Temperature-Independent Fault Detection of Solenoid-Actuated Proportional Valve

Pedersen, Henrik C.; Bak-Jensen, Terkil; Jessen, Rasmus H.; Liniger, Jesper

Published in:
IEEE/ASME Transactions on Mechatronics

DOI (link to publication from Publisher):
10.1109/TMECH.2022.3158483

Publication date:
2022

Document Version
Accepted author manuscript, peer reviewed version

Temperature Independent Fault Detection of Solenoid Actuated Proportional Valve

Henrik C. Pedersen, Terkil Bak-Jensen, Rasmus H. Jessen, Jesper Liniger

***Abstract*—Most electrically actuated hydraulic valves utilize solenoids as the actuating element due to their robustness and simplicity. This goes for both on-off and proportional type valves. However, despite their robustness, solenoid coil malfunction is the largest single failure mode in solenoid actuated valves. An outspoken fault is here solenoid winding short-circuit, i.e. two windings short-circuiting, which may ultimately lead to solenoid failure if more windings short-circuit. Research has therefore also focused on detecting winding short-circuits. Common for the approaches are that they, directly or indirectly, depend on the coil winding temperature, as this directly influences the coil resistance. Alternatively, the approaches are based on injection of high-frequency signals, which is typically a costly solution that is not a feasible approach for use in hydraulic valves, with the limitations imposed by the control electronics. Therefore, this paper focuses on a temperature independent algorithm to detect coil winding short-circuit, which is easy to implement and only relies on existing position and current sensors. The proposed algorithm is based on an Extended Kalman Filter (EKF), which estimates the coil resistance. As this resistance estimate is indirectly dependent on the coil temperature, a window-based CUmulative SUM (CUSUM) fault detection method is included to detect transient changes in the coil resistance while compensating for the temperature variations. The algorithm is developed based on an experimentally validated model of the valve and has been tested for several different situations through both simulations and experimentally. Based on the presented results, it is found that the algorithm can consistently detect resistance changes down to $0.11\,\Omega$ for constant input signals, and down to $0.17\,\Omega$ for sinusoidal varying input signals. This while still being robust to parameter variations, like increased valve friction, spring coefficients and sensor signal deviations.**

***Index Terms*—Fault detection, temperature independent, solenoid, valve, winding short-circuit.**

I. Introduction

SOLENOID valves are used in a long range of applications due to their simplicity, high force capacity and robustness, where the solenoid coil is separated from the pressurized fluid chamber(s). Despite their simplicity, the solenoids and solenoid coils are reported to be one of the major failure sources for solenoid operated valves, [1], [2], accounting for more than 50% of the failures reported in [1] for valves used in nuclear systems. Similarly, Angadi *et al.*, [3], points to coil malfunction due to thermal overheating and break-down of insulation material as the main failure cause for the valves they investigate. Hence, typical causes for insulation break-down are thermal effects, vibrations and/or chemical degradation of the insulation material. In the short run, this will lead to winding short-circuits, which then again leads to increased coil current, increasing temperature and ultimately coil burn-out. Typically, at least five coil (short circuit) faults are seen before coil failure (burn-out), [4]. Being able to detect these coil faults will hence be a good indicator for the coil's degradation and may aid in replacing/servicing the valve before failure.

Detecting solenoid faults is not a new idea, and several approaches have been presented. However, it should be noted that winding short-circuits will yield both a drop in resistance and change the inductance and capacitance of the coil. Therefore, most of the approaches presented have been based on estimating/determining one or more of these parameters. Therefore, one group of methods effectively considers resistance changes due to coil faults, whereas other (basically impedance-based) methods focus on the transient and/or high-frequency changes primarily resulting from the inductance also being affected by coil faults. Starting with the latter Kryter, [1], investigated several detection methods for solenoid valve failure finding that the most promising method relied on monitoring the coil voltage when disconnecting the power system. Jameson *et al.*, [2], [5], utilized frequency analysis and accelerated lifetime testing to investigate a variety of coils, measuring reactance and resistance in a large frequency span. Utilizing Spearman correlation analysis, it was found that the best excitation frequencies to detect insulation break-down was in the range $20 - 100\,kHz$. However, while these methods may show promising results, the applicability to hydraulic valves is very limited due to the limitations imposed by the control electronics, which are not designed to handle high-frequency signals for cost reasons.

Considering the methods relating to resistance estimation, then Jo *et al.*, [6], presented a fault detection method for detecting short circuit faults in solenoid operated on/off valves. The method utilizes ambient temperature and maximum solenoid supply current to calculate a temperature-dependent health indicating current value, which is compared to a tolerance value to detect faulty valves. The method yielded a fault detection accuracy of 100% for ambient temperatures of -20° to 60° for the considered batch. Liniger *et al.* utilize a state augmented EKF-approach to describe not only the coil current but also predict the coil temperatures based on measurement of ambient temperature, fluid temperature, voltage and coil current. Based on the current residuals, a cumulative sum was used to indicate when a winding short circuit occurred. The approach was shown to be effective with detection probabilities above 97% and was also shown to be robust to changes in the ambient temperature, convection, and fluid flow conditions. However, the method relies on temperature measurements and the need for fairly accurate

Henrik C. Pedersen is with the Department of Energy Technology, Aalborg University, DK-9220 Aalborg East, Denmark, e-mail: hcp@et.aau.dk

T. Bak-Jensen, R.H. Jessen and J. Liniger are with Aalborg University.

Manuscript received April, 2021; revised -

thermal models for the solenoid. Khoshzaban-Zavarehi, [7], also utilized an EKF-based approach to estimate coil resistance based on coil voltage and current but did not address the effect of temperature drift of the resistance. The same applies to the results of Jung *et al.*, [8], and Dülk and Kovácsházy, [9], which present methods for estimating the coil resistance, but do not focus on fault detection and therefore do not need to consider the resistance change due to temperature drift. On a more general scale, insulation break-down has also been considered in relation to, e.g. stator windings in induction motors. However, these methods generally rely on multiple phases, whereby the temperature drift is out compensated, which is hence not possible for a single coil solenoid. For an overview of these methods, see e.g. [10], [11].

From the above, it is clear that multiple approaches have been taken to detect solenoid short circuit faults. However, the high-frequency measurement based methods are not suited for hydraulic valves, whereas the resistance estimation based methods are temperature dependent, as a result of this imposing requirements for extra sensors or limiting their applicability. Therefore, the focus of the current paper is on a novel algorithm for *temperature independent* detection of early short circuit faults in solenoid coils, with minimal sensor and sampling frequency requirements. The temperature independence here means that the algorithms detect short circuit faults fully independent on the temperature in both coil and surroundings, and unlike the approach in [4] does not rely on determining the temperatures or thermal time constants for the different parts in the system. On the contrary, the algorithm compensates for the resistance changes in the coil due to temperature variations without applying any kind of thermal elements, advanced sensors or the like. As the computational requirements are also low, this yields a widely applicable algorithm. The method is based on an EKF for joint state and parameter estimation of the coil current, spool position and coil resistance while compensating for multiple non-linearities in the model. Based on the EKF resistance estimation, a windowed CUSUM approach is used to detect winding short circuits while disregarding the influence of resistance drift caused by temperature variations of the coil. Furthermore, the algorithm's robustness is tested towards common parameter variations for the considered valve, including gain, noise, and bias variations in the sensor signals. It should be noted that while the algorithm is developed and tested on a proportional solenoid valve, it may readily be applied on on-off solenoid valves as well, as long as the plunger position is properly estimated or measured.

