Journal of Power Sources 195 (2010) 4743–4752

Contents lists available at ScienceDirect

Journal of Power Sources

journal homepage: www.elsevier.com/locate/jpowsour

Modeling of an automotive fuel cell thermal system

John Nolan [a,*], Jason Kolodziej [b,1]

[a] General Motors Company, Fuel Cell Activities, 11 Carriage Street, Honeoye Falls, NY 14472, United States
[b] Department of Mechanical Engineering, Kate Gleason College of Engineering, Rochester Institute of Technology, Rochester, NY 14623, United States

ARTICLE INFO

Article history:
Received 5 January 2010
Received in revised form 23 February 2010
Accepted 24 February 2010
Available online 3 March 2010

Keywords:
PEM fuel cells
Modeling
Thermal system
Automotive

ABSTRACT

This work develops an 8th order, non-linear thermal model of an automotive proton exchange membrane (PEM) fuel cell system. Subsystem models are developed from first principals where ever possible and validated against data from a physical system. The system model is validated against data from an automotive 120 kW fuel cell system, with good agreement. Next, a reduced order model is constructed from the full model and the performance of the two models are compared. The reduced order linear model provided an acceptable representation of the full non-linear model.

1. Introduction

As resources become increasingly scarce, government regulatory agencies and the public are demanding non-fossil fuel based transportation solutions. Automakers have responded by electrifying vehicles through the use of hybrid technologies, battery-electric vehicles (BEVs) and fuel cell vehicles. Hybrids, which combine a traditional internal combustion engine with electric motors and storage devices, are popular due to the fact that they return better fuel economy than non-hybrids and do not require a special refueling infrastructure. The downside to hybrids is they still use fossil fuels and release CO_2. BEVs are emission free, but are hindered by the limited range of their battery packs. Since they currently suffer from long recharge times, BEVs are thought to be a solution only for intra-city transportation needs. Fuel cells, in contrast, combine the advantages of both BEVs and hybrids. A fuel cell vehicle does not use fossil fuels or release CO_2, and can be refueled in the same manner as traditional internal combustion engines, negating the range issues associated with BEVs.

Of the many varieties of fuel cells, the proton exchange membrane (PEM) fuel cell has shown the most promise for automotive applications based on their lower operating temperatures and fast start-up times. However, they are not without challenges. In addition to requiring an entirely new hydrogen infrastructure to facilitate refueling, there are many technical challenges to integrating a fuel cell into an automobile. Along with the high cost of precious metals used as catalyst in the membranes and long-term operational durability of the fuel cell stack, there are several difficult control challenges, including thermal management. Precise thermal control of a PEM fuel cell system (FCS) is critical for several reasons. If the fuel cell stack internal temperature rises too high, the membrane can suffer mechanical damage. Elevated temperatures also result in faster catalyst degradation due to high platinum transport resulting in a loss of active area [1]. Additionally, the stack temperature has a large influence on the relative humidity in the system. Humidification of the membrane is critical since proton conduction is directly proportional to the water content in the polymer electrolyte [2]. Proton conduction is the basis of a PEM fuel cell's operation; without conduction, no power is produced. However, too much water in the cell can lead to blocking of the electrolyte pores which results in flooding. Flooding leads to a loss of reactant flow to the cell, which causes both a decrease in cell performance and the possibility of catalyst degradation. It is therefore critical that a balance is struck between removing water produced by the reaction and maintaining proper humidity of the electrolyte.

An additional requirement is to minimize the thermal gradient between the inlet and outlet of the fuel cell stack. A large temperature gradient across the stack results in a large relative humidity change between the inlet and outlet, which introduces mechanical stress into the membrane. In order to facilitate control design, it is desired to have an accurate system model to use in simulations. This work develops a thermal system model and validates that model against actual data from an automotive fuel cell system.

This work attempts to fill a gap in the literature by presenting a thermal model that is validated against physical system data. There has been great interest in fuel cells systems over the last decade, and as a result there is an extensive amount of research available in the

* Corresponding author. Tel.: +1 585 645 8225; fax: +1 585 624 6804.
E-mail addresses: john.p.nolan@gm.com (J. Nolan), jrkeme@rit.edu (J. Kolodziej).
[1] Tel.: +1 585 475 4313.

0378-7753/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpowsour.2010.02.074

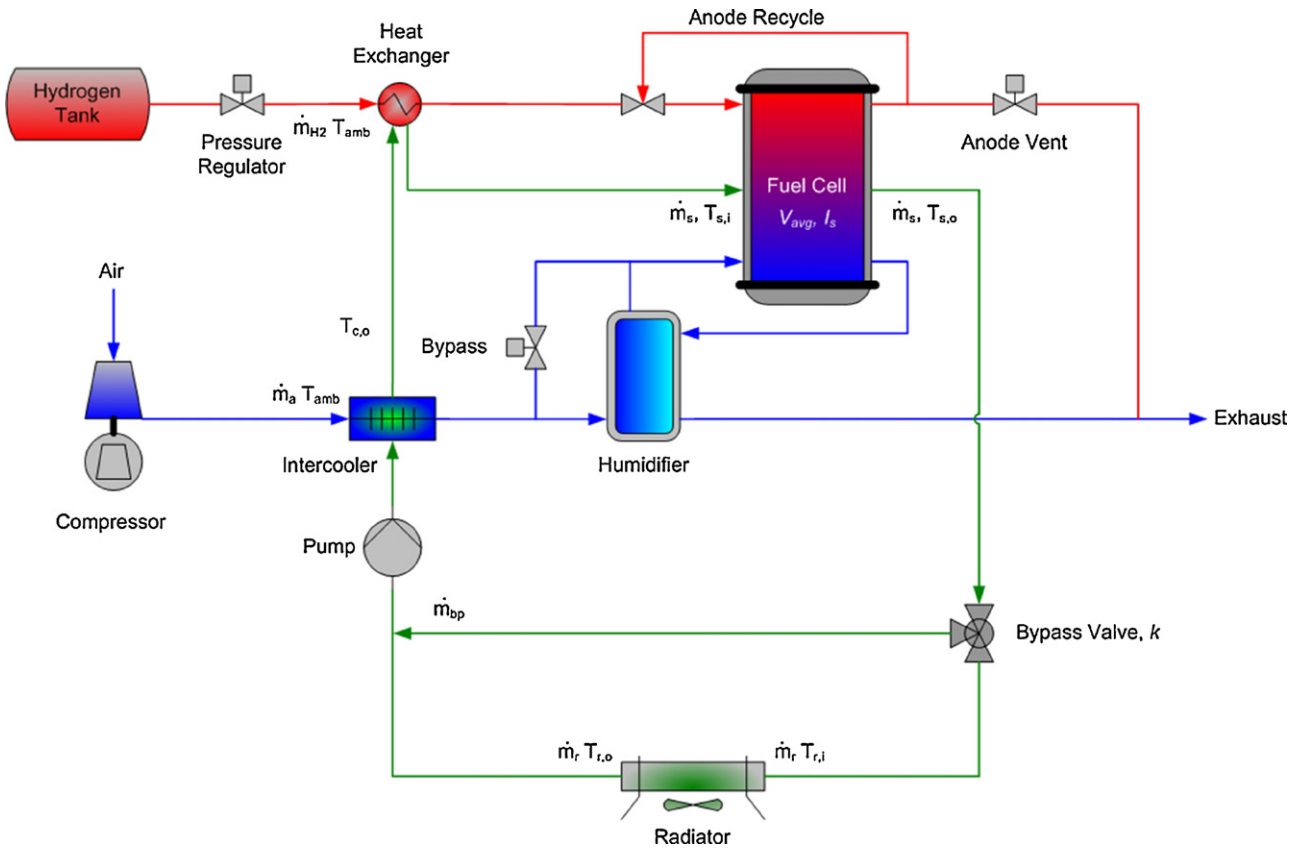

Fig. 1. Typical automotive fuel cell system mechanization.

literature. For the purpose of this research, the principal interests are with models of the fuel cell stack itself, thermal system modeling and the control of fuel cell systems. For an overview of fuel cells relating to automotive applications and the trade-offs associated with the design of the thermal system, the reader is referred to Fronk et al. [3]. Several authors have presented the development of thermal system models, but most lack validation. The cooling system found in this FCS, which uses a bypass valve and a variable flow coolant pump, has been studied recently for automotive internal combustion (IC) applications. Generally, the goal of this research is to reduce fuel consumption. Allen and Lasecki [4] present a good overview of both traditional mechanical and advanced electromechanical coolant systems. They detail both the components of these advanced coolant systems and their benefits to the system. They also consider the need for this type of electromechanical coolant system in future powertrains, including fuel cells.

