mathbf{W}_1 \mathcal{O}_i \mathbf{W}_2 = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} \tag{19}$$

$$\mathbf{\Gamma}_i = \mathbf{W}_1^{-1} \mathbf{U}_1 \mathbf{S}_1^{1/2} \tag{20}$$

From the definition in (13), the state matrix ($\mathbf{C}$) is given by the top block row of the extended observability matrix ($\mathbf{\Gamma}_i$). The system matrix ($\mathbf{A}$) is related to it by the relationship $\underline{\mathbf{\Gamma}_i}\mathbf{A} = \overline{\mathbf{\Gamma}_i}$ where $\overline{\mathbf{\Gamma}_i}$ and $\underline{\mathbf{\Gamma}_i}$ is the extended observability matrix with the top and bottom block row removed respectively. Thus, the state matrix ($\mathbf{A}$) is simply given by $\mathbf{A} = \underline{\mathbf{\Gamma}_i}^\dagger \overline{\mathbf{\Gamma}_i}$.

The sequences $\mathcal{Z}_{i+1}$ and $\mathcal{Z}_i$ are defined by a set of oblique projections shown in (21). They are related to the Kalman filtered states $\hat{\mathbf{X}}_{i+1}$ and $\hat{\mathbf{X}}_i$by (22).

$$\mathcal{Z}_{i+1} = \mathbf{Y}_f^- / \begin{bmatrix} \mathbf{W}_p^+ \\ \mathbf{U}_f^- \end{bmatrix}, \quad \mathcal{Z}_i = \mathbf{Y}_f / \begin{bmatrix} \mathbf{W}_p \\ \mathbf{U}_f \end{bmatrix} \tag{21}$$

$$\mathcal{Z}_{i+1} = \mathbf{\Gamma}_i \hat{\mathbf{X}}_{i+1} + \mathbf{H}_{i-1}^d \mathbf{U}_f^-, \quad \mathcal{Z}_i = \mathbf{\Gamma}_i \hat{\mathbf{X}}_i + \mathbf{H}_i^d \mathbf{U}_f \tag{22}$$

The stochastic component of the output is assumed to be uncorrelated to the Kalman filtered states. As such the state equations first introduced in (6) can be written in block form given by (23) where $\mathbf{R}_w$ and $\mathbf{R}_v$ matrices are the stochastic component of the output. Equation (24) is the state equations expressed in terms of $\mathbf{\Gamma}_i$, $\mathbf{A}$, $\mathbf{C}$ and $\mathbf{K}$ by using the definitions (22) to substitute for the states in (23). $\mathbf{K}$ is defined as the Kalman gain sequence as shown in (25).

$$\begin{bmatrix} \hat{\mathbf{X}}_{i+1} \\ \mathbf{Y}_i \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{X}}_i \\ \mathbf{U}_i \end{bmatrix} + \begin{bmatrix} \mathbf{R}_w \\ \mathbf{R}_v \end{bmatrix} \tag{23}$$

$$\begin{bmatrix} \Gamma_{i-1}^\dagger \cdot \mathcal{Z}_{i+1} \\ \mathbf{Y}_i \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} \cdot \Gamma_i^\dagger \cdot \mathcal{Z}_i + \mathcal{K} \cdot \mathbf{U}_f + \begin{bmatrix} \mathbf{R}_w \\ \mathbf{R}_v \end{bmatrix} \tag{24}$$

$$\mathcal{K} = \begin{bmatrix} \left(\mathbf{B} | \Gamma_{i-1}^\dagger \cdot \mathbf{H}_{i-1}^d\right) - \mathbf{A} \cdot \Gamma_i^\dagger \cdot \mathbf{H}_i^d \\ (\mathbf{D}|\mathbf{0}) - \mathbf{C} \cdot \Gamma_i^\dagger \cdot \mathbf{H}_i^d \end{bmatrix} \tag{25}$$

$\mathbf{\Gamma}_i$, $\mathbf{A}$ and $\mathbf{C}$ are already determined from the earlier analysis which leaves the Kalman gain sequence to be determined. $\mathcal{K}$ is obtained from a least square regression of (24). This is possible since $\mathbf{R}_w$ and $\mathbf{R}_v$ are assumed to be stochastic noise sequence uncorrelated to $\mathcal{Z}_i$ and $\mathbf{U}_f$. Thus, $\mathbf{R}_w$ and $\mathbf{R}_v$ are the residual errors from the least square regression process.

$$\begin{bmatrix} \mathbf{K}_{1|1} \\ \vdots \\ \mathbf{K}_{1|i} \\ \hline \mathbf{K}_{2|1} \\ \vdots \\ \mathbf{K}_{2|i} \end{bmatrix} = \mathbf{N} \begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix} \tag{26}$$

Equation (26) shows $\mathcal{K}$ as a linear combination of $\mathbf{B}$ and $\mathbf{D}$ where $\mathbf{N}$ is a function of the already known $\mathbf{\Gamma}_i$, $\mathbf{A}$ and $\mathbf{C}$. The final step to determine $\mathbf{B}$ and $\mathbf{D}$ uses the least squares regression on (26). The preceding discussion gives the general description of the steps involved to determine the state matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ as first defined in (6) using the combined stochastic-deterministic subspace method. A solution is guaranteed since only linear algebraic operations such as the least square algorithm is used to solve the inverse mathematical problem. Subspace algorithms are non-iterative which means there are no convergence issues. The method is robust and efficient which makes it suitable for large data sets.

2) Modified Kalman Filter: In order to obtain an optimal temperature estimate, sensor feedback as applied in a Kalman Filter (KF) is used to minimize the state space model output error. The output error first presented in section II is reintroduced in (27) with a similar definition where y and $\hat{y}$ are the measured and model output respectively. An additional term, $\bar{y}$, is the adjusted output from applying the KF. Quantities derived from the adjusted output are termed posterior while those which are not are termed priori and indicated by a superscript '-'. The error and error covariance are defined by (27)-(28).

$$\mathbf{e}^- = \mathbf{y}(n) - \hat{\mathbf{y}}(n),\ \mathbf{e}(n) = \mathbf{y}(n) - \bar{\mathbf{y}}(n) \tag{27}$$

$$\mathbf{P}^-(n) = E\left\{\left[\mathbf{e}^-(n)\right]\left[\mathbf{e}^-(n)\right]^T\right\}, \mathbf{P}(n) = E\left\{[\mathbf{e}(n)][\mathbf{e}(n)]^T\right\} \tag{28}$$

Temperature is estimated from the state space model during the 'time' update. During the 'measurement' update a weighted correction of the output is performed based on the measured output error and the Kalman gain ($\mathbf{K}$) as illustrated in Fig. 3.