The paper is organized as follows: An experimentally validated model is first presented, upon which the EKF is based. The validity of the EKF and its ability to detect resistance changes are then presented. This is followed by a description of the windowed CUSUM approach used, and validation of its ability to consistently detect coil short-circuits even in case of parameter variations. Finally, experimental results are presented along with a conclusion summarizing the main findings.

II. System description

The valve utilized in this study is a solenoid driven 4/3 proportional 4WRE-10 valve from Bosch-Rexroth. This is a common type solenoid valve, which is finding widespread use in both industry and pitch systems in wind turbine applications. An illustration of the exploded valve is shown in Fig. 1. The spool assembly consists of a spool guiding the oil in specified directions, two plungers adjusting the spool position, and four springs centering the spool when current is not applied to either solenoid. Furthermore, it may be seen from the figure that the valve incorporates a position transducer in the form of an LVDT. The electrical part consists of two solenoids, one on each side to move the spool in the given direction. An illustration of the right side solenoid and the attached LVDT is shown in Fig. 2. The solenoid resistance given in the datasheet is 4.55Ω.

It should be stressed that, although the valve considered here is a proportional valve, the algorithm presented is generally applicable and may also be used on, e.g. solenoid driven on-off type valves. Hence, for the latter case, the plunger position estimation, as described later, will simply be left out in the EKF-model.

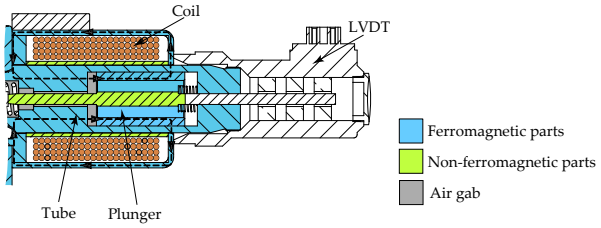

Fig. 2. Illustration of right side solenoid and the attached LVDT.

A. System Model

A model of the system is first presented, which is used as basis for describing the proposed EKF based Fault Detection and Diagnostics (FDD) algorithm, and for testing the robustness of this. Although the focus is on electrical winding short-circuit faults, the model also includes the mechanical part of the valve, as the flux linkage is dependent on both the plunger position and current. The model is governed by the solenoid voltage equation, the electromagnetic force and Newton's second law of motion, and may be illustrated as shown in Fig. 3

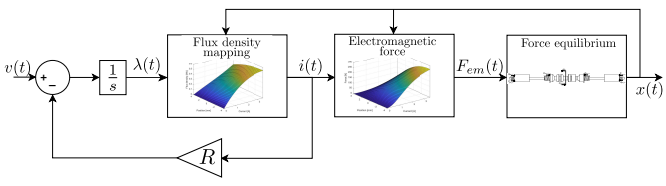

Fig. 3. Model structure.

The solenoid voltage equations is given by:

$$v = Ri + \frac{d\lambda(i,x)}{dt} \tag{1}$$

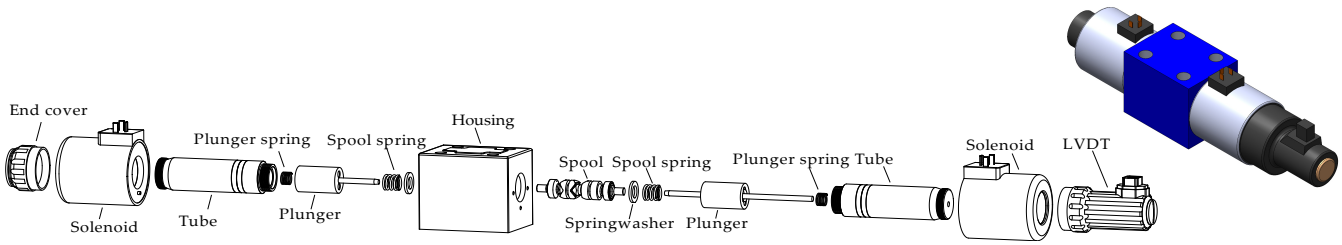

Fig. 1. Exploded view of the valve with notation.

where all states, of course, are time-dependent, but this is omitted to ease the notation. The flux linkage, λ, is a function of the plunger position and the current and may be written as $\lambda(i,x) = L(i,x)i$ and R is the coil resistance. The flux linkage is hysteresis dependent. Therefore, determining the flux linkage and electromagnetic force analytically is complicated and earlier attempts at this have shown to yield poor results. Therefore, both the flux linkage and electromagnetic force have been determined experimentally. The flux linkage is determined by mechanically fixing the plunger position and applying applying slow ($1\,Hz$) sinusoidal input voltages to the valve with different amplitudes, while at the same time measuring the current. This experiment is then repeated for different plunger positions. A plot of the results from these experiments is shown in Fig. 4, where the flux density is plotted as function of the current.

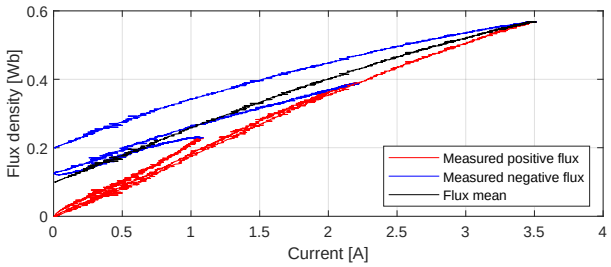

Fig. 4. Flux density for sinusoidal voltage with three different amplitudes.

Since the flux linkage is influenced by hysteresis, the mean value is utilized in the following. The mean flux linkage is mapped as a function of plunger position and current, and a polynomial fit of this is made to avoid interpolating in the algorithm. In a similar way, a mapping and polynomial fit of the electromagnetic force has been made. This has been done by controlling the current and linearly increasing this from $0A$ to $3A$ over a 10 second period, and again repeating this for several different fixed plunger positions. The resulting mappings of the flux density and electromagnetic force as functions of plunger position and current are shown in Fig. 5.

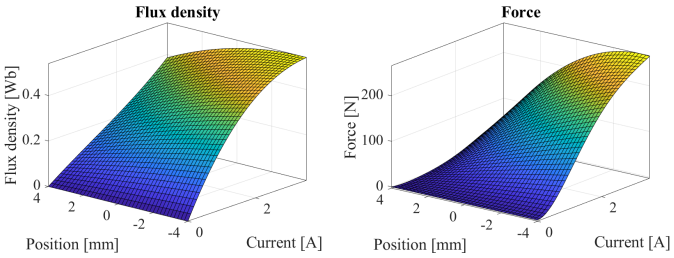

Fig. 5. Flux density (Coefficient of determination: $R^2 = 0.98$) and electromagnetic force ($R^2 = 0.99$) mapping.