Gurski [5] presents a dynamic stack model using a lumped capacitance assumption and finite difference methods. Gurski also uses Kroger's [6] graphical, steady-state method for the heat exchanger model used. The main goal of Gurski's work is to quantify the impact of low temperature operation and start-up conditions on efficiency and performance of the FCS. Zhang et al. [7] present the results of a thermal system model that consists of a fuel cell stack, coolant pump and heat exchanger but no bypass. The stack model from this work follows the well-established transient lumped capacitance type, taking into account the heat generated by the chemical reaction in the stack. The heat exchanger model is formed similarly, using an energy balance with an accumulation term. It does not specify how the capacitance of the radiator is determined. The model is validated against data from a fuel cell city bus with good agreement.

Ahn and Choe [1] present interesting work which begins by developing a FCS system model. The system mechanization used is similar to that which is used in this research, however, Ahn and Choe lack an intercooler, anode heat exchanger and transport dynamics. They do include a coolant reservoir which is not present in the current system. The stack model used is the same first order, non-linear model mentioned in previous works cited [7–9]. Two controllers are developed and compared in their performance to maintain the stack inlet and outlet temperature at a fixed value. Both a traditional PI controller and a state feedback LQR controller are presented. The full system model is second order, and due to the non-linearity of the stack model, the model is linearized at the operating point. Ahn and Choe found improved temperature control and reduced power consumption of the coolant pump.

This work will present an analysis of the thermal systems of a 120 kW automotive PEM fuel cell system. Section 2 presents the development of the full, non-linear system model. Section 3 presents the model validation and Section 4 contains the development and validation of the reduced order model.

2. System modeling

A typical FCS for automotive applications is show in Fig. 1 and consists of three main subsystems [10]. The anode subsystem supplies compressed hydrogen from an onboard storage tank to the fuel cell stack. A pressure regulator is used to maintain the pressure inside of the stack. The anode heat exchanger helps warm the incoming hydrogen to the stack operating temperature, since the temperature is well below the stack operating temperature due to the adiabatic expansion of the gas through the pressure regulator. The recycle leg on the anode is used to both humidify the incoming hydrogen as well as increase the hydrogen utilization of the system by recirculating unconsumed hydrogen. Finally, there is a vent valve used to purge water and waste gas (N_2) from the anode.

The cathode subsystem consists of a compressor to supply air to the stack. Air then flows through an intercooler to a water vapor transfer unit to humidify the inlet air stream using the wet cathode exhaust stream. The cathode intercooler is needed due to the fact the temperature of the air exiting the compressor can be up to 50 °C above the stack operating temperature. There is also a bypass leg around the humidifier, used for both stack purges and to control inlet humidity.

The thermal subsystem, which is the focus of this study, regulates the temperature of the stack. The thermal system consists of a bypass valve to direct flow either to the heat exchanger or through a bypass leg, and a pump which circulates the coolant through the system. Transport dynamics are included between the bypass valve and radiator and the radiator and pump to account for the plumbing between these components, which is relatively long and can add significant lag to the temperature response. There are three heat exchangers in the thermal system: the main radiator for dissipating heat to the environment; the cathode intercooler; and the anode heat exchanger. There are three external inputs that affect the thermal system: the system power request; ambient temperature; and vehicle speed. Based on the power request, the stack current, air flow requirement and hydrogen flow requirement is determined. These values affect the heat produced by the stack and the cooling loads caused by the anode and cathode heat exchangers. The ambient temperature primarily affects the performance of the radiator. Finally, the vehicle speed significantly determines the airflow across the radiator which affects its performance. In the following sections, dynamic models of each component are developed and validated against actual system data.

The stack used throughout this research is a 440-cell stack with composite bipolar plates and an active area of 360 cm^2. The net power of the system is approximately 120 kW. This system is typical of those found in automotive applications and is based on the system found in General Motor's (GM) Equinox FC vehicle [11]. The data set used for validation was obtained from test cell validation runs.

2.1. Stack model

The fuel cell stack is the main source of heat generation in the system. For this work, a lumped-parameter approach is employed to model the stack based on a continuously stirred tank reactor (CSTR) model, which results in a non-linear, first order differential equation [9]. This method is similar to Amphlett's [8] but uses a novel optimization technique to estimate the stack thermal mass and coolant volume in one parameter. Performing an energy balance around the stack:

$$\dot{q}_{store} = \dot{q}_{in} - \dot{q}_{out} + \dot{q}_{gen} \tag{1}$$

where

$$\dot{q}_{store} = \frac{d}{dt}(\rho V_{eff} C_p T_{s,o}) \tag{2}$$

$$\dot{q}_{in} = \dot{m}_{s,i} C_p T_{s,i} \tag{3}$$

$$\dot{q}_{out} = \dot{m}_{s,o} C_p T_{s,o} \tag{4}$$

where ρ and C_p is the density and specific heat of the coolant, $T_{s,o}$ and $T_{s,i}$ are the stack coolant outlet and inlet temperatures. $\dot{q}_{in}$ and $\dot{q}_{out}$ is the energy carried into and out of the stack by the coolant and $\dot{q}_{store}$ is the accumulation term. The coolant outlet temperature, $T_{s,o}$, is used in Eq. (2) based on the CSRT model assumption that the temperature within the control volume is well-mixed and uniform. This assumption allows the use of a simpler model form by neglecting the thermal gradient within the stack. The term V_{eff} in the store term is the effective volume, not the actual volume of the coolant in the stack. In much of the other literature, this storage term is a summation of the mass components, including the bipolar plates and end units. The estimated volume is found to account for all of these masses in addition to the coolant volume. This simplifies the model to a first order equation suitable for model-based control applications since it reduces all of the accumulation terms down to a single value. Further, this simplification reduces the state variable to a single, measureable quantity, as opposed to a higher order model which would require multiple temperature measurements or estimates of the state variable. This would drive up cost in an automotive system and is not feasible.