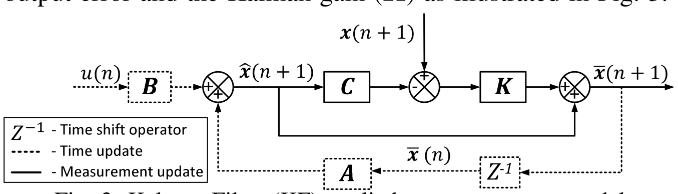


Fig. 3: Kalman Filter (KF) applied to a state space model

'Time update'

$$\begin{aligned}\hat{\mathbf{x}}(\mathbf{n+1}) &= \mathbf{A}\bar{\mathbf{x}}(n) + \mathbf{B}u(n)\\ \mathbf{P}^-(n+1) &= \mathbf{A}\mathbf{P}(n)\mathbf{A}^T + \mathbf{R}_w\end{aligned} \tag{29}$$

'Measurement update'

$$\begin{aligned}&\mathbf{K}(n+1) = \mathbf{P}^-(n+1)\mathbf{C}^T\left[\mathbf{C}\mathbf{P}^-(n+1)\mathbf{C}^T + \mathbf{R}_v\right]^{-1}\\ &\mathbf{P}(n+1) = \mathbf{P}^-(n+1) - \mathbf{K}(n+1)\mathbf{C}\mathbf{P}^-(n+1)\\ &\bar{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n+1)\\ &\qquad\qquad + \mathbf{K}(n+1)\left[\mathbf{y}(n+1) - \mathbf{C}\hat{\mathbf{x}}(n+1)\right]\end{aligned} \tag{30}$$

The mathematical operations during the 'time' and 'measurement' update are shown in (29)-(30) [9]. The amount of output correction is determined by the Kalman gain ($\mathbf{K}$). The value of $\mathbf{K}$ is derived iteratively based on a recursive least square algorithm which minimizes $\mathbf{P}$ over time. $\mathbf{R}_w$, is added to $\mathbf{P}$ as shown in (29) at each iteration. It is represents the model output error during the 'time' update. Similarly, $\mathbf{R}_v$, is added to $\mathbf{P}$ during the 'measurement' update which reflects the amount of measurement uncertainty. $\mathbf{K}$ and $\mathbf{R}_v$ have an inverse relationship as shown by (30). Thus, output correction will be less when the measurement uncertainty is high. $\mathbf{K}$, $\mathbf{R}_w$ and $\mathbf{R}_v$ was first introduced in the preceding section. The use of the same notation indicates that they are the same variable. The difference lies in way it is calculated; direct or iterative.

The state space model is represented in (31)-(33) with altered notations to illustrate the physical quantities it represents. $\hat{\mathbf{y}}_T$ is the model estimated temperature while u_q is the input heat generation to the system and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ are the state matrices defined in (6). $\mathbf{D}$ is set to be a null matrix during the identification process since the output is not directly correlated with the input. Equation (32) simplifies to give $\hat{\mathbf{x}} = \mathbf{C}^{-1}\hat{\mathbf{y}}_T$ which directly relates the internal states to the output. Substitution of this relation into (31) leads to the transformed state equations (33).

$$\hat{\mathbf{x}}(n+1) = \mathbf{A}\hat{\mathbf{x}}(n) + \mathbf{B}u_q(n) \tag{31}$$

$$\hat{\mathbf{y}}_T(n) = \mathbf{C}\hat{\mathbf{x}}(n) + \mathbf{D}u_q(n) \tag{32}$$

$$\hat{\mathbf{y}}(n+1) = \mathbf{C}\mathbf{A}\mathbf{C}^{-1}\hat{\mathbf{y}}_T(n) + \mathbf{C}\mathbf{B}u_q(n) \tag{33}$$

The KF is presented earlier to highlight the changes made in the Modified Kalman Filter (MKF) as illustrated by Fig. 4 and represented by (34)-(35) with altered notation for differentiation. The modification of the KF is aimed at reducing the number of sensor measurement needed for output correction. The following is a summary of the modifications made.

1) Replace the state ($\hat{\mathbf{x}}$) with the output ($\hat{\mathbf{y}}_T$) by substitution of the original state matrices ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$) with the transformed state matrices ($\alpha = \mathbf{C}\mathbf{A}\mathbf{C}^{-1}$, $\beta = \mathbf{C}\mathbf{B}$)
2) Output correction using an output mapping (ζ_σ) of the measured error signal

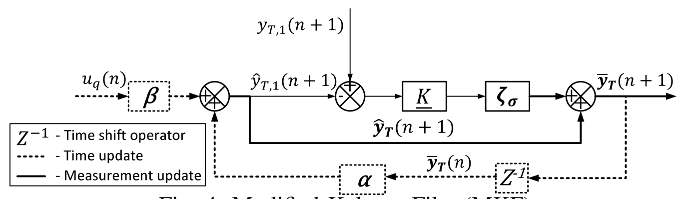


Fig. 4: Modified Kalman Filter(MKF)

'Time update'

$$\begin{aligned}\hat{\mathbf{y}}_T(n+1) &= \alpha\bar{\mathbf{y}}_T(n) + \beta u_q(n)\\ \underline{P}^-(n+1) &= \underline{\alpha}\underline{P}(n)\underline{\alpha}^T + R_w\end{aligned} \tag{34}$$

'Measurement update'

$$\begin{aligned}&\underline{K}(n+1) = \underline{P}^-(n+1)\left[\underline{P}^-(n+1) + R_v\right]^{-1}\\ &\underline{P}(n+1) = \underline{P}^-(n+1)\left[1 - \underline{K}(n)\right]\\ &\bar{\mathbf{y}}_T(n+1) = \hat{\mathbf{y}}_T(n+1)\\ &\qquad\qquad + \underline{K}(n+1)\left[y_{T,1}(n+1) - \hat{y}_{T,1}(n+1)\right]\zeta_\sigma\end{aligned} \tag{35}$$

The output mapping (ζ_σ) shown in (35) is derived from the covariance between the error signal measured at coil to the other temperature points as defined in (36). The covariance for sequences of data ($\mathbf{z}_1, \cdots, \mathbf{z}_i, \cdots, \mathbf{z}_M$) can be calculated using (37) where μ is the mean of the each data sequence. The output error sequence is obtained from the results of the model identification test which will be presented in section V.

$$\zeta_\sigma = \begin{bmatrix} 1 & \cdots & \frac{\sigma_j}{\sigma_1} & \cdots & \frac{\sigma_M}{\sigma_1}\end{bmatrix}^T \tag{36}$$

$$\sigma_j = \frac{1}{N}\sum_{i=1}^{N}\left[z_1(i) - \mu_1\right]\left[z_j(i) - \mu_j\right] \tag{37}$$

In the original KF, the state matrix ($\mathbf{A}$) in (29) serves as a weighting factor for error covariance updating. In the MKF, the Kalman gain ($\underline{K}$) and error covariance ($\underline{P}$) are calculated based on a single measured error signal. Equation (34) reflects the reduced dimension in which $\underline{\alpha}$ is obtained from the first row of the transformed state matrix. In the MKF, $\mathbf{R}_v$ corresponds to the thermocouple measurement uncertainty. $\mathbf{R}_w$ is estimated from the measurement uncertainty and the residual error resulting from the subspace identification process. The measurement uncertainty and residual error will be presented in section IV and V respectively.