Considering the mechanical part of the model, the forces acting on the spool and plungers may be seen in Fig. 6. As the plunger springs are pre-compressed by more than the possible movement of the spool, the plungers and spool may be treated as a single element.

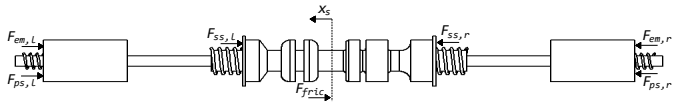

Fig. 6. Free body diagram of valve spool and plungers.

The motion of the spool is determined from Newton's second law of motion as:

$$m_s\ddot{x}_s = F_{em} - F_s - F_{fric}(\dot{x}) - F_d \tag{2}$$

where F_{em} is the combined electromagnetic force from the two solenoids. $F_s = (2\,k_{ps} + k_{ss})\,x_s$ is the combined spring force, and F_{fric} is the friction force modelled as a combination of viscous and Coulomb friction. Finally F_d is a lumping included to account for unmodelled effects and external disturbances such as the flow force.

B. Model validation

To illustrate the validity of the model, the simulated current and position is compared to the experimentally measured data. The same input voltage signal is used for the simulation as in the experiments, which is a sinusoidal signal that increases in frequency from 0.125 to $10\,Hz$. The result from the validation may be seen in Fig. 7.

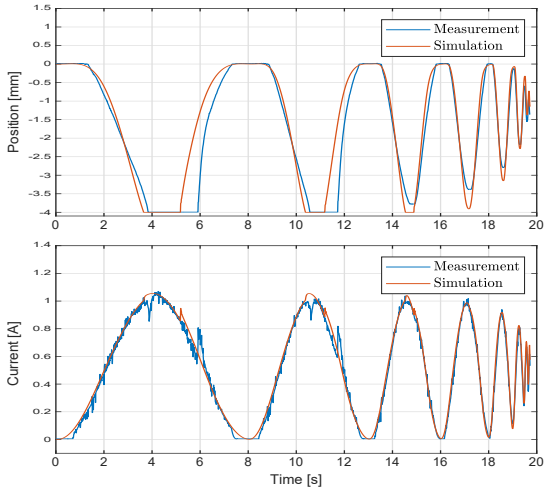

Fig. 7. Simulated and measured position and current. The spool end position is at -4 mm. The small current spikes aligns with when spool hits and leaves the end position.

From the graphs, it may be seen that there is a fair correspondence between the model and the measured response. Considering the spool position, there are some unmodelled effects near the spool end-stop, which the model does not capture. This deviation is due to the spool "sticking to the end" when in the end positions, i.e. it needs to overcome Stribeck friction before moving. However, this is mainly a problem in this model validation experiment, as the experiment is made without oil in the valve. Before the experiment the valve has been filled with oil, but the experiments are done without oil in the valve. Thus, there is only a thin oil film left and no oil pressure to help lubricate and radially center the spool, which results in the spool sticking, as the fluid tension forces in the oil film has to be overcome, which is not captured by the model. The small deviation between the measured and simulated current, particularly the misalignment of the spikes when the spool starts moving, is therefore also a result of the spool sticking, which is not captured. However, for the development and validation of the solenoid winding short-circuit FDD algorithm this deviation is of insignificant importance. Therefore, the model is also sufficient for the development and testing of the algorithm, as it captures the current correctly, and this is the most important of the two states, as the resistance estimation is based on this.

III. State and Resistance Estimation (FDD part 1)

The presented FDD algorithm is comprised of two parts. The first part is an Extended Kalman Filter (EKF) for estimating the valve's states, i.e. coil current, spool position and velocity, and augmented with a state parameter estimation part for estimating the coil resistance. The reason for using this kind of observer is that, besides its noise benefits, the EKF is well suited to accommodate the non-linearities inherent in the system resulting from both the flux-linkage and electromagnetic force, but also the non-linearities in from the parameter estimation building on these. However, as the functions are still continuous there in no need to apply, e.g. an Unscented Kalman Filter. The benefit of the EKF approach is in this regard, that it is easy to implement and augment with the parameter estimation. Still, it is fairly easy to tune, as the EKF parameters directly rely on measurement and process noise and considerations of the model uncertainties.

A. EKF algorithm

The EKF is based on the non-linear stochastic model given below, where the external disturbance force is omitted, and the current derivative is found from (1) by taking the partial derivative of the flux linkage. Furthermore, the model is augmented with the state estimation for the solenoid resistance:

$$\mathbf{f}(\mathbf{x},u,\mathbf{w}) = \begin{bmatrix}\dot{i}\\ \ddot{x}_s\end{bmatrix} + \mathbf{w} = \begin{bmatrix} \frac{v - Ri - \frac{\partial \lambda(i,x_s)}{\partial x_s}\dot{x}_s}{\frac{\partial \lambda(i,x_s)}{\partial i}} \\ \frac{F_{em}(i,x_s) - (2\,k_{ps}+k_{ss})\,x_s - B\dot{x}_s}{m_s}\end{bmatrix} + \mathbf{w} \tag{3}$$

$$\mathbf{g}(\mathbf{x},\mathbf{v}) = \begin{bmatrix} i & x_s \end{bmatrix}^T + \mathbf{v} \tag{4}$$

$$\mathbf{x} = \begin{bmatrix} i & x_s & \dot{x}_s & R \end{bmatrix}^T \tag{5}$$

Here $\mathbf{w}$ is the process noise and $\mathbf{v}$ the measurement noise. The EKF algorithm is separated into three steps. In the *initialization* step the initial state and parameter estimate, $\hat{\mathbf{x}}_\mathbf{0}$, the initial predicted state estimate covariance, $\mathbf{P_0}$, as well as both the measurement noise, $\mathbf{N}$, and state transition matrix process noise, $\mathbf{M}$, are defined. The initial state estimates are based on the datasheet value for R, whereas the other states are zero at start-up. The measurement noise, $\mathbf{N}$, is based on analysis of the steady-state measurement noise, whereas the process noise, $\mathbf{M}$, is adjusted considering the model uncertainty and to obtain a good compromise between filter dynamics and steady-state oscillations. It is further assumed that the noise characteristics for each state are independent, resulting in diagonal matrices. Finally, $\mathbf{P_0}$ is adjusted to obtain an acceptable convergence speed of the filter, here resulting in using the identity matrix:

$$\hat{\mathbf{x}}_0 = \begin{bmatrix} 0 & 0 & 0 & R \end{bmatrix}^T \tag{6}$$

$$\mathbf{N} = \begin{bmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_x^2 \end{bmatrix} \tag{7}$$

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 1\cdot 10^{-7} \end{bmatrix} \tag{8}$$

$$\mathbf{P_0} = \mathbf{I} \tag{9}$$

The sensor noise variation is measured through a series of steady-state tests, and the mean is $\sigma_i = 15.4\,mA$ and $\sigma_x = 4.56\,\mu m$. The initial test resistance used is the datasheet value of $R = 4.55\Omega$.