Assuming incompressible flow allows the assumption that the mass flow in equals the mass flow out. Additionally, it is assumed that the specific heat of the coolant is constant, which results in Eq. (5):

$$\frac{d}{dt}(\rho V_{eff} C_p T_{s,o}) = \dot{m}_s C_p (T_{s,i} - T_{s,o}) + \dot{q}_{gen} \tag{5}$$

The generation term, $\dot{q}_{gen}$, results from inefficiency within the stack in generating power due to activation overvoltage and ohmic losses. These losses can be estimated by [2]:

$$\dot{q}_{gen} = (1.25 - V_{avg}) \times n \times I_s \tag{6}$$

where V_{avg} is the average cell voltage, n is the number of cells in the stack, I_s is the stack current and 1.25 is the maximum open circuit voltage of the cell without any losses. This is derived from the Gibb's free energy of the reversible reaction between the hydrogen and the air in the cell. However, due to the activation, ohmic and mass transport losses, the actual voltage in the cell will be less than the reversible, ideal voltage. That difference manifests itself as heat generated within the cell. This heat needs to be rejected to keep the cell at a desired temperature. Combining Eqs. (5) and (6) yields the first order, non-linear stack model:

$$\frac{d}{dt}(\rho V_{eff} C_p T_{s,o}) = \dot{m}_s C_p (T_{s,i} - T_{s,o}) + (1.25 - V_{avg}) \times n \times I_s \tag{7}$$

Taking the coolant density and estimated volume to be constant, Eq. (7) reduces to the final state equation:

$$\dot{T}_{s,o} = \frac{\dot{m}_s}{\rho V_{eff}} (T_{s,i} - T_{s,o}) + \frac{1}{\rho V_{eff} C_p} (1.25 - V_{avg}) \times n \times I_s \tag{8}$$

To estimate the effective stack volume, an optimization routine is utilized. Using coolant data from a fuel cell stack, a sweep of effective volumes was simulated against the actual data and the fit of the data was evaluated using a mean squared error algorithm. The effective volume is varied from 1 to 20 l. Fig. 2 illustrates the result from the optimization. Transient temperature data is used since the effective volume acts as a thermal capacitance to changes in temperature.

The vertical axis in Fig. 2 represents the value of the cost function, J, for that corresponding volume. The minimum value of the function is at an effective volume of 8.5 l, which is larger than the actual coolant volume. This is expected since this effective volume takes into consideration the volume of the coolant along with the mass of the bipolar plates and end units.

To verify the effective volume, the stack model was simulated against system data. The fit of the model using the effective volume is quite good, as shown in Fig. 3 which compares the actual and simulated stack outlet temperatures. While there is some steady-state error in the simulated temperature, the transient sections of the data match very well. Since the transient response is the focus of the model, the small steady-state error can be accepted. The normalized Root Mean Squared Error, $RMSE_n$, is used to evaluate the model fit, with 0 indicating a perfect fit.

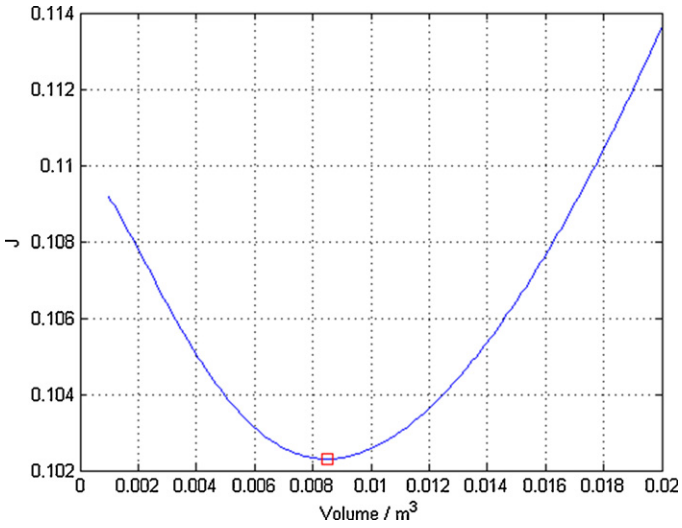

Fig. 2. Volume optimization results.

2.2. Bypass valve

The bypass valve is used to control the flow of coolant between the heat exchanger and the bypass leg. Since the response of the valve is much faster than the temperature dynamics of the system, the dynamic response of the valve is ignored and the valve is considered ideal. The flow through the valve is treated as a linear function of the valve command, with a 0% command, k, indicating all flow goes to the radiator and a 100% command indicating all flow is through the bypass. Eqs. (9) and (10) model the coolant flow distribution:

$$\dot{m}_{bp} = \dot{m}_s \left(\frac{k}{100}\right) \tag{9}$$

$$\dot{m}_r = \dot{m}_s \left(1 - \frac{k}{100}\right) \tag{10}$$

where $\dot{m}_{bp}$ is the flow through the bypass leg and $\dot{m}_r$ is the flow through the radiator.

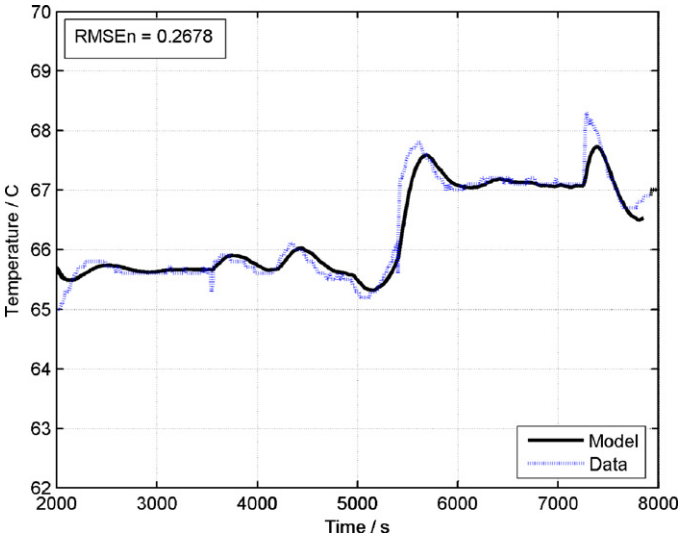

Fig. 3. Simulated and actual stack coolant outlet temperature.

2.3. Radiator

The radiator used in this system is a standard automotive style heat exchanger. The main function of the radiator is to dissipate waste heat to the environment. The model for the radiator is a second order, lumped capacitance model with experimentally derived factors. The model consists of two parts. The first part describes the heat transfer between the radiator itself and the environment as a function of ambient temperature, air mass flow, coolant flow and coolant inlet temperature using an experimentally derived QITD look-up table. This table estimates the heat transfer, q, based on the inlet temperature difference between the coolant and the ambient air. The data for this table was obtained from experiential measurements of the GM fuel cell system. The second part is a lumped capacitance model which describes the heat transfer between the coolant and the radiator. Using an energy balance around the radiator and adding an accumulation term:

$$\dot{q}_{store} = \dot{q}_{in} - \dot{q}_{out} + \dot{q}_{Thermal} \tag{11}$$

with

$$\dot{q}_{store} = \frac{d}{dt}(\rho V c_p T_{r,o}) \tag{12}$$

$$\dot{q}_{in} = \dot{m}_{r,i} c_p T_{r,i} \tag{13}$$

$$\dot{q}_{out} = \dot{m}_{r,o} c_p T_{r,o} \tag{14}$$

where $\dot{q}_{in}$ and $\dot{q}_{out}$ is the energy carried into and out of the radiator by the coolant and $\dot{q}_{store}$ is the accumulation term. The $\dot{q}_{Thermal}$ term describes the heat transfer between the coolant and the thermal mass of the radiator.

$$\dot{q}_{Thermal} = G(T_{r,w} - T_{r,o}) \tag{15}$$

The coefficient G represents the thermal conductance of the radiator and is determined experimentally from test data. Combining Eq. (11) through Eq. (15) and assuming that the specific heat is constant yields:

$$\frac{d}{dt}(\rho V c_p T_{r,o}) = \dot{m}_r c_p (T_{r,i} - T_{r,o}) + G(T_{r,w} - T_{r,o}) \tag{16}$$

Defining the heat capacity as $C = \rho V c_p$ and taking that term to be a constant further simplifies Eq. (16) to be:

$$\dot{T}_{r,o} = \frac{\dot{m}_r c_p}{C}(T_{r,i} - T_{r,o}) + \frac{G}{C}(T_{r,w} - T_{r,o}) \tag{17}$$

Eq. (17) is the final dynamic equation for the coolant outlet temperature of the radiator. The second equation represents the heat capacity of the thermal mass of the radiator. It is a function of the heat transfer between the coolant and the environment and the heat transfer between the coolant and the radiator. Taking an energy balance around the radiator:

$$\dot{q}_{store} = \dot{q}_{in} - \dot{q}_{out} \tag{18}$$

with

$$\dot{q}_{store} = \frac{d}{dt}(\rho V c_p T_{r,w}) \tag{19}$$

$$\dot{q}_{in} = G(T_{r,o} - T_{r,w}) \tag{20}$$

$$\dot{q}_{out} = f_{QITD}(T_{amb} - T_{r,i}) \tag{21}$$

Here $\dot{q}_{in}$ is the heat transfer from the coolant to the radiator and $\dot{q}_{out}$ is the heat transfer to the environment from the experimentally derived QITD table. Combining Eq. (18) through Eq. (21) and taking C as defined above results in:

$$\dot{T}_{r,w} = \frac{G}{C}(T_{r,o} - T_{r,w}) + \frac{1}{C} f_{QITD}(T_{r,i} - T_{amb}) \tag{22}$$

Taken together, Eqs. (17) and (22) describe the dynamics of the radiator.