Subsequent application of the filter would require only the measurement feedback from the embedded temperature sensor at the coil. Application of the MKF will reduce the total number of sensors needed for estimation of the thermal expansion. The assumption is that unmeasured error signals are correlated based on a statistical measure, given by the covariance of the output error predetermined during a calibration test before the actual filter application. The improvement in estimation accuracy will be discussed in section V.

III. Compensation model and position controller

The compensation model is used as an estimate of the position error based on the measured heat generation (u_q) and coil temperature ($\bar{y}_{T,1}$). The model can be described entirely by (38)-(39). $\bar{\mathbf{y}}_T$ refers to the adjusted temperature estimate after Kalman filtering. The estimated thermal expansion ($\hat{y}_L$) is simply a product of the adjusted temperature estimate against the model parameter (θ) shown in (39). The model is illustrated in a block diagram in Fig. 5.

$$\begin{aligned}\bar{\mathbf{y}}(n+1) &= \alpha\bar{\mathbf{y}}_T(n) + \beta u_q(n) \\ &+ \underline{K}(n)\left[y_{T,1}(n+1) - \hat{y}_{T,1}(n+1)\right]\zeta_\sigma\end{aligned} \tag{38}$$

$$\begin{aligned}&\hat{y}_L(n+1) = \varphi(n)\theta \text{ where} \\ &\theta = \left[b_{1,1}\cdots b_{M,1}\ b_{1,2}\cdots b_{M,2}\ f_1\right], \\ &\varphi(n) = \left[\bar{y}_T^1(n-1)\cdots\bar{y}_T^M(n-1)\,\bar{y}_T^1(n)\cdots\bar{y}_T^M(n)\,\hat{y}_L(n)\right]\end{aligned} \tag{39}$$

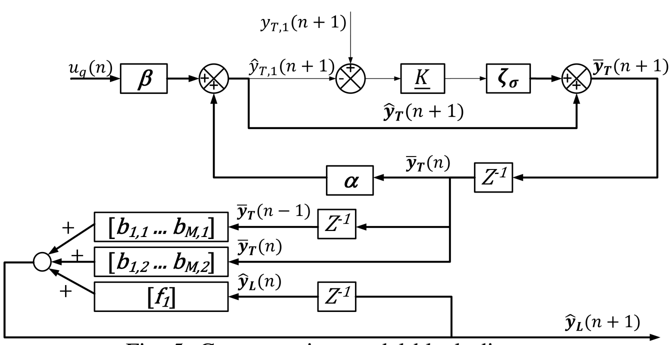


Fig. 5: Compensation model block diagram

The actuator can be represented by a series combination of an electrical and mechanical subsystem as illustrated in Fig. 6. The electrical subsystem is a series connection of a resistor and inductor where R and L are the resistance and inductance respectively. The mechanics can be modeled as a mass in series with springs and a damper. The moving mass (m) is made up of the coil and bobbin mass while the spring force is provided by the flexure mechanism with a spring constant k. Damping is caused by eddy current induced in the bobbin-coil assembly with a damping coefficient of b. K_a is the amplifier gain while k_T is the force constant which is dependent on strenght of the coupling magnetic field and the number of turns of the coil.

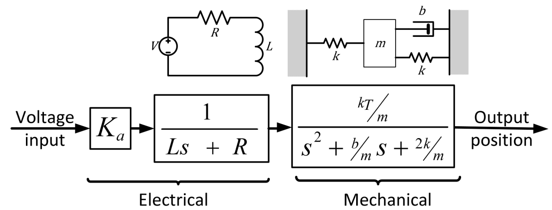


Fig. 6: Actuator electromechanical model

The electro-mechanical and thermo elastic interaction in the device is illustrated in Fig. 7 by the system block diagram. The uncompensated position is measured by an optical position encoder. The compensation model estimated thermal expansion is combined with the position measurement to give the compensated position. Positioning of the output shaft is controlled with a PID controller using the compensated position as a feedback. A second position encoder is used to measure the reference output position for performance characterization. The position error is defined as the difference between this reference output position and the target position.

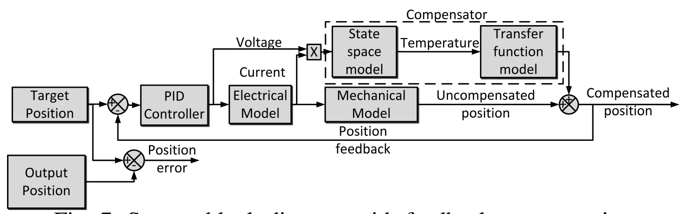


Fig. 7: System block diagram with feedback compensation

All the quantities indicated in the block diagram are measured or derived from measurements except for the compensated position which is obtained from model estimation. Since the dynamic response of the actuator is measured, an empirical PID controller tuning method is used to determine the controller gains. The method is based on the Ziegler-Nichols standard tuning approach by analyzing the open loop step response of the system. The continuous form of a standard PID controller is given by (40) where ε is defined as the feedback error signal and υ is the output corrective action. K_p is defined as the proportional gain while T_i and T_d are the integral and derivative time respectively.

$$D(s) = \frac{\upsilon(s)}{\varepsilon(s)} = K_p\left(1 + \frac{1}{T_i s} + T_d s\right) \tag{40}$$

Fig. 8 shows the open loop step response of the system. The first step towards a tuned controller starts with analyzing the open loop step response of the system. The three parameters obtained from the output response are termed the delay time (τ_d), time constant (τ_c) and output gain (K_0) as illustrated in Fig. 8. The controller gains are estimated from these three parameters based on the relationship presented in Table I. A discrete version of the PID controller is necessary for implementation on the actual system. Equation (41) describes the general discrete form where, $T_S = 1\text{ms}$, is the sampling time of the system [17].

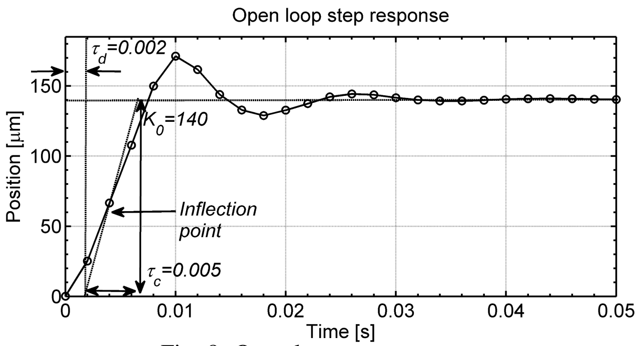


Fig. 8: Open loop step response

TABLE I: Controller parameters

	Proportional gain	Integral time	Derivative time
	$K_p = 1.2\tau_c/(K_0\tau_d)$	$T_i = 2\tau_d$	$T_d = 0.5\tau_d$
Inital estimate	0.021	0.004	0.001

$$v(n) = v(n-1) + K_P\left[\left(1+\frac{T_s}{T_i}+\frac{T_d}{T_s}\right)\varepsilon(n)-\left(1+\frac{2T_d}{T_s}\right)\varepsilon(n-1)+\frac{T_d}{T_s}\varepsilon(n-2)\right] \tag{41}$$

The controller gains are adjusted based on the initial estimates using the following guidelines; (i) proportional gain is used to decrease the rise time (ii) integral gain eliminates steady state error and (iii) derivative gain reduces overshoot and settling time. The close loop system incorporated with the compensation model has altered dynamics. Further tuning would be necessary, until a satisfactory dynamic performance is achieved. A speed limiter sets the maximum actuation speed to 1mm/min to prevent large step changes and instability.