In the *prediction* phase, an estimate of the system states in equation (10) is created, along with the state transition covariance matrix in equation (11). The subscript $k|k-1$ indicates that the intermediate value for the current time step

is being calculated based on information from the previous timestep:

$$\hat{\mathbf{x}}_{k|k-1} = \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{f}(\mathbf{x}_{k-1}, u_{k-1})T_s \tag{10}$$

$$\mathbf{P}_{k|k-1} = \mathbf{J}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{J}_{k-1}^T + \mathbf{WMW}^T \tag{11}$$

In Eq. (10) $\mathbf{W}$ is a distribution matrix which is chosen as a 4×4 identity matrix since the process noises are assumed to be independent from each other. $\mathbf{J}$ is the Jacobian matrix:

$$\mathbf{J}_{k[i,j]} = \left.\frac{\partial \mathbf{f}_i}{\partial \mathbf{x}_j}\right|_{\hat{\mathbf{x}}_{k-1}, u_{k-1}} \tag{12}$$

In the *update* stage a posteriori estimation is done based on the predicted (*a priori*) estimates:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{C}_{k-1}^T(\mathbf{C}_{k-1}\mathbf{P}_{k|k-1}\mathbf{C}_{k-1}^T + \mathbf{VNV}^T)^{-1} \tag{13}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}) \tag{14}$$

$$\mathbf{P}_{k+1} = (\mathbf{I} - \mathbf{K}_k\mathbf{C}_k)\mathbf{P}_{k|k-1} \tag{15}$$

Here $\mathbf{V}$ is a 4×4 identity matrix as the measurement noises are assumed independent from each other. Throughout the update stage, $\mathbf{C}$ is the Jacobian matrix:

$$\mathbf{C}_{k[i,j]} = \left.\frac{\partial \mathbf{g}_i}{\partial \mathbf{x}_j}\right|_{\hat{\mathbf{x}}_{k-1,0}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{16}$$

With the algorithm derived and the necessary matrices tuned, the next sections will focus on the performance of the state and parameter estimation of the EKF.

B. Validation of EKF

To show the performance of the EKF Figs. 8 & 9 shows the estimated, simulated, and measured current, position and resistance.

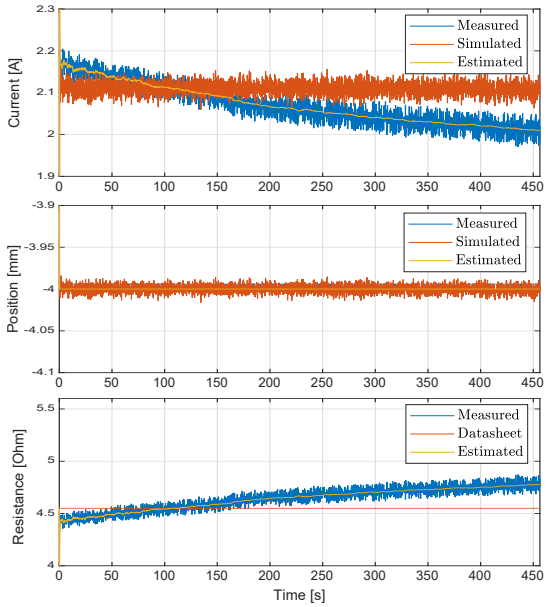

Fig. 8. EKF state estimations for constant input.

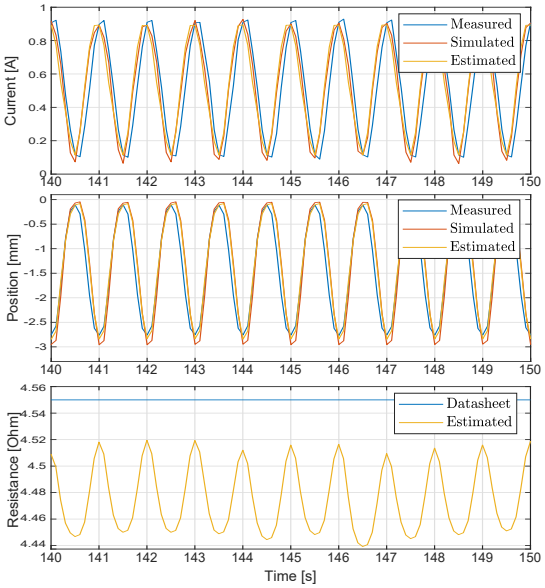

Fig. 9. EKF state estimations for $1\,Hz$ sine input.

Overall, the state estimation of the EKF is seen to nicely track both current and position, displaying errors of 40 mA and 10 μm respectively for the constant input due to noise. Also, the resistance is tracked nicely, even with a starting guess from $4\,\Omega$. The drift in current and resistance is due to the temperature increase of the coil. However, the EKF ensures that the measured current is included in the estimation of the resistance. For the sine input, the errors are larger. The current error is here $200mA$, and the position error is $600\mu m$. This is mainly due to the flux linkage approximation, which is based on a mean flux linkage hence not accounting for the hysteresis effects. The figure shows that the resistance estimate is influenced by the state oscillations, as the EKF is adjusted to yield a good compromise between accuracy (due to flux linkage approximation) and fast detection capability. While improving the flux linkage model may yield better results, it is actually desirable for the FDD algorithm to make this robust towards uncertainties in the flux and model parameters. Therefore the implemented flux linkage model is sufficient as it is.

C. Fault injection

In order to test the ability to detect a resistance drop a change is imposed in the measurement data at $t = 250\,s$. The input to the EKF is the voltage, but to emulate the resistance change in the measured data, which is feed back in the update stage of the EKF, in Eq. (14), an offset is added to the measured current corresponding to a resistance decrease of $0.1\,\Omega$. The result of this emulated resistance change is seen in Fig. 10.

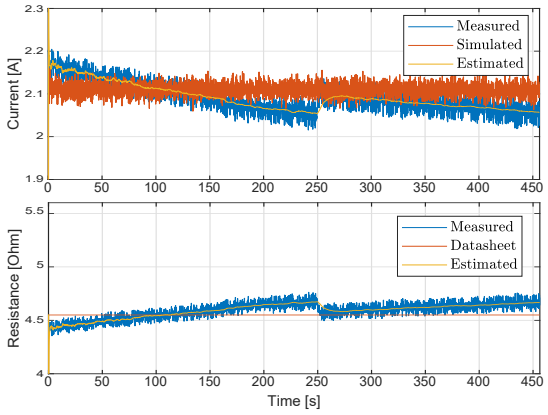

Fig. 10. EKF state and parameter estimations for a $0.1\,\Omega$ fault injection.

A similar plot is seen for the estimated resistance, when the resistance change is imposed in the sine signal. The results of this fault injection can be seen in Fig. 11.