2.4. Transport delays

Due to the length of plumbing between the stack, the bypass valve and the pump to the heat exchanger, it is necessary to include some transport dynamics in the system model to account for the delay in the temperature response. Neglecting any heat loss in the plumbing and only considering the lag of the temperature, a pure time delay can be used. The delay is a function of the plumbing length and the flow of the coolant.

$$\theta = \frac{V}{\nu} \tag{23}$$

where θ is the delay, V is the volume of the plumbing and ν is the volumetric flow rate of the coolant. Using a 1/1 Pade approximation allows the delay to be modeled in transfer function form as [12]:

$$G(s) = \frac{1 - (\theta/2)S}{1 + (\theta/2)S} \tag{24}$$

For this work, there are two transport delay terms; one for the plumbing from the bypass valve to the radiator and the second from the radiator back to the fluid mixer. It is important to include these delays in the system model because they will have a significant impact on the controller performance, since the delay is not constant. From measurements from the physical system, the plumbing leading to the radiator has an approximate volume of 0.5 l. The piping leading from the radiator has an approximate volume of 0.6 l. Using the expected flow range of 20–196 l m^{-1}, the transport delays can vary between 0.2 and 2 s.

2.5. Fluid mixer

The fluid mixer is the connector where the bypass leg and the heat exchanger leg join together before going to the pump. Physically, the mixer is nothing more than a T junction, but dynamically it is where the two fluid streams join and mix. To model this component, a zero capacitance energy balance is performed:

$$\dot{q}_{out} = \dot{q}_{bp,s} + \dot{q}_{bp,r} \tag{25}$$

$$\dot{q}_{out} = \dot{m}_s c_p T_{m,o} \tag{26}$$

$$\dot{q}_{bp,s} = \dot{m}_{bp} c_p T_{s,o} \tag{27}$$

$$\dot{q}_{bp,r} = \dot{m}_r c_p T_{r,o} \tag{28}$$

Assuming constant specific heat and solving for the mixed stream outlet temperature, Eqs. (25)–(28) can be simplified:

$$T_{m,o} = \frac{\dot{m}_{bp} T_{s,o} + \dot{m}_r T_{r,o}}{\dot{m}_{bp} + \dot{m}_r} \tag{29}$$

Eq. (29) represents the outlet temperature of the fluid mixer in terms of the temperature and flow of each incoming coolant stream.

2.6. Coolant pump

The coolant pump provides flow in the system. A normal pump model would yield the coolant flow rate as a function of the pressure difference across the pump based on the pump's characteristic curve. However, in this analysis, the pressure drops through the system are not modeled to reduce complexity, so a different approach is taken. Two critical parameters of interest in the pump model are the time rate of change of flow to changes in the pump set point and the flow rate of the pump for various set points. Using an empirical modeling approach and step response data of the pump, a first order plus dead time characteristic response is observed.

$$G_{pump} = \frac{Flow}{Speed} = \frac{Ke^{\theta s}}{\tau s + 1} \tag{30}$$

where K is the gain, θ is the time delay and τ is the pump time constant. These parameters can be estimated from pump step change response data. From empirical data, an estimated time constant, τ, of 1.35 s is found. Additionally, the gain K was found to be 1.96, with the dead time easily determined from inspection and found to be approximately 2.0 s.

Admittedly, this method has some drawbacks, namely, it relies on system data to develop the model, so it is not useful for evaluating changes to the system's hardware. Moreover, the flow rate is only an approximation based on data, so fluctuations in system pressure that effect the coolant flow will not be captured. However, for the purpose of this research, these are acceptable trade-offs. Since the goal is to develop a model for this specific system, as long as the model can accurately represent the flow dynamics of the pump in response to changes in set points, then it serves its purpose. Comparing the model to empirical data results in an $RMSE_n$ of 0.1067, indicating excellent agreement between the model and data. Furthermore, the simplifications that these assumptions allow further validate the reasoning for choosing the modeling approach. To capture the pressure dynamics for each component in the system would add an additional state to each component model, almost doubling the model order. Since the essential behavior of the system can be accurately represented using this modeling approach, this simplification is justified.

2.7. Cathode intercooler

In the cathode subsystem, air is compressed and fed into the fuel cell using a compressor, which can cause a significant temperature rise in the air. To bring the temperature of the incoming air to the stack operating condition (80 °C), the air passes through an air-to-water intercooler which uses the coolant on the water side. The effect on the coolant is a function of the air flow, which results from the power request of the system. The cathode flow required for the stack can be determined by the oxygen consumption [2]:

$$\dot{m}_{O_2} = \frac{I_s n}{4F} \tag{31}$$

where I_s is the stack current, n is the number of cells and F is Faraday's constant. This equation gives the mass flow requirement in moles per second. Using the molar mass of oxygen, the equation can be represented in kg s^{-1} as:

$$\dot{m}_{O_2} = 8.29 \times 10^{-8} I_s n \tag{32}$$

Since the oxygen is delivered as air, it needs to be converted to an air basis. Assuming the oxygen content of air is 21%, Eq. (32) is converted to the air mass flow rate:

$$\dot{m}_a = 3.57 \times 10^{-7} \lambda I_s n \tag{33}$$

where the stoichiometric ratio, λ, is included. For a ratio of $\lambda = 1$, the exact amount of oxygen needed for the reaction is supplied to the stack, with no oxygen in the outlet stream. This is not practical because localized starvation can occur, which damages the membrane. More often, a ratio of 2 or higher is used [2]. The temperature rise due to the compressor can be computed from [13]:

$$T_{air,o} = T_{amb} + \frac{T_{amb}}{\eta c_p}(P_r^{(\gamma-1)/\gamma} - 1) \tag{34}$$

where γ is the specific heat ratio (1.4 for air), η_{cp} is the compressor efficiency, and P_r is the pressure ratio for the given mass flow. Using the compressor map supplied by the manufacturer, look-up tables for efficiency and pressure ratio can be constructed versus mass flow. From this data, Fig. 4 shows the relationship between the stack current and outlet temperature for values of lambda from $\lambda = 1$ to $\lambda = 4$, assuming an ambient temperature of 25 °C. The unusual shape of the response is a function of the changing compressor efficiency

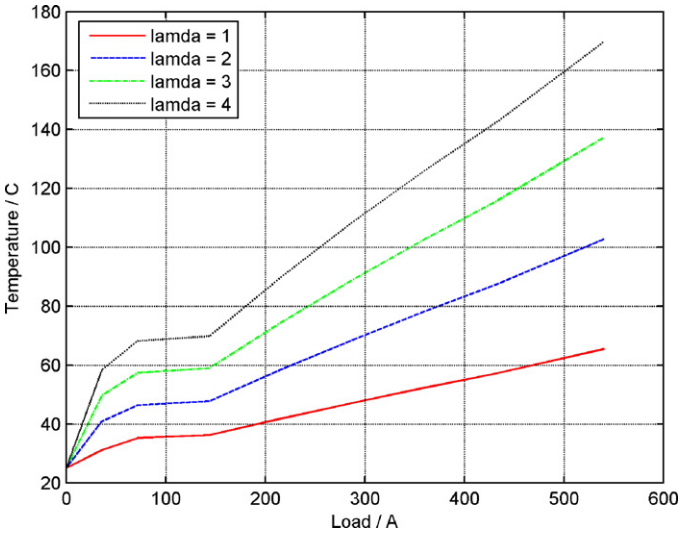

Fig. 4. Compressor outlet temperatures as a function of load for different lambda values.