IV. Measurement and Experimental Test Cases

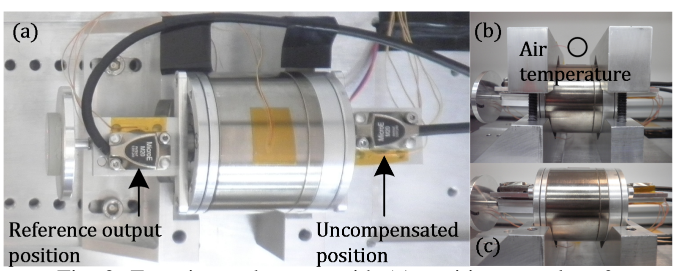


Fig. 9: Experimental set up with (a) position encoders & thermocouples (b) air temperature measurement and (c) side view

The experimental set up for device testing is shown together with the measurement system in Fig. 9. The optical linear position encoders and K type thermocouples can be seen in Fig. 9(a). The air temperature is measured at the point shown in Fig. 9(b) and (c) shows the side view of the device. The measurement uncertainties associated with the position and temperature measurement are summarized in Table II.

TABLE II: Measurement accuracy and resolution

Instrument	Measured quantity	Accuracy	Resolution
MicroE Mercury 3500 optical linear encoder	Position	$\pm 1\mu$m (long range) $\pm 0.15\mu$m (short range)	± 5nm
K type thermocouple	Temperature	$\pm 1.0°C$	$\pm 0.25°C$

The optical encoders have a built in digital interpolator which gives position measurement of up to 5nm in resolution. However, the uncertainty of the measurement typically amounts to 1% of grating size (20μm) of the scale for short range measurement. Moving average filters are used to remove some of the unwanted noise due to digitization. A total of four tests were conducted for model identification and validation. The control parameter, input signal, measured and derived quantities together with the objective of each test are shown in Table III.

TABLE III: Tests conducted for model identification and validation

No	Input signal	Derived quantity	Measured quantity	Objective	Control parameter
1	Step	Electrical power	Voltage, Current, Temperature	Thermal time constant determination	Current
2	Varying pulse width	Electrical power	Voltage, Current, Temperature	State space model identification, Kalman filter parameter determination	Current
3	Varying period trapezoidal	Electrical power	Voltage, Current, Temperature, Position	State space model validation	Position
4A	Varying period trapezoidal	Electrical power, Position error	Voltage, Current, Temperature, Position	Transfer function model identification	Position
4B	Fixed period trape-zoidal	Electrical power, Position error	Voltage, Current, Temperature, Position	Compensation model validation	Position

V. Results and Discussion

Test 1 is a step input of electrical power (3.2W) to obtain the temperature response and shortest thermal time constant of the system. This step response analysis serves as a basis for selecting the input signal for model identification in test 2. Fig. 10 shows the coil and bracket temperature which are the nearest and furthest points from the heat source respectively.

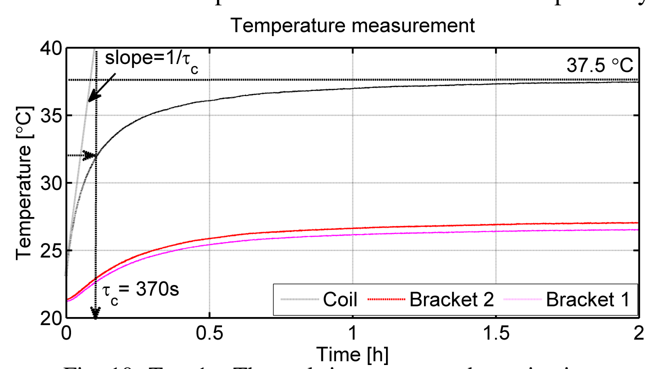


Fig. 10: Test 1 - Thermal time constant determination

The temperature response resembles a first order system which is characterized by an initial gradient of $1/\tau_c$ and a zero gradient approaching the steady state output. τ_c is the time constant defined as the time taken for the output to reach 63.2% of the final steady state value as illustrated in Fig. 10. The system response will be attenuated beyond the cut off frequency which is defined as $f_c = 1/2\pi\tau_c$.

Therefore, to enhance the Signal to Noise Ratio (SNR) of the measured output for the thermal model identification in test 2, the input waveform is designed based on a frequency domain analysis. A varying pulse width input signal is chosen because of its flexibility in design. The frequency content of the signal is controlled by changing the duration of the pulses. A total of 3 pulses of varying period is used for input excitation during the 'heat up' period followed by a period of zero input to allow the system to 'cool down' as shown in Fig. 11(a). The device undergo a similar 'cool down' period during actual operations. The heat loss at the natural convective boundary undergo changes between these two periods. Thus, it would lead to a more robust model if they are accounted for during the identification test.

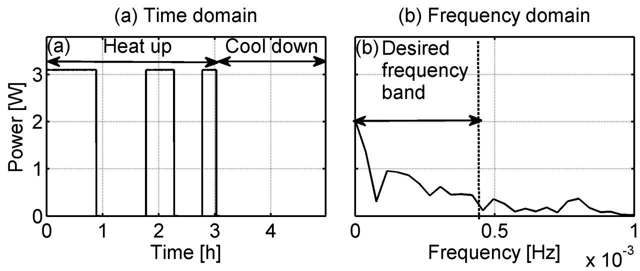


Fig. 11: Test 2 - State space model identification: input waveform in (a) time and (b) frequency domain

The input power spectrum is expected to be low due to the long period of zero input. It is critical to engineer the input signal to enhance its power content in the desired frequency band. The shortest time constant recorded, $\tau_c = 370$s, corresponds to a cut off frequency of $f_c = 4.3 \times 10^{-4}$Hz which gives an upper limit estimate of the device thermal response bandwidth. The selected waveform has an increasing power spectrum in the desired frequency band as shown in Fig. 11(b). This input signal is used for thermal model identifcation in test 2. Fig. 12(a)-(b) shows the measured temperature response and corresponding output error plot is shown in Fig. 12(d)-(f). This error sequence is used to determine the Modified Kalman Filter (MKF) parameter.

$$\langle \mathbf{e} \rangle = \begin{bmatrix} e_1 & \cdots & e_m & \cdots & e_{12} \end{bmatrix}^T$$
$$\text{where } e_m = \sqrt{\frac{\sum_{n=1}^{N} \left[y_{T,M}(n) - \bar{y}_{T,M}(n) \right]^2}{N}} \tag{42}$$

$\langle \mathbf{e} \rangle$ is a measure of the output error averaged over the whole measurement period as defined by (42). $\langle \mathbf{e} \rangle$ is termed the residual error in this instance since it is calculated from the output error of this identification test. The estimated coil temperature has a residual error of $0.46°C$ which is used for estimating the filter parameter, R_w, as defined in (34). The output mapping defined as $\zeta_\sigma = \left[1 \cdots \frac{\sigma_j}{\sigma_1} \cdots \frac{\sigma_{12}}{\sigma_1} \right]^T$ is calculated using (37) and tabulated in Table IV.