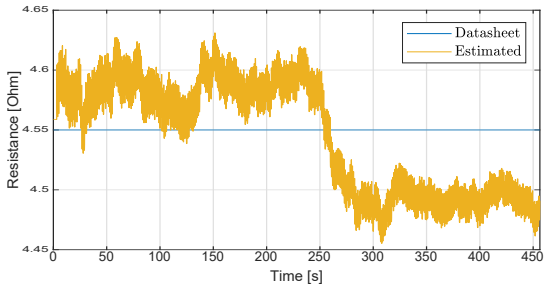

Fig. 11. Resistance estimation using a sinusoidal input voltage. $V_{peak} = 2.4V$

From both the above results, it is seen that the EKF detects the change in resistance very well, although the current fluctuations still influence the estimation for the sine input. Still, the mean value indicates the resistance drop, which is used in the following part of the FDD algorithm.

IV. FDD USING WINDOWED CUSUM (FDD PART 2)

A typical method concerning fault detection is the CUSUM method. The CUSUM method depends on detecting deviations from a calculated target mean. The cumulative sum is generated from an input, in this case the resistance estimate, to determine if the mean changes throughout the dataset. The output of the CUSUM is a lower and upper CUSUM, where the CUSUM denotes the number of standard deviations in which the cumulative signal deviates from the cumulative mean. The problem with this approach is the resistance drift due to changes in temperature, which means that in its basic form, the CUSUM will also detect resistance changes due to temperature drift. This hence needs to be compensated for, like, e.g. using a temperature model as presented in [4]. However, an accurate temperature model requires an ambient temperature sensor and detailed knowledge of the thermal characteristics of the solenoid and valve. Therefore, a detection method, which is temperature independent but still robust to parameter uncertainties, is introduced in the following. The method is here based on calculating a CUSUM twice; first for the resistance deviation within a given window, where the output considered is the number of samples before the threshold is exceeded (if exceeded), and secondly for determining the deviations in the outputs of the first CUSUM for a series of windows. Conceptually this corresponds to a moving target mean as illustrated in Fig. 12, where the tolerance band, which is usually static, is modified, so it follows the slow drift of the temperature. In these CUSUM plots, the red circles indicate the samples where the threshold is exceeded. It should here be noted that while the presented algorithm has been implemented offline for the testing, the algorithm may easily be modified to be implemented online instead. This is done by only looking at a specified set of windows in the second application of the CUSUM and/or applying moving windows.

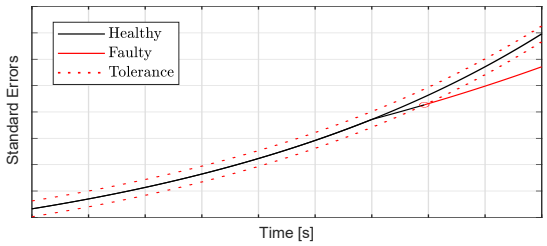

Fig. 12. Concept sketch of the windowed CUSUM.

A. Fault detection using windowed CUSUM

The algorithm has to take into account the resistance change due to temperature, which is slower than the change due to a winding short-circuit. The concept is to use this information to make the resistance variation due to the temperature allowable while being able to detect short circuit faults. This is handled by separating the estimated resistance into smaller windows and applying a cumulative sum to each window. As a result of this, it will be utilized that a short circuit will result in a fast drop in resistance, which will trigger the lower CUSUM, while a temperature drift will stay within the "target" limits for the considered window. The adjustment is here to determine the correct window size (relative to the convergence rate of the EKF and the sampling frequency.).

Denoting the number of samples in a window by L, the allowable number of standard deviations by $n_{\sigma,w}$ and the detection target standard deviation by σ_w then the cumulative sum for the i'th sample within the window is calculated as (with both $U_0 = 0$ and $L_0 = 0$):

$$U_i = \max\{0, U_{i-1} + x_i - \mu_{w,i} - \frac{n_{\sigma,w}\sigma_w}{2}\} \quad , \; i > 0 \quad (17)$$

$$L_i = \min\{0, U_{i-1} + x_i - \mu_{w,i} + \frac{n_{\sigma,w}\sigma_w}{2}\} \quad , \; i > 0 \quad (18)$$

where μ_w is the estimated average and σ_w the estimated standard deviation. The estimated average $\mu_{w,i}$ is here calculated as the mean of the samples within the given window, i.e. $\mu_{w,i} = \sum_{i=1}^{L} x_i/L$.

Based on the cumulative sum, the indices k and l, denote the points where the cumulative sum exceeds the allowed threshold, i.e. $U_k > n_{\sigma,w}\sigma_w$ and $L_l < -n_{\sigma,w}\sigma_w$ respectively.

The effect is illustrated in Fig. 13, showing a window for both a healthy case and where a fault is detected.

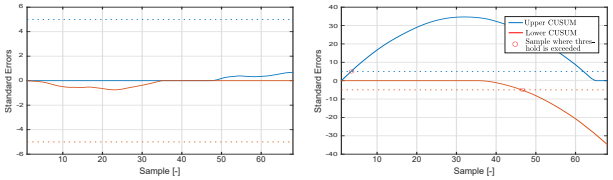

Fig. 13. CUSUM of respectively healthy and faulty window. The sample, where the CUSUM thresholds are exceed are shown by red circles.

From Fig. 13 it may be seen that for the faulty case (shown to the right), the fault happens early in the window because the upper CUSUM early exceeds the threshold, i.e. the mean is well below the estimated samples in the beginning. Therefore as a short circuit will result in a resistance drop, and the mean is based on the entire window, it is the upper CUSUM value, which is of interest. Based on this detection index, a new vector, $\alpha = [\alpha_1 \dots \alpha_j \dots \alpha_n]$, is generated, where α_j stores the detection samples for the j'th window and n is the number of windows considered. In the case no detection occurs, the window length is stored instead, i.e. for the j'th window: $\alpha_j = \min\{k, L\}$, where k was determined from the first application of the CUSUM. The size of α is determined by the number of windows analyzed and may, for online implementation, be adjusted according to the used sample rate, window length and EKF convergence rate. In the case of external cooling of a coil, this may also need to be considered when adjusting the window length and number of windows to avoid triggering false positives.

Based on the generated vector, α, a second CUSUM is applied, similar to above, but where the target mean is equal to the window size, L, and the allowed number of standard deviations is $n_{\sigma,\alpha} < 1$. All deviations in the entries in α (from L) will hence be triggered. This results in a triangular shape in the CUSUM chart, as, e.g. seen in the right part of Fig. 14. This is a result of the EKF convergence time being longer than the length of the window. Hence, it is the consecutive windows in the α-vector exceeding the threshold, which yields the triangular shape in the CUSUM plot, indicating a winding short-circuit and not false triggerings due to noise and/or temperature variation. Here, the short circuit fault is related to the first window where the lower CUSUM standard error becomes negative. Therefore, the window size and the first set of CUSUM parameters have to be adjusted to achieve the desired fault detection. Adjusting the window size is here related to the sampling frequency, the settings of the EKF and its convergence rate, and the compromise between fast and robust detection relative to higher numbers of false or missed detections. Therefore, the compromise is to include sufficient samples in the window to capture the drop in resistance without making the window too long. The window length is adjusted to be in the same range as the time it takes for the EKF to converge to a new resistance level, which is around 7 seconds with the current settings. As the system is operated with an update frequency of 10 Hz, small adjustments of the window length has yielded the best results with a window length of 68 samples. The values have been slightly adjusted to filter out false detections caused by noise or fluctuations in the estimated resistance. Contrary, the second set of CUSUM parameters (related to α) just have to be large enough to result in consistent detection regardless of the magnitude of the deviation from the default α_j-value. The used parameters are given in table I.