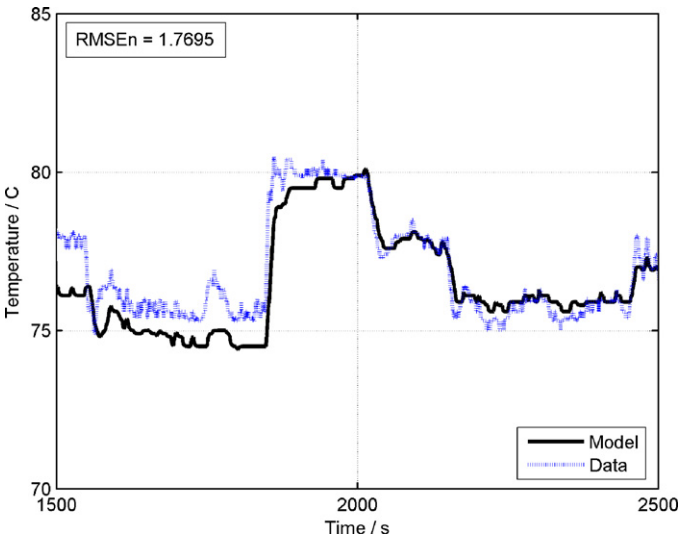

Fig. 5. Cathode heat exchanger outlet temperature model comparison.

at different flow rates and pressure ratios. As shown in the figure, at low loads the intercooler must supply heat to the incoming air to bring it up to stack operating temperature, and at high loads heat must be removed to bring the incoming air down to approximately 80 °C.

The temperature dynamics of the heat exchanger are modeled in a similar fashion to the stack thermal model. Performing an energy balance around the heat exchanger and assuming the specific heat of the coolant is constant results in:

$$\frac{d}{dt}(\rho V c_p T_{c,o)}) = \dot{m}_s c_p (T_{m,o} - T_{c,o}) + \dot{q}_c \tag{35}$$

The generation term, $\dot{q}_c$, represents the heat transfer from the incoming air to the coolant. From the compressor outlet temperature and the coolant inlet temperature, the amount of energy transfer to the coolant can be estimated [14]:

$$\dot{q}_c = \varepsilon \dot{m}_a c_{p,air} (T_{air,o} - T_{m,o}) \tag{36}$$

where ε is the effectiveness of the heat exchanger. From an analysis of the heat exchanger provided by the manufacturer, the effectiveness can be reasonably approximated as 0.88 for all operating conditions. Combining Eqs. (35) and (36) and assuming the coolant volume and density to be constant yields the first order, non-linear heat exchanger model:

$$\dot{T}_{c,o} = \frac{\dot{m}_s}{\rho V}(T_{m,o} - T_{c,o}) + \frac{\dot{m}_a c_{p,air}}{\rho V c_p}(T_{air,o} - T_{m,o}) \tag{37}$$

Fig. 5 shows the simulated heat exchanger coolant outlet temperature compared to the actual outlet temperature. From the figure it is clear there is good agreement between the simulated and actual temperatures, with an $RMSE_n$ value of 1.7695.

2.8. Anode heat exchanger

The onboard fuel storage system contains gaseous hydrogen in a tank and controls the flow of hydrogen into the stack using a regulating control valve. Due to the high pressure of the gas in the tank and the expansion it undergoes as it passes through the valve, the temperature of the incoming gas is very low. Injecting cold hydrogen has a negative impact on the stack performance by significantly changing the relative humidity of the membranes near the anode inlet, so it is necessary to preheat the hydrogen before it enters the stack. To accomplish this, an air-to-water heat exchanger is selected which uses the coolant to warm the hydrogen. This heat exchanger is a shell and tube type, with the coolant as the shell fluid and the hydrogen as the tube fluid. At high power, the flow of hydrogen is significant and can cause an additional disturbance to the coolant control. At low power, the flow of hydrogen is small, and this heat exchanger does not affect the coolant temperature significantly. A simplified modeling scheme is used for the anode heat exchanger as compared to the coolant radiator. Here a static effectiveness-NTU model is used, where NTU stands for *number of transfer unit* and is a dimensionless parameter commonly used in heat transfer analysis. This model was selected due to the heat exchanger being much smaller than the coolant radiator and thus is only a small disturbance on the coolant system. The heat transfer in the heat exchanger is given by [14]:

$$\dot{q} = \varepsilon C_{\min}(T_{c,o} - T_{\mathrm{H}_2,i}) \tag{38}$$

with

$$C_{\min} = c_{p,\mathrm{H}_2} \dot{m}_{\mathrm{H}_2} \tag{39}$$

where ε is the efficiency of the heat exchanger and is estimated from [14] to be 0.85 for a shell and tube heat exchanger with two tube passes, which closely approximates the anode heat exchanger. Using a lumped-parameter assumption, the heat transfer from the coolant is given by:

$$\dot{q} = \dot{m}_s c_p (T_{c,o} - T_{a,o}) \tag{40}$$

Equating Eqs. (38) and (40), and solving for the anode heat exchanger coolant outlet temperature yields:

$$T_{a,o} = T_{c,o} - \frac{\varepsilon c_{p,\mathrm{H}_2} \dot{m}_{\mathrm{H}_2}}{c_p \dot{m}_s}(T_{c,o} - T_{\mathrm{H}_2,in}) \tag{41}$$

The hydrogen mass flow rate can be calculated by the consumption rate based on the stack load [2]. The hydrogen consumption for a single cell is:

$$\dot{m}_{\mathrm{H}_2} = \frac{I_s}{2F} \tag{42}$$

Multiplying by the number of cells in the stack and converting from mol s^{-1} to kg s^{-1} yields:

$$\dot{m}_{\mathrm{H}_2} = 1.05 \times 10^{-8} I_s n \tag{43}$$

Taking this result and assuming the hydrogen inlet temperature is equal to the ambient temperature and letting $\alpha = 1.05 \times 10^{-8}$

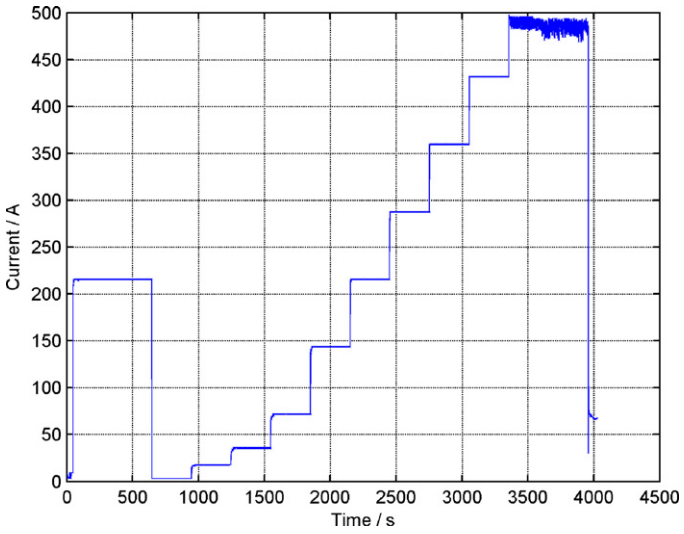

Fig. 6. Validation polarization curve load profile.