TABLE IV: Output mapping (ζ_σ)

σ_1	σ_2/σ_1	σ_3/σ_1	σ_4/σ_1	σ_5/σ_1	σ_6/σ_1
0.155	0.763	0.729	0.495	0.353	0.264
σ_7/σ_1	σ_8/σ_1	σ_9/σ_1	σ_{10}/σ_1	σ_{11}/σ_1	σ_{12}/σ_1
0.148	0.312	0.185	0.131	0.372	0.661

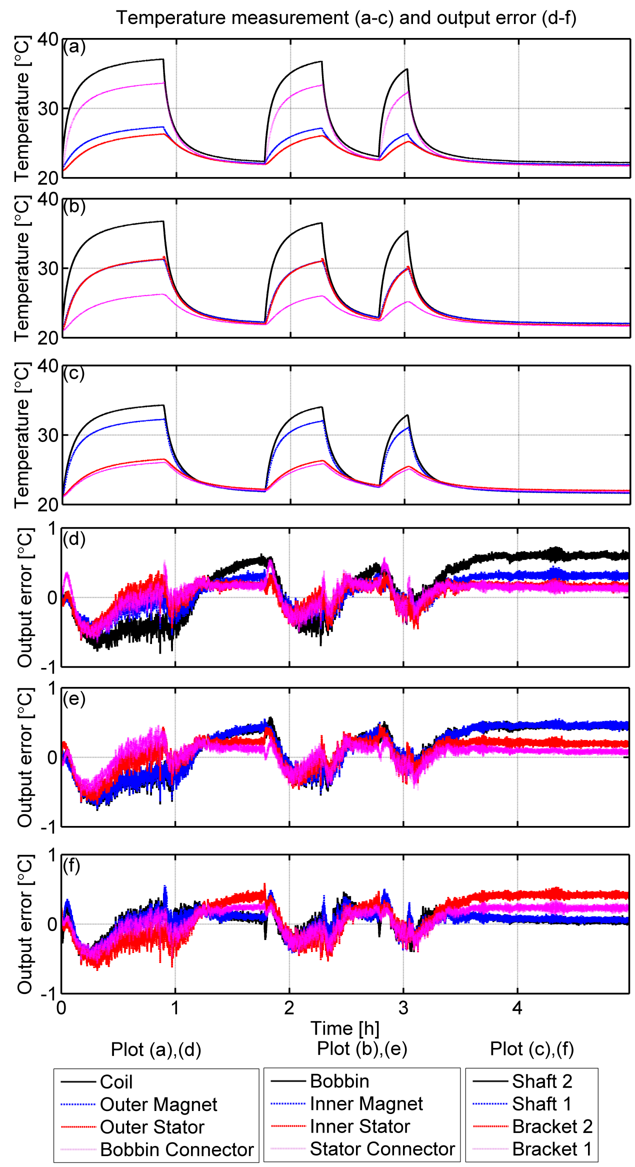


Fig. 12: Test 2 - State space model identification: (a-c) temperature response and (d-f) output error

Test 3 is designed for the characterization of the temperature estimation accuracy with and without Kalman filtering. The input motion profile is designed to induce significant thermal loading on the device. It is separated into the 'heat up' and 'cool down' period as shown Fig. 13(a). The first 5 cycle(s) of the trapezoidal motion path during the 'heat up' period is shown in Fig. 13(b). This path simulates a repeated maximum forward and backward stroke with varying duration of hold in between as illustrated in Fig. 13(c). This ensures that the device is put through the typical range of motion expected during the actual operation.

There is significant temperature rise recorded at all the points of interest as shown in Fig. 14(a)-(c) due to this input motion profile. The largest temperature rise is recorded at

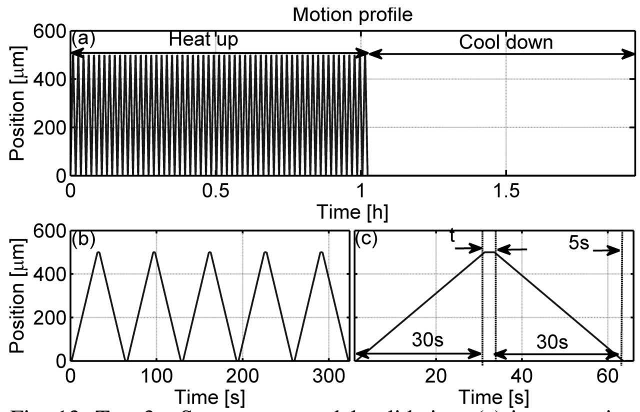


Fig. 13: Test 3 - State space model validation: (a) input motion profile, (b) 5 cycle(s) and (c) 1 cycle of the varying period trapezoidal motion path

the coil with a final temperature of 27.9°C after an hour of operation. There is a large temperature gradient between coil and outer stator as shown in Fig. 14(a). This indicates a high resistance to heat flow from the internal components outwards. One of the major barrier to heat transfer is the internal air gap between the coil and magnet assembly (refer to Fig. 1). The resulting temperature rise represents an accumulation of heat within the device which is undesirable as it gives rise to the eventual thermal expansion and positioning error.

A comparison of the model output is made against the measurement where the output error plot is shown in Fig. 14(d)-(f). A maximum output error of 0.56°C is recorded at the inner magnet as shown in Fig. 14(e) with a time averaged error of 0.26°C. From the error plots, it is apparent that the model estimation accuracy is affected by the device operating condition. The 'heat up' and 'cool down' period registers very different error signature. During the 'heat up' period the model is overestimating the temperature while the opposite is true during the 'cool down' period. There is also a slow drift in the temperature estimate away from the measurement during the 'cool down' period which is caused by the changing conditions within the test chamber. Some of these modeling inadequacies can be overcome with model output adjustment using sensor feedback in the Modified Kalman Filter (MKF).

Fig. 15 shows the output error after applying the MKF. The maximum error originally recorded at the inner magnet is now reduced to 0.16°C with an average of 0.05°C over the measurement period. With application of the MKF, the maximum error recorded across the 12 temperature points of interest amount to 0.37°C which is equivalent to 1.8% of the measurement range. The maximum and time averaged output error for other temperature points are tabulated in Table V and Table VI respectively. The error remaining after Kalman filtering is shown in brackets in both tables. As expected, the largest reduction in error is observed for the coil temperature. The model validation result show that the MKF is able correct the model output towards the actual temperature based on output mapping of the measured error signal.