TABLE I
CONFIGURABLE PARAMETERS AND VALUES USED.

Parameter	L	n_{σ_w}	μ_w	σ_w	$n_{\sigma,\alpha}$	μ_α	σ_α
Value	68	5	$\sum_{i=1}^{L} x_i/L$	0.015	0.1	L	1

V. RESULTS

This section presents and evaluates the results generated by the algorithm through simulations, where faults are artificially injected. The results of different steady-state cases are shown in Fig. 14, showing the results for both a healthy valve and valve where a short-circuit fault of $\Delta R = 0.1\Omega$ is injected at $t = 250\,s$. The fault is artificially injected. The figure shows the α CUSUM chart for both cases. It should be noted that this corresponds to the first winding short-circuit seen, and the method is, therefore, excellent as an early fault detection method before the coil may eventually fail.

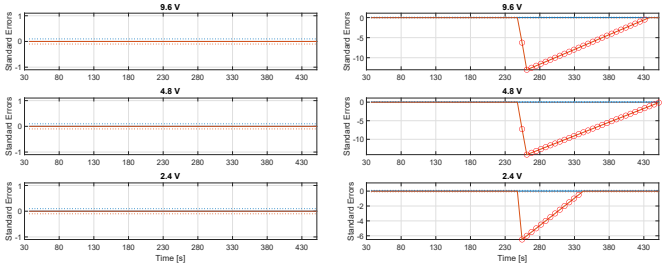

Fig. 14. Windowed CUSUM based fault detection for constant input voltages. Left: no fault injections, right, with fault injected at $t = 250\,s$.

Figure 14 shows that for all fault injections, the fault is detected while no false detections occur for the healthy tests. The utilized parameter adjustment results in a detection time of $< 4\,s$ for all three test cases. Other tests have been made, where the fault is injected at other time instances. However, similar results are obtained for these cases, why they are not included in the paper. However, it should be noted that in special cases, where the fault occurs late in a window, there may be two consecutive detections rather than a single detection, i.e. the second detection shows a larger drop than the first.However, this does not influence the detection. Overall, the FDD algorithm is shown to be effective for steady-state fault detection. Therefore, the following section will focus on detecting faults when a sinusoidal input is applied to the solenoid.

A. Sinusoidal test

In this section, a sinusoidal input is applied to the solenoid. As the estimated resistance is influenced by the input oscillations, due to the mean flux linkage approximation

described in sec. III-B, a 40 sample moving average is applied to the estimated resistance in order to reduce signal variance as shown in Fig. 15. It is the filtered resistance that is then used as input for the CUSUM part of the method.

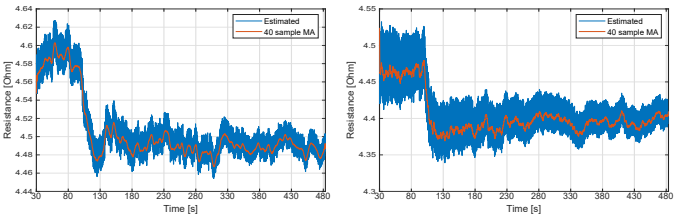

Fig. 15. Moving average resistance filtering. Left $4.8\,V$ right $2.4\,V$.

From Fig. 15 it may also be seen that the variation in the resistance estimation is slightly higher for the low voltage input, as the signal-to-noise ratio is lower. This also means that the required resistance change, which may be consistently detected, will be larger for small signal inputs for the same settings of the EKF-algorithm. The noise sensitivity may, of course, be altered through the settings in the EKF-algorithm. However, this will come at the expense of a longer convergence time and is also limited by the used flux linkage approximation. Longer convergence times may be handled by reducing σ_w to identify lower gradient changes in the resistance, but this comes at the expense of the algorithm being more sensitive to temperature drift. Considering that both the EKF and CUSUM parameters are adjusted to yield a good compromise between convergence time, noise sensitivity and robustness, it is not desirable to significantly alter these in the present case. Instead, the focus has here been on determining the limitations of the algorithm. In general, the change in resistance for a winding short-circuit will be dependent on the specific coil in question and is not known for the specific coil. Therefore the analysis has been made, where the least possible resistance change has been determined, for which faults are consistently detected. For small input signals, the required resistance change is hence 0.17Ω for consistent detection.

In the following the results for an input voltage of $4.8\,V$ are shown in Fig. 16, for which resistance changes down to 0.11Ω are consistently detected, i.e. with a 100% detection rate under the considered conditions. Similar results are obtained for all other tested cases and input voltages when the resistance change detection is set to 0.17Ω.

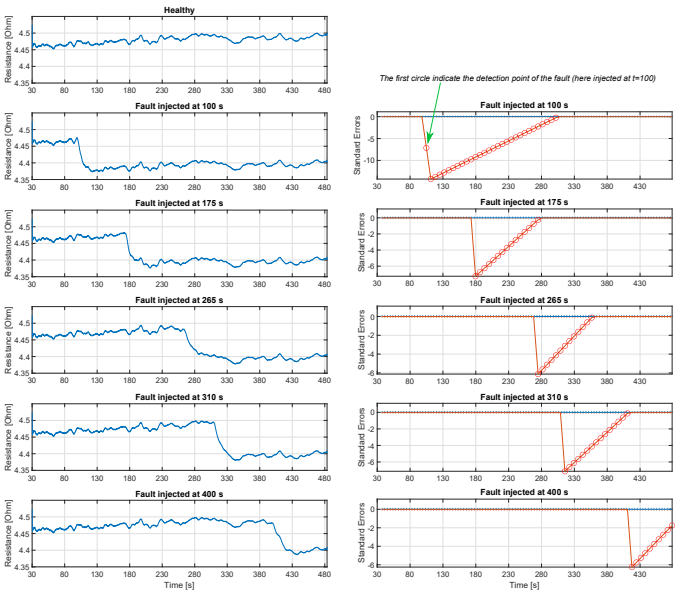

Fig. 16. To the left the estimated resistance, when faults are emulated at different time instances. To the right the corresponding CUSUM output of the algorithm.

Figure 16 shows the results for the $4.8\,V$ sinusoidal test with fault injection magnitudes of $0.11\,\Omega$. The left part of the graphs shows the estimated resistance for the different cases, whereas the corresponding fault detection is shown to the right. The graphs show that the FDD algorithm correctly flags all the injected faults shortly after the injection of the given fault.