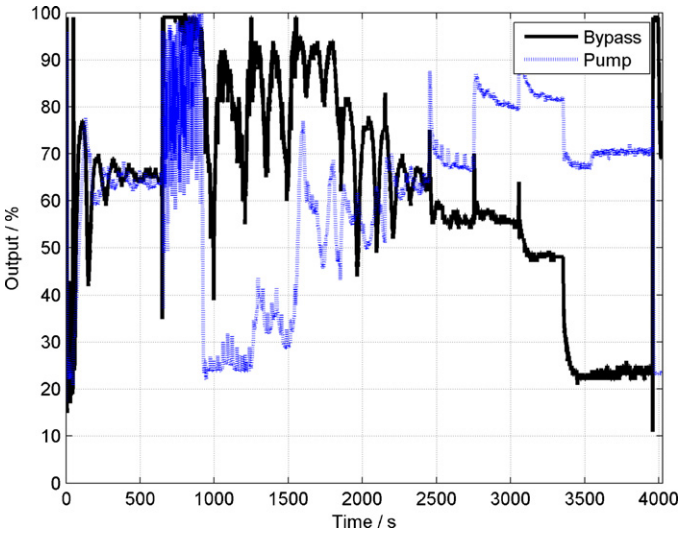

Fig. 7. Validation test actual actuator commands.

reduced Eq. (43) to:

$$T_{a,o} = T_{c,o} - \frac{\varepsilon C_{p,H_2} \alpha I_s n}{c_p \dot{m}_s}(T_{c,o} - T_{amb}) \quad (44)$$

2.9. System equations and assumptions

The final equations used to simulate the FCS are listed above as Eqs. (8), (17), (22), (29), (30), (37) and (44). Several assumptions are made in developing these models. First, it is assumed that there is no heat loss or gain to the environment from piping. Second, any place where transport delays were neglected, it is assumed that the coolant volume between the components is small and the dynamics can be neglected. It is assumed that there is no temperature change in the coolant across the pump. Also, as stated before, pressure drops across components are neglected. This assumption limits this models ability to be used to evaluate component changes, but since this work is interested in modeling a specific system, this is an acceptable trade-off in light of the reduced model complexity.

3. Model validation

For the system model to have value for simulation and control design use it must be validated. System data taken from a polarization curve is used to validate the non-linear model developed in Section 2. A polarization curve is a common test used to measure a fuel cell's performance. The test consists of a series of fixed load points representative of the operating range the system expects to see. The model developed for this work is validated against data from the automotive FCS described in Section 2. Fig. 6 shows the load profile of the polarization curve used to validate the model and Fig. 7 shows the commanded set points for the bypass valve and coolant pump from the physical system. Figs. 8 and 9 show the stack outlet temperature and temperature differential, respectively, for both the physical system and the model, with Fig. 10 showing the error.

The model shows good agreement for both $T_{s,o}$ and ΔT. The normalize $RMSE_n$ is 1.0658 for the coolant ΔT, which is extremely good since a value of zero would indicate a perfect fit. The $RMSE_n$ for the coolant temperature is 4.7385, which also indicates a very good fit. As Fig. 10 shows, the differential error is within $\pm 1^\circ$ over the entire test. The outlet temperature error shows slightly more error, especially at some of the transients. Closer inspection reveals the model

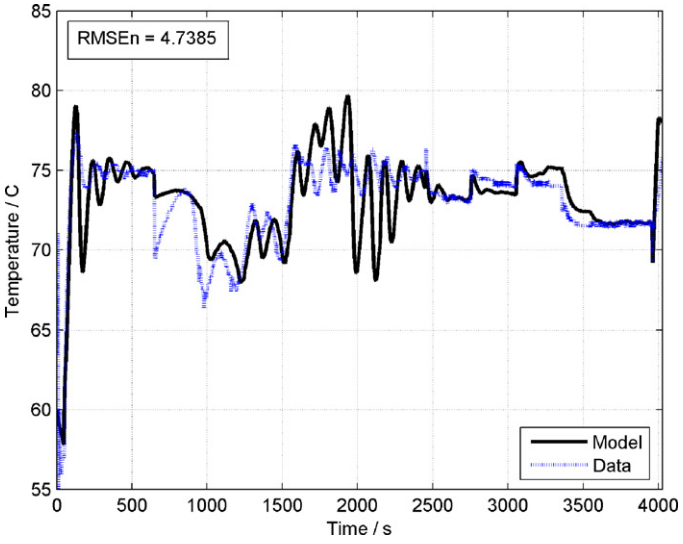

Fig. 8. Validation test stack outlet temperature, measured versus modeled.

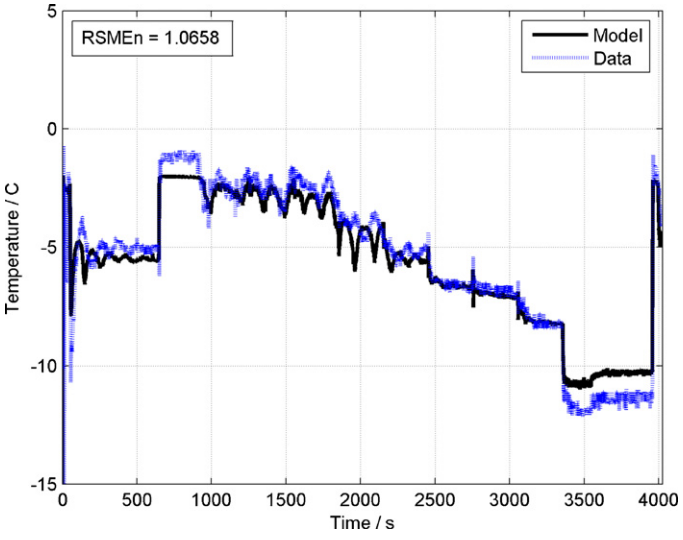

Fig. 9. Validation test stack coolant ΔT, measured versus modeled.

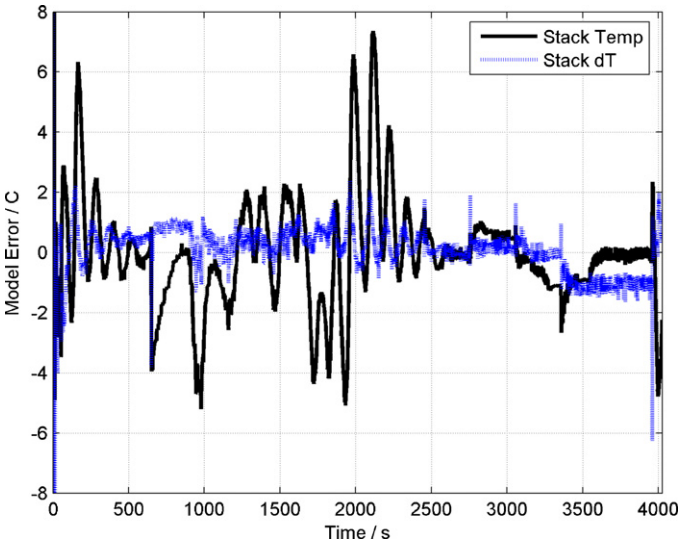

Fig. 10. Model validation error.