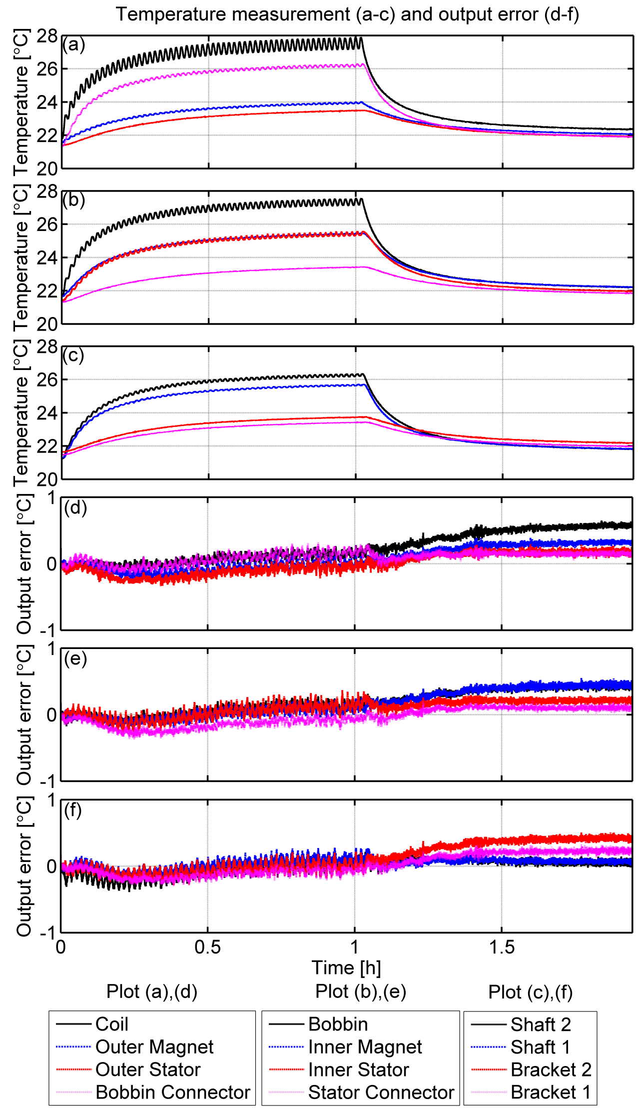


Fig. 14: Test 3 - State space model validation: (a-c) temperature response and (d-f) output error before filtering

TABLE V: Maximum output error (with) and without filtering

Coil	Bobbin	Inner Magnet	Outer Magnet	Inner Stator	Outer Stator
0.65 (0.01)	0.52 (0.16)	0.56 (0.16)	0.39 (0.24)	0.37 (0.27)	0.34 (0.32)
Stator Connector	Bobbin Connector	Shaft 1	Shaft 2	Bracket 1	Bracket 2
0.38 (0.35)	0.32 (0.22)	0.28 (0.21)	0.39 (0.37)	0.33 (0.28)	0.51 (0.25)

TABLE VI: Time averaged output error (with) and without filtering

Coil	Bobbin	Inner Magnet	Outer Magnet	Inner Stator	Outer Stator
0.33 (0.00)	0.26 (0.05)	0.26 (0.05)	0.26 (0.09)	0.19 (0.08)	0.15 (0.12)
Stator Connector	Bobbin Connector	Shaft 1	Shaft 2	Bracket 1	Bracket 2
0.14 (0.14)	0.13 (0.09)	0.09 (0.09)	0.11 (0.13)	0.16 (0.12)	0.25 (0.09)

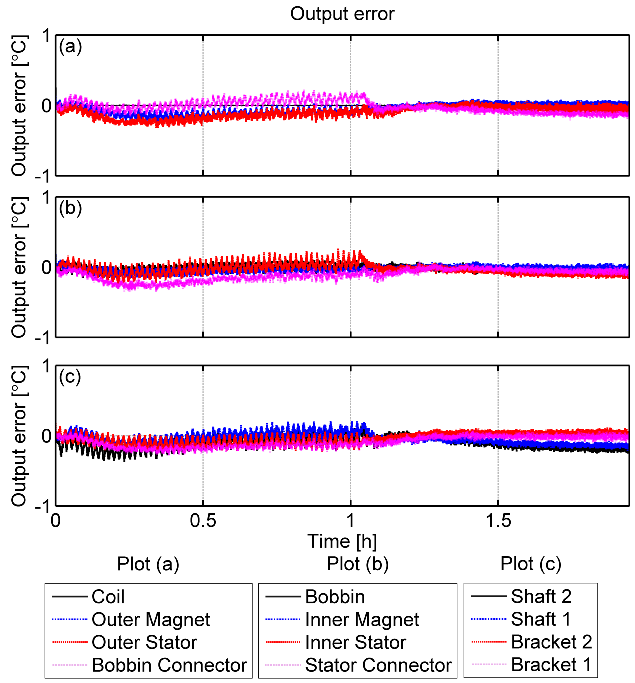


Fig. 15: Test 3 - State space model validation: output error after filtering

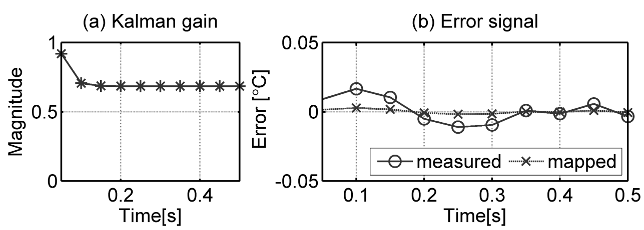


Fig. 16: (a) Kalman gain and (b) error signal

Fig. 16(a) shows the Kalman gain reaching a steady value beyond the first 5 samples of data as the model output approaches the measured value. This means a decreasing error signal over time. The mapped error signal is reduced proportionately due their linear relationship as shown by a one of the selected signal in Fig. 16(b). Output correction will be reduced over time until a large error is induced possibly by some external disturbance.

Output mapping provides an indirect estimate of the output variation at the unmeasured temperature point based on the measured error signal and the statistical characterization of the system in a prior test. However, if the cause for temperature variation is external, such as an additional heat source, it will for example have more effect on the outer stator temperature than the coil temperature. This disturbance will not be adequately captured by the embedded sensor. So long as the system is not subjected to significant changes, the MKF is an effective tool to reduce the short term tracking and long term drift error in the thermal model output. At this point, the thermal model is fully calibrated from the previous three tests. The remaining temperature sensors are removed except the embedded sensor at the coil. Temperature is estimated using the state space model coupled with the MKF in this final test.

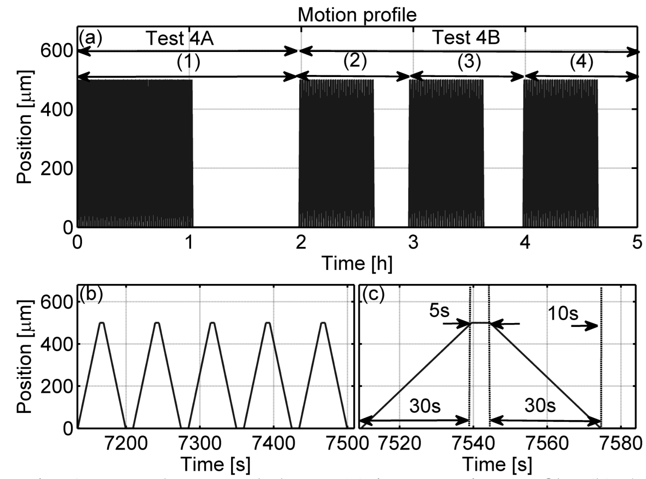


Fig. 17: Test 4 - Extended test: (a) input motion profile, (b) 5 cycle(s) and (c) 1 cycle of the fixed period trapezoidal motion path

Test 4 is divided into two parts (4A and 4B) with a total of four segments as illustrated in Fig. 17. Test 4A is a single two hour segment which consists of a continuous cycle of varying period trapezoidal motion path for one hour followed by one hour of cool down. The varying period trapezoidal motion path is similar in design to the one used in test 3 (refer to Fig. 13). This test is designed for the identification of the transfer function model. The model output is taken as the measured position error defined in section III. The temperature response resulting from this motion profile is shown in Fig. 18(a)-(c). There is significant temperature variation due to the cyclical thermal loading which is ideal as the input for fine tuning the transfer function model parameters.