VI. Robustness Evaluation

As a typical valve may be subject to some parameter uncertainty, the algorithm's robustness is tested in the following. This is done for the steady-state tests with $\Delta R = 0.1\,\Omega$ and is done with both increasing and decreasing parameters, these being: ± 50 % viscous friction coefficient change, ± 10 % spool spring constant change (these values are chosen based on the magnitude of their influence). The results show that mechanical parameter variations have only minimal impact on the estimated resistance. Thus the performance of the FDD algorithm remains unchanged for both variations. Given that the EKF was adjusted to rely heavily on the position measurement and not the modelled position, these results are expected. For this reason, the robustness is also tested by introducing three different measurement faults to the system. Sensor faults considered are offset, bias drift, and noise faults. The results of sensor offsets are seen in Fig. 17, where the current sensor offset is $\pm 0.05\,A$ while the position offset is $\pm 0.2\,mm$. The values used are approximately 10% of the working ranges. As may be seen in the graphs, these offsets are not enough to impact the detection, but the (actual value of the) resistance estimation in the EKF is influenced by the current offset. The second sensor fault implementation is based on sensor calibration drift. This is implemented as a change in the sensor gain. The sensor gains are 0.95 and 1.05 for both sensors, and the results of these gain changes can be seen in Fig. 18.

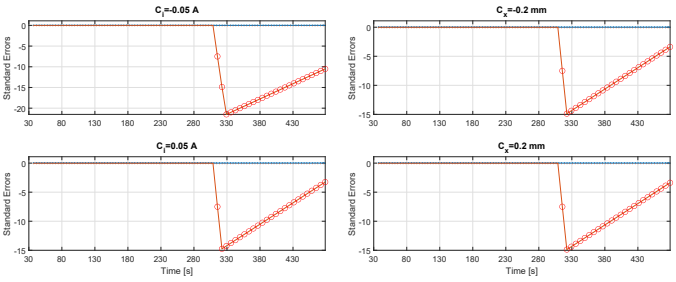

Fig. 17. Fault detection with sensor offset on both current (left) and position (right) sensor with 0.05 A and 0.2 mm offset respectively.

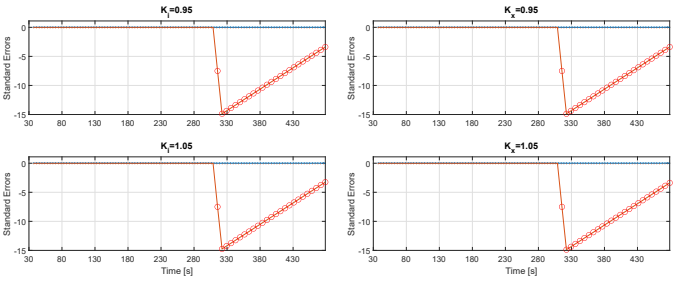

Fig. 18. Fault detection with 0.95 and 1.05 gain on both current (left) and position (right) sensor.

Similarly to what was observed for sensor offsets, the current gain variations have an impact on the estimated resistance magnitude, but not the change in resistance and the fault is still consistently detected. Furthermore, the current feedback is more influential than the position sensor regarding the estimated resistance. Finally, the influence of increased sensor noise is considered in Fig. 19, showing that the impact of the noise magnitude is little to none.

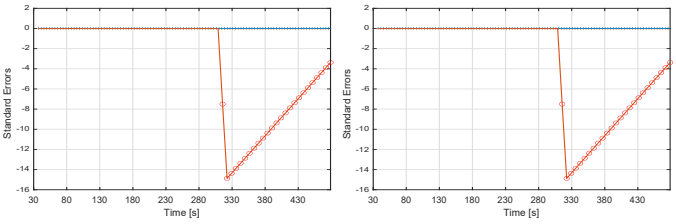

Fig. 19. Fault detection with increased sensor noise added on both sensors. The noise added to the current (left) and position (right)sensors are $\sigma = 2\sigma_i$ and $\sigma = 5\sigma_x$ respectively.

From the above, it may be concluded that the FDD algorithm is robust within the considered margins but will, of course, be limited if much larger parameter variations are encountered. Furthermore, since the FDD algorithm relies entirely on the estimated resistance gradient, and therefore the current gradient, it is also expected to be weak against variations in the flux linkage approximation.

VII. Experimental Fault Injection and Detection

To experimentally validate the algorithm, it is tested in the laboratory, here using a $9.6\,V$ steady-state input. As it is not possible to directly alter the solenoids to create a short circuit, a fault is emulated by connecting a resistor, R_e, and a switch in parallel with the coil as shown in Fig. 20. Closing the switch will hence emulate a winding short-circuit as the equivalent resistance of the system drops. The resistor connected in parallel has a resistance of $180\,\Omega$, resulting in an emulated resistance drop of $0.12\,\Omega$.

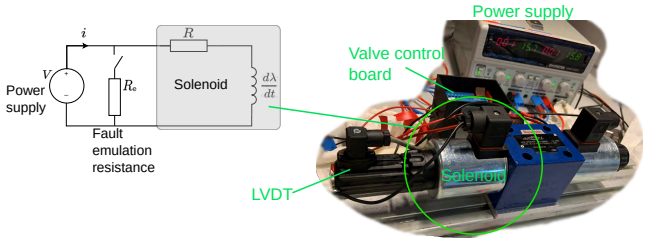

Fig. 20. Electric circuit for fault emulation and the actual valve in the laboratory. The NI-DAQ-equipment and fault emulation circuit is located outside the picture.

As a result of emulating a winding short-circuit through an external resistor, the system is significantly more noise sensitive, as seen from Fig. 21 where a measurement of the current is shown when a fault is injected.

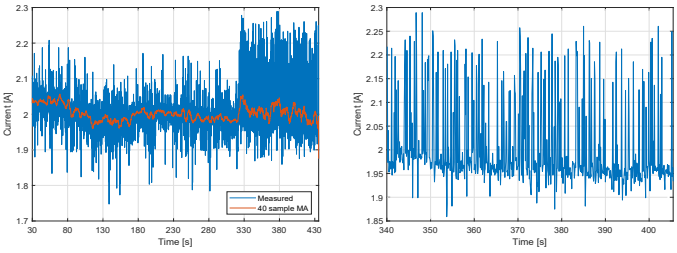

Fig. 21. Measured current, when fault is injected at $t \approx 320\ s$.

From Fig. 21 it may be seen that the current is much more oscillatory with fluctuations of approximately $0.1\,A$. Beyond that, the sensor noise is also significantly larger than seen for the previous steady-state tests with $\sigma_i = 15.4 \longrightarrow 42.8\,mA$. Furthermore, it can be seen that after the injection of the fault, the sensor noise increases further, resulting in $\sigma_i = 42.8 \longrightarrow 90.3\,mA$. This hence has a direct influence on resistance and current estimation in the EKF algorithm. Usually, this would call for a readjustment of the EKF parameters. However, as seen in the below results (and the robustness evaluation), the algorithm is fairly robust to the increased noise, which shows that the algorithm detects the fault when injected after 323 seconds. However, the increased noise does mean that there is a false positive detection in the beginning. This is not included in the graph, but it may be circumvented by readjusting the EKF parameters. However, as this noise increase will not be present if the fault was not emulated, but the fault was a real winding short-circuit, this has not been done.