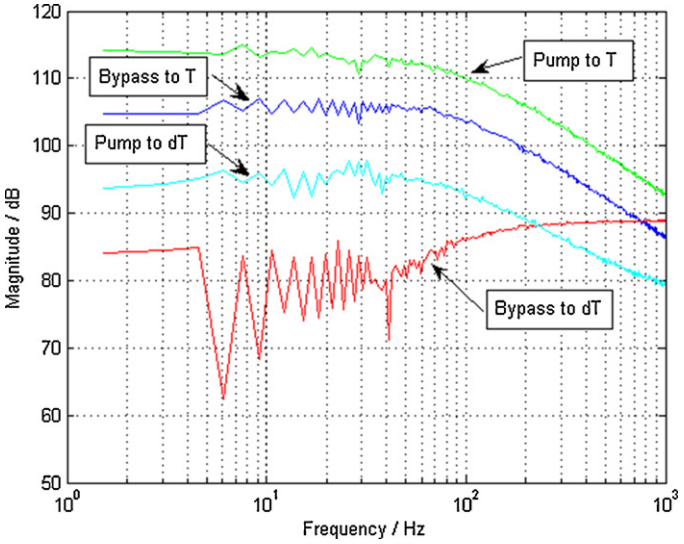

Fig. 12. Model frequency response magnitude at 0.2 A cm^{-2}.

is more sensitive to changes in the bypass set point than the actual system, which exhibits more dampening. This can be explained by the modeling assumptions used. Since pressure drops through the system are neglected, simplifications are made to the coolant pump and flow models. This results in the coolant flow changing instantaneously through each component, which is not realistic. This could be improved by adding lag to the flow in each modeled component; however the error does not degrade the model's performance to warrant the change, since the magnitude and shape of the transient response does show good agreement.

Fig. 11 shows a close-up of a transient for the coolant outlet temperature. This figure shows the good agreement between the model and the actual system in terms of the transient response. Overall, the model provides a very good representation of the physical system, both in terms of steady-state magnitude, and more importantly, in transient response characteristics.

Along with the validation of the model outputs versus system data, the frequency response function (FRF) is generated for the system. The FRF is useful to evaluate the model, and the system response to varying input signals. Additionally, the FRF can be used in some frequency design controller methods. This analysis is conducted around three different load cases; 0.2, 0.6 and 1.2 A cm^{-2}.

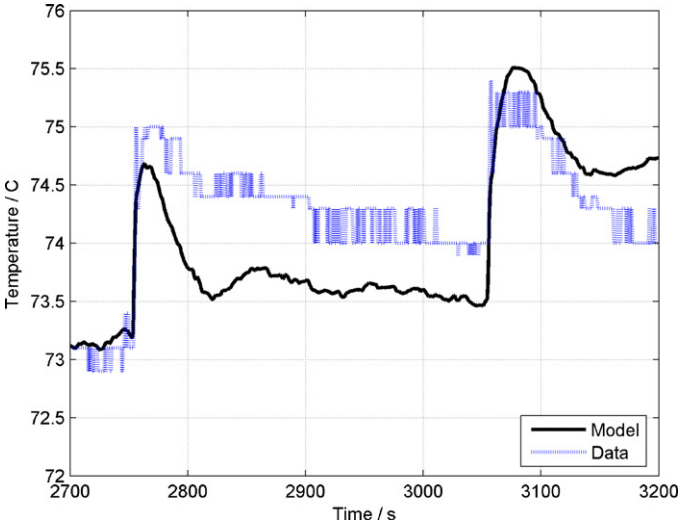

Fig. 11. Model validation temperature response.

These load cases are selected since they are good representations of a FCS at low, mid and high power. Figs. 12–14 show the magnitude response for the FRF at each load case.

An analysis of these plots yield important information about the system. For the two higher load cases, the largest static gain is between the bypass valve and the stack outlet temperature (Pair 1), followed by the pump and the stack outlet temperature (Pair 2), the pump and the temperature differential (Pair 4) and finally the bypass valve and the stack temperature differential (Pair 3). Pairs 1, 2 and 4 all exhibit a first order response with a roll-off starting around 200 Hz. The fourth pairing's response is more interesting, with a magnitude response similar to a non-minimum phase system. All three plots show similar forms, but with varying low frequency magnitudes. Pairs 3 and 4, which relate to the stack temperature differential, both show some fluctuation in their responses in the 8–60 Hz range, especially in the 0.6 and 1.2 A cm^{-2} load cases. These results also highlight one of the difficulties in controlling this system. The stack coolant outlet temperature has higher gain from either of the actuators than does the coolant ΔT. This indicates both actuators have a stronger effect on the outlet temperature than on the coolant ΔT. The strong coupling between actuator effort

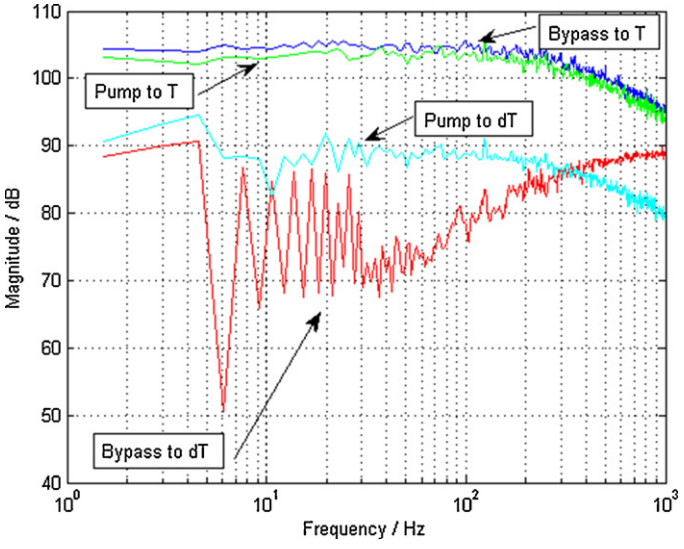

Fig. 13. Model frequency response magnitude at 0.6 A cm^{-2}.

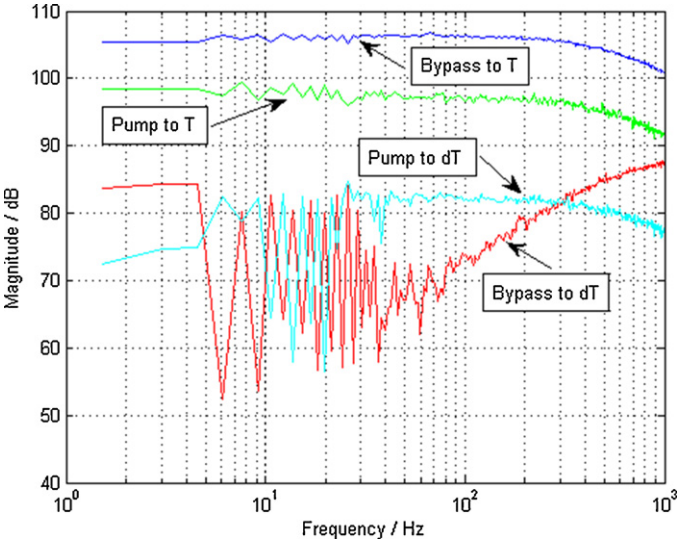

Fig. 14. Model frequency response magnitude at 1.2 A cm^{-2}.