Test 4B has 3 one hour segments which each consist of a continuous 32 cycle(s) of 75s fixed period trapezoidal motion path followed by a 'cool down' period as shown in Fig. 17(a). The fixed period trapezoidal motion path is shown in Fig. 17(b)-(c). This test is used for validation of the compensation model. This validation test is designed to simulate a continuous operation. The device will undergo cycles of thermal loading during the 'heat up' period and significant temperature variation is observed as a result of this motion profile. The maximum temperature recorded during this extended test is 27.6°C as shown in Fig. 18(a). The physical thermal boundary conditions such as the chamber air temperature is expected to change over the course of the test. This would in turn affect the compensation model estimation accuracy. However, the effect of changing boundary conditions can be minimized with minimal temperature sensing using the Modified Kalmam Filter. This would help improve the positioning error tracking ability of the compensation model.

The positioning error before compensation is shown in Fig. 18(d). The mean unidirectional position deviation before compensation is $+5.7\mu$m with a repeatability of $\pm 7.3\mu$m for the duration of the extended test (calculated based on ISO230-2). A SISO transfer function model which uses only the

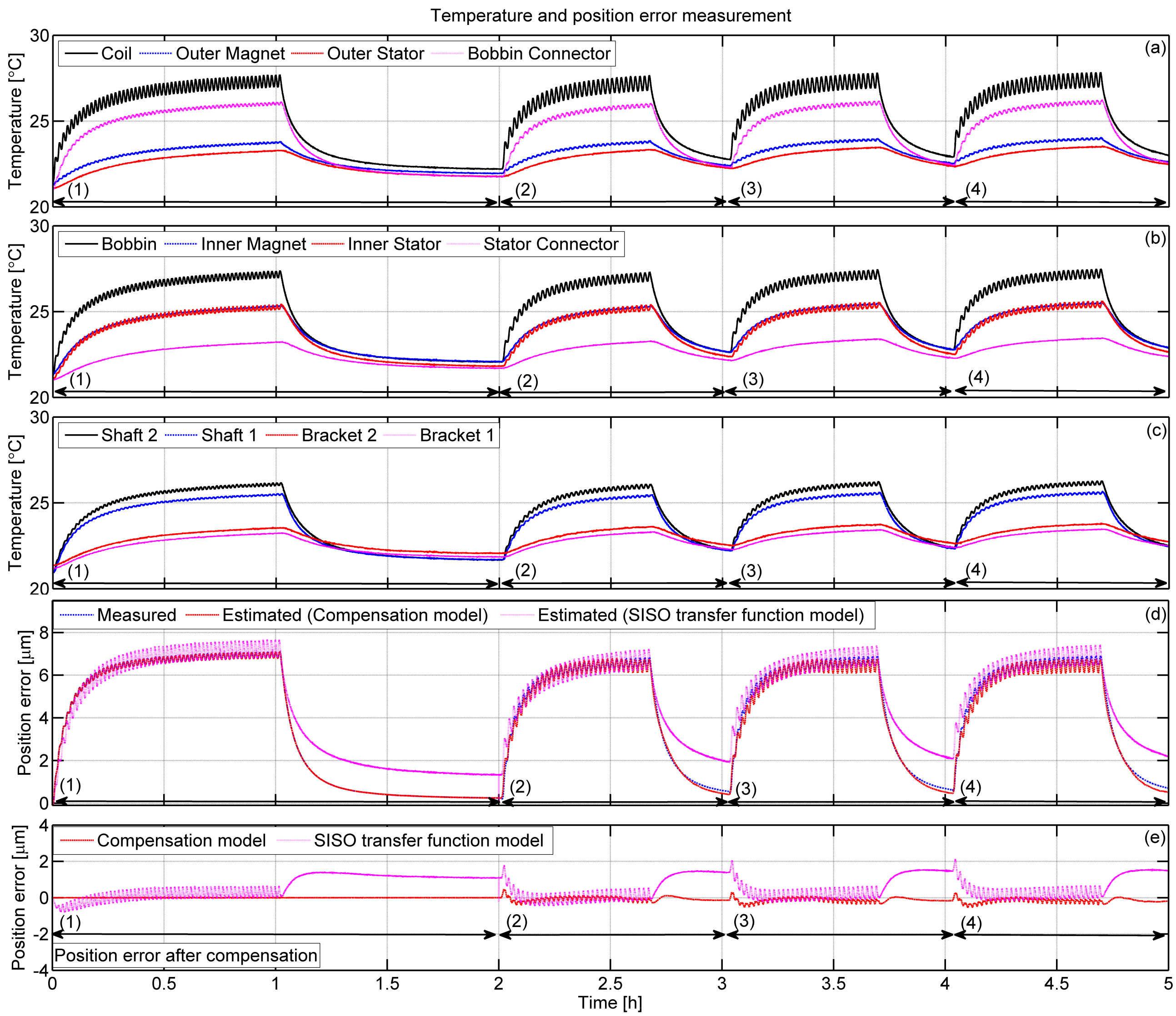


Fig. 18: Test 4 - Extended test: (a-c) temperature estimate with Modified Kalman Filtering, (d) position error and (e) position error residual after compensation

measured coil temperature as an input is used as the estimator of position error for comparison with the compensation model. The model estimated position error from both models are plotted against the measurement in Fig. 18(d) and the residuals are shown in Fig. 18(e). The residual is minimized during the first test segment as the position error is measured and used for model calibration. Thereafter, the measurement is used solely for validation as shown in test segments 2-4 of Fig. 18(e).

The SISO transfer function model is able to reduce the position error down to a mean deviation of $+0.7\mu$m with a repeatability of $\pm 2.3\mu$m. With implementation of the compensation model in a feedback arrangement, the mean unidirectional position deviation have been reduced even further to -0.2μm with a repeatability of $\pm 0.7\mu$m. The positioning accuracy achieved with the proposed method shows a threefold increase over the expected improvement if a SISO transfer function model is implemented instead. The model based compensation method is effective in reducing a large component of the position error caused by thermal loading. It shows good short term tracking ability as well as long term stability.