From Fig. 22 it may be seen that the current noise is increased significantly from previous tests as a result of the increased sensor noise. More significantly, it may be seen that the fluctuations in the current directly influences the estimated resistance. However, the tendencies are the same as the original conducted test, and it is seen that the proposed FDD algorithm correctly detects the fault injected. Similar results are obtained

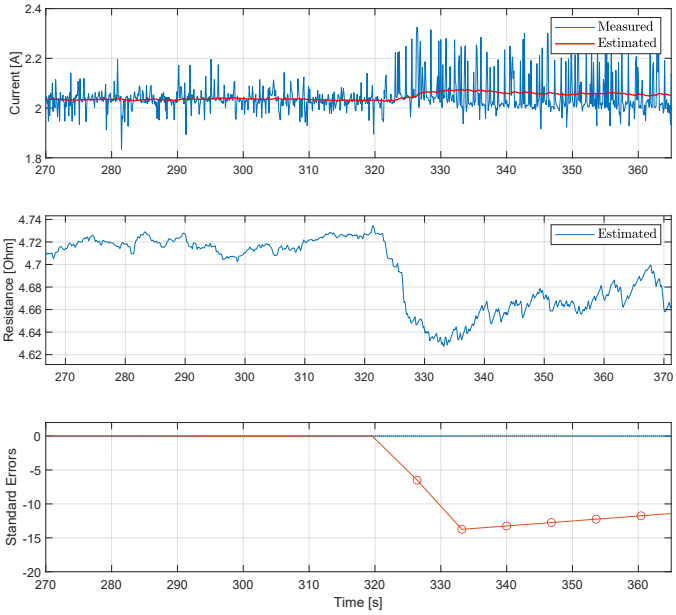

Fig. 22. Estimated current, resistance and windowed CUSUM based on the results of the experimental test.

for other input voltages and when the fault is emulated at different time instances. Overall, this proves the validity of the algorithm and indicates that the FDD algorithm is functional when implemented in the laboratory with injected faults being detected.

VIII. Conclusion

In the current paper, a fault detection algorithm for detecting coil short circuits has been presented. The method, which is temperature independent, is based on an EKF-approach and a windowed dual-CUSUM implementation. The EKF is augmented to estimate coil resistance, and the windowing technique is implemented to accommodate for the variations in the estimated resistance, which are caused by the EKF and the drift of the resistance due to temperature variations. The approach was shown to be able to consistently detect resistance changes down to $0.11\,\Omega$ and $0.17\,\Omega$ for a constant and a sinusoidal varying input signal, respectively, correctly detecting all changes for the considered test cases. This was done even in the presence of parameter variations and increased noise in the system. However, the EKF/method is dependent on a fair approximation of the flux linkage in the solenoid, which directly affects the model precision. Including hysteresis in the flux description may, therefore, improve the results further. Finally, experimental results were presented, showing that the algorithm could detect a resistance drop emulating a coil short circuit despite the noise in the experimental setup is significantly increased by implementing the external resistance used to emulate the fault.

References

[1] R. C. Kryter, *Aging and Service Wear of Solenoid-Operated Valves Used in Safety Systems of Nuclear Power Plants - Evaluation of Monitoring Methods (NUREG/CR-4819 ORNL/TM-12038 Vol. 2)*. Oak Ridge National Laboratory, 1992.

[2] N. J. Jameson, M. H. Azarian, and M. Pecht, "Fault diagnostic opportunities for solenoid operated valves using physics-of-failure analysis," pp. 1–6, 2014.

[3] S.V.Angadi, R. Jackson, Song-yulChoe, G.T.Flowers, J.C.Suhling, Y.-K. Chang, J.-K. Ham, and J. il Bae, "Reliability and life study of hydraulic solenoid valve. part 2: Experimental study," *Engineering Failure Analysis*, vol. 16, pp. 944–963, April 2009.

[4] J. Liniger, S. Stubkier, M. Soltani, and H. C. Pedersen, "Early detection of coil failure in solenoid valves," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 2, pp. 683–693, 2020.

[5] N. Jameson, M. Azarian, and M. Pecht, "Impedance-based health monitoring of electromagnetic coil insulation subjected to corrosive deterioration," 10 2016.

[6] S. Jo, B. Seo, H. Oh, B. D. Youn, and D. Lee, "Model-based fault detection method for coil burnout in solenoid valves subjected to dynamic thermal loading," *IEEE Access*, vol. 8, pp. 70 387–70 400, 2020.

[7] M. Khoshzaban-Zavarehi, "On-line condition monitoring and fault diagnosis in hydraulic system components using parameter estimation and pattern classification," Ph.D. dissertation, Dept. of Mechanical Engineering, University of British Columbia, Vancouver, BC, Canada, 1997.

[8] H. G. Jung, J. Y. Hwang, P. J. Yoon, and J. H. Kim, "Resistance estimation of a PWM-driven solenoid," *International Journal of Automotive Technology*, vol. 8, no. 2, pp. 249–258, 2007.

[9] I. Dülk and Kovácsházy, "Resistance estimation in solenoid actuators by considering different resistances in the pwm paths," *Periodica Polytechnica, Electrical Engineering*, vol. 58, no. 3, pp. 109–120, 2014.

[10] S. Grubic, J. M. Aller, B. Lu, and T. G. Habetler, "A survey on testing and monitoring methods for stator insulation systems of low-voltage induction machines focusing on turn insulation problems," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 12, pp. 4127–4136, 2008.

[11] A. Siddique, G. S. Yadava, and B. Singh, "A review of stator fault monitoring techniques of induction motors," *IEEE Transactions on Energy Conversion*, vol. 20, no. 1, pp. 106–114, 2005.

Henrik C. Pedersen Since 2016 Professor (with special responsibilities) at the Department of Energy at Aalborg University, with a speciality in Fluid Power and Mechatronic Systems. Research areas include modelling, analysis, design, optimization and control of mechatronic systems and fluid power systems in particular. Author of >160 publications within these areas. Head of the section for Mechatronic Systems and program leader for several research projects within this area.

Terkil Bak-Jensen Received his M.Sc. in Energy Engineering with specialization in Mechatronic Control Engineering at Aalborg University, 2021. Currently, he is a PhD fellow at the Department of Energy at Aalborg University, where he is working with the design and optimization of self-contained hydraulic drives.

Rasmus H. Jessen Received his M.Sc. in Energy Engineering with specialization in Mechatronic Control Engineering at Aalborg University, 2021. Currently working as R&D engineer at Schmidt Innovation.

Jesper Liniger Received the PhD degree in reliable fluid power systems from the Department of Energy, Aalborg University, Denmark, in 2018. Since 2018, he has been an Assistant Professor with the Department of Energy, Aalborg University, Denmark. His research interests include modelling, control and fault detection in relation to mechatronic systems with special attention to fluid power systems.