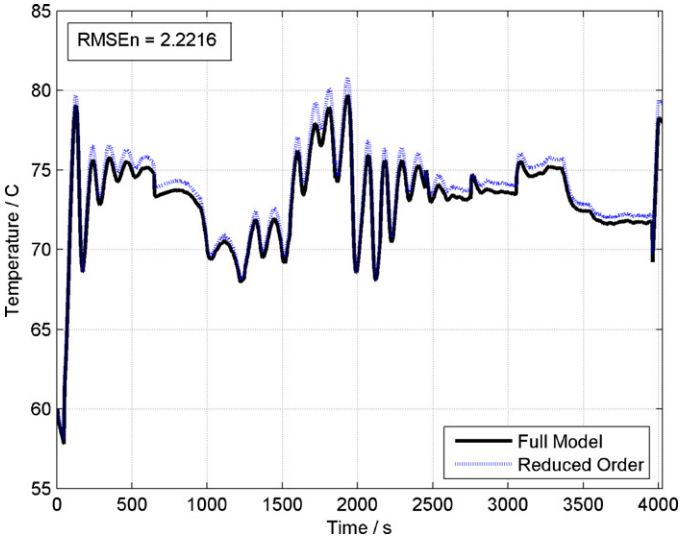

Fig. 15. Stack coolant outlet temperature comparison.

and outlet temperature means any attempt to regulate the ΔT will result in a change in the outlet temperature.

4. Reduced order model

The full non-linear model is an accurate representation the physical system. However, the higher order nature of the model makes it difficult to linearize and use for control design requiring full-state feedback. As such, it is desired to develop a reduced order model. Additionally, it is desirable if the model's states are all physically realizable. The components that contribute the most to the dynamic response of the thermal system are the stack and the radiator. These two components contain three states; the stack coolant outlet temperature, the radiator coolant outlet temperature and the radiator wall temperature. These two components, along with the bypass valve and fluid mixer form the basis of the reduced order model. The pump dynamics are neglected for this model, with a fixed gain being used to convert the pump set point into a flow value. This model is desirable since it has only three states, two of which are easily measured. The third state, the radiator wall temperature can be estimated from other available measurements.

To test the reduced order model, a simulation was run to compare the reduced order model against the full non-linear model using the polarization curve shown in Fig. 8. The results of the simulation are found below in Figs. 15–17. As Fig. 16 shows, the stack coolant ΔT agreement is almost exact between the two models, with an $RMSE_n$ of 0.0377, indicating a high degree of fit. The coolant outlet temperature of the reduced order model has an $RMSE_n$ of 2.2216, due to a 0.5–1 °C offset over most of the simulation, as shown in Fig. 15. This is due to the reduced order model neglecting the anode and cathode heat exchangers. Both of these components effect the coolant temperature, where the cathode intercooler has a larger impact on the coolant, tending to raise the temperature which accounts for the negative offset of the reduced order model.

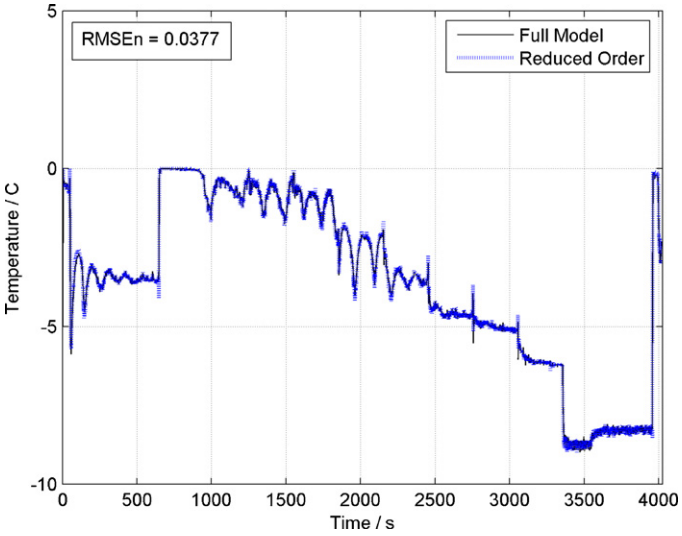

Fig. 16. Stack coolant ΔT comparison.

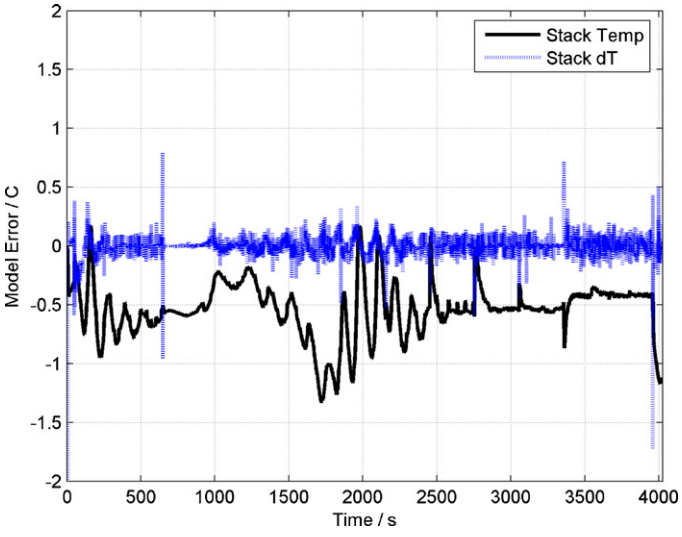

Fig. 17. Reduced order model error.

5. Conclusions and further work

This work has attempted to fill a gap in the published literature of a higher order model of an automotive, PEM fuel cell thermal system. Section 2 presented the development and validation of this model. Each subsystem was modeled using first principal models where ever possible and validated using data from a physical system. A high order, non-linear model was developed and compared

to physical system data. This model was shown to have very good agreement with the physical system, especially the stack coolant temperature delta. The frequency analysis of the full model was also conducted to gain a better understanding of the system's characteristics. Next, a reduced, 3rd-order non-linear model was developed and compared to the full, non-linear model. This reduced order model was found to have very good transient agreement.

This work is the first of two parts. Whereas this work is concerned with the modeling of the system, the follow-up develops a full-state feedback controller to demonstrate performance improvements possible using an advanced control strategy.

References

[1] J.-W. Ahn, S.-Y. Choe, Journal of Power Sources 179 (2007) 252–264.
[2] J. Larminie, A. Dicks, Fuel Cell Systems Explained, Wiley, s.l., 2003.
[3] M.H. Fronk, et al., PEM Fuel Cell System Solutions for Transportation, Society of Automotive Engineers, Warrendale, PA, 2000, 2000-01-0373.
[4] D.J. Allen, M.P. Lasecki, Thermal Management Evolution and Controlled Coolant Flow, SAE International, Warrendale, PA, 2001, 2001-01-1732.
[5] D. Gurski, Cold-start Effect on Performance and Efficiency for Vehicle Fuel Cell Systems, Virginaia Polytechnic Institue and State University, Blacksburg, 2002.
[6] D.G. Kroger, Radiator Characterization and Optimization, SAE International, Warrendale, PA, 1984, 840380.
[7] Y. Zhan, et al., A Model Predicting Performance of Proton Exchange Membrane Fuel Cell Stack Thermal Systems, vol. 24, Applied Thermal Engineering, s.l., 2004.
[8] J.C. Amphlett, et al., Journal of Power Sources (1996) 183–188.
[9] J. Kolodziej, Dynamic modeling and nonlinear control of a PEM fuel cell stack, Journal of Fuel Cell Science and Technology, vol. 4, ASME, s.l., 2007.
[10] D.A. Masten, A.D. Bosco, System Design for Vehicle Applications: GM/Opel, vol. 4, John Wiley & Sons, Chichester, 2003.
[11] General Motors Co. GM Fuel Cell Vehicles. [Online] http://www.chevrolet.com/experience/fuel-solutions/fuel-cell/.
[12] D.E. Seborg, T.F. Edgar, D.A. Mellichamp, Process Dynamics and Control, Wiley, New York, 1989.
[13] J.T. Pukrushpan, A.G. Stefanopoulou, H. Peng, Control of Fuel Cell Power Systems, Springer, London, 2004.
[14] F.P. Incropera, D.P. DeWitt, Fundamentals of Heat and Mass Transfer, Wiley, New York, 2002.