VI. Conclusion

This work illustrates the effect of dynamic thermal disturbance on positioning accuracy of a precision linear electromagnetic actuator. A model based compensation method is applied to the test device. It resulted in a final mean unidirectional position deviation of -0.2μm with a repeatability of $\pm 0.7\mu$m during a continuous 5 hour test. The significant improvement in positioning accuracy is due to the robust estimation of the position error using the compensation model. The model is determined from an identification procedure. Online model tuning is applied in an extended test to illustrate its practical usefulness. Output mapping of the measured error signal from the embedded temperature sensor in a Modified Kalman Filter (MKF) resulted in minimal temperature sensors required for active position compensation. A total of only 1 sensor is needed from an original 12. Application of a MKF resulted in improved temperature estimates with absolute error of $0.37°C$ which is equivalent to 1.8% of the measurement range.

REFERENCES

[1] J. Mayr, J. Jedrzejewski, E. Uhlmann, M. Alkan Donmez, W. Knapp, F. Härtig, K. Wendt, T. Moriwaki, P. Shore, R. Schmitt, C. Brecher, T. Würz, and K. Wegener, "Thermal issues in machine tools," *CIRP Annals - Manufacturing Technology*, vol. 61, no. 2, pp. 771–791, 2012.

[2] R. Ramesh, "Error compensation in machine tools a review Part II: thermal errors," *International Journal of Machine Tools and Manufacture*, vol. 40, no. 9, pp. 1257–1284, Jul. 2000.

[3] M. U. Khan, N. Bencheikh, C. Prelle, T. Beutel, and B. Stephanus, "A Long Stroke Electromagnetic XY Positioning Stage for Micro Applications," *IEEE Transactions on Mechatronics*, vol. 17, no. 5, pp. 866–875, 2012.

[4] T. J. Teo, "Cylindrical Electromagnetic Actuator," WO2012/074482 A1.

[5] Z. Chen, B. Yao, and Q. Wang, "Accurate Motion Control of Linear Motors With Adaptive Robust Compensation of Nonlinear Electromagnetic Field Effect," *IEEE Transactions on Mechatronics*, vol. 18, no. 3, pp. 1122–1129, 2012.

[6] L. Bascetta, P. Rocco, and G. Magnani, "Force Ripple Compensation in Linear Motors Based on Closed-Loop Position-Dependent Identification," *IEEE Transactions on Mechatronics*, vol. 15, no. 3, pp. 349–359, 2010.

[7] H. Yang and J. Ni, "Adaptive model estimation of machine-tool thermal errors based on recursive dynamic modeling strategy," *International Journal of Machine Tools and Manufacture*, vol. 45, no. 1, pp. 1–11, Jan. 2005.

[8] R. Zhou, B. Gressick, J. T. Wen, M. Jensen, J. Frankel, G. Lerner, and M. Unrath, "Active Thermal Management for Precision Positioning," in *Proc. IEEE Conf. on Automation Science and Engineering*, Scottsdale, AZ, Sep. 2007, pp. 45–50.

[9] Z. Gao, T. Habetler, R. Harley, and R. Colby, "An adaptive kalman filtering approach to induction machine stator winding temperature estimation based on a hybrid thermal model," in *Proc. IEEE Industry Applications Conf.*, Hong Kong, Oct. 2005.

[10] K. Ito, W. Maebashi, M. Yamamoto, M. Iwasaki, and N. Matsui, "Fast and precise positioning by sequential adaptive feedforward compensation for disturbance," in *11th IEEE International Workshop on Advanced Motion Control (AMC)*, Japan, Mar. 2010.

[11] L. Ljung, *System identification: Theory for the user*. Prentice Hall, 1999.

[12] O. Peter and M. Bart, "Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems," *Automatica*, vol. 30, pp. 75–93, 1994.

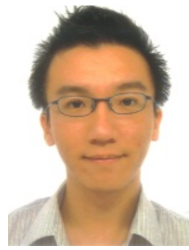
[13] L. Ljung and T. Mckelvey, "A Least Squares Interpretation of Sub-space Methods for System Identification," in *Proc. of the 35th Conference on Decision and Control*, 1996.

[14] P. Van Overschee and L. De Moor, *Subspace identification for linear systems: theory, implementation, applications*. Kluwer Academic Publishers, 1996.

[15] J. Nerg, M. Rilla, and J. Pyrhonen, "Thermal Analysis of Radial-Flux Electrical Machines With a High Power Density," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 10, pp. 3543–3554, Oct. 2008.

[16] J. Gieras, Z. Piech, and B. Tomczuk, *Linear Synchronous Motors: Transportation and Automation Systems*. Taylor and Francis, 2012.

[17] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital control of dynamic systems*. Addison-Wesley, 1998.



Jonathan Hey received the B.Eng (Hons) Degree in Mechanical Engineering from Nanyang Technological University (NTU), Singapore, in 2008. He is a recipient of the A*STAR graduate scholarship. He is currently a candidate for the Ph.D. degree at Imperial College London, UK. His research interest is on thermal management of electromechanical devices focused on disturbance model identification, compensation methods and thermal design analysis.

Choon Meng Kiew received the B.Eng (Hons) Degree in Electrical and Computer Engineering from National University of Singapore (NUS) in 2003. After which, he was awarded A*STAR Graduate Scholarship to pursue his Ph.D. studies at NUS and was conferred the degree in 2008. He joined Singapore Institute of Manufacturing Technology since Sept 2007. His research interests are in the area of thermal process control, precision control of nano-positioning stages, control of web handling system and also thermal related analysis

Guilin Yang received the B. Eng degree and M. Eng degree from Jilin University, China, in 1985 and 1988 respectively, and Ph.D. degree from Nanyang Technological University in 1999. From 1988 to 1995, he had been with the School of Mechanical Engineering, Shijiazhuang Tiedao University, China, as a lecturer, a division head, and then the vice dean of the school. Since 1998, he has been with Singapore Institute of Manufacturing Technology (SIMTech), Singapore. Currently, he is a senior scientist and the manager of the Mechatronics Group. His research interests include precision mechanisms, electromagnetic actuators, parallel-kinematics machines, modular robots, industrial robots, and rehabilitation devices. He has published over 190 technical papers in referred journals and conference proceedings, 8 book chapters, and 2 books. He has also filed 12 patents. He was the committee chair of Singapore IEEE Robotics and Automation Chapter (2011- 2012) and the Technical Editor of IEEE/ASME Transaction on Mechatronics (2008-2012). He is now the Editorial Board Members of Frontiers of Mechanical Engineering and International Journal of Mechanisms and Robotic Systems, and an associate editor of IEEE Access.

Ricardo Martinez-Botas is a Professor of Turbomachinery at Imperial College London. He has an MEng (Hons) Degree in Aeronautical Engineering from Imperial College London and a doctoral degree from the University of Oxford. He has developed the area of unsteady flow aerodynamics of small turbines, with particular application to the turbocharger industry. The contributions to this area centre on the application of unsteady fluid mechanics, instrumentation development and computational methods. He has expanded his research in the area heat transfer of electrical machines and batteries. He has published extensively in journals and peer reviewed conferences. He is currently Associate Editor of the Journal of Turbomachinery and is a member of the editorial board of two other international journals. He is currently the Theme Leader for Hybrid and Electric Vehicles of the Energy Futures Lab at Imperial College.