Journal of Manufacturing Processes 23 (2016) 165–174

Contents lists available at ScienceDirect

Journal of Manufacturing Processes

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

Technical Paper

Dynamic modeling of friction stir welding for model predictive control

Brandon S. Taysom[a,*,1], Carl D. Sorensen[a,1], John D. Hedengren[b,2]

[a] Department of Mechanical Engineering, Brigham Young University, Provo, UT, USA
[b] Department of Chemical Engineering, Brigham Young University, Provo, UT, USA

ARTICLE INFO

Article history:
Received 21 April 2016
Received in revised form 4 June 2016
Accepted 6 June 2016
Available online 13 July 2016

Keywords:
Model predictive control
Friction stir welding
Heat source method
Temperature control
Temperature models

ABSTRACT

Controlling temperature in Friction Stir Welding (FSW) is important for consistent post-weld properties. PID temperature control of FSW has been previously implemented once the process is at a quasi-steady state, but has not worked well during either starting transients or during process changes that significantly alter the system dynamics. This work develops models and theories for the application of Model Predictive Control (MPC) to FSW and assesses temperature predication capabilities in simulation.

Two different model forms are developed for MPC and are evaluated in simulation. The first model is a first-order plus dead time (FOPDT) model. The second is the Hybrid Heat Source model that combines the heat source method and a 1D discretized thermal model of the FSW tool. Model parameters are determined by fitting model predictions to weld data. This is done both manually and via optimization-based curve fitting. The models' fits are compared quantitatively by calculating the mean-subtracted SSE (MSSSE) and average absolute derivative error. The manually tuned parameter sets result in a better fit by both metrics for both models. The FOPDT model matches the post-startup-transient data better than the Hybrid Heat Source, and is expected to have superior control in this region of the weld. The Hybrid Heat Source model is expected to have superior temperature control during the startup transient.

1. Introduction

Friction Stir Welding (FSW) is a solid-state hot-deformation process that is used to join two pieces of metal. In this process a tool is rotated and pushed into the seam of two work pieces (the plunge). This action creates heat and softens the metal. Once the metal is sufficiently soft, the tool starts traversing; slowly at first, then transitioning to full speed (the traverse ramp). Once at full traverse speed, the tool continues to travel along the seam of the two pieces and joins them by stirring the metal together. Significant thermal transients are present during the plunge and traverse ramp, and often persist after a constant traverse speed is reached. Because FSW does not melt the weld zone, post-weld properties such as strength, ductility, and fracture toughness are much better than in traditional welding [1,2].

FSW was first implemented by selecting a depth, travel speed, and spindle rotation speed; these methods have been used successfully for many years [3,4]. However, if a weld is run at constant input parameters, the temperature in the weld may fluctuate over time due to transients and disturbances. Because FSW is an inherently temperature-dependent process, if the temperature in the weld is too high or too low, the strength and quality of the weld is negatively affected. In some cases, the welded piece is completely unusable if temperature control is poor [5].

Recently, work has been done to actively control weld temperature [6,7]. Ross [3,8] used a PID controller to control temperature, with tool temperature as the output variable, and power to the rotating tool as the input variable. Ross also showed that directly controlling the motor torque to achieve a set power – and letting the rotational speed float – is an effective means of controlling power. This method of controlling power is consequently used for all models presented in this paper. Marshall [9,10] refined the process of determining PID gains by using a relay feedback test to perform system identification and then used established PID gain rules [11]. Both authors were able to control weld temperatures in both aluminum and steel within one degree Celsius [3,8–10].

PID controllers do not perform well in the presence of large initial errors or in highly transient situations. Due to the significant thermal transients present during and immediately after the start of a weld, controlling temperature with a PID controller has been difficult.

* Corresponding author at: 435 CTB BYU, Provo, UT 84602, USA.
E-mail addresses: scott.taysom@gmail.com (B.S. Taysom), c_sorensen@byu.edu (C.D. Sorensen), johnhedengren@byu.edu (J.D. Hedengren).
[1] Address: 435 CTB BYU, Provo, UT 84602, USA.
[2] Address: 350 CB BYU, Provo, UT 84602, USA.

http://dx.doi.org/10.1016/j.jmapro.2016.06.004
1526-6125/

Nielsen [12,13] has done significant theoretical work in controlling FSW for the sealing of nuclear waste copper canisters. In this application, the machine compliance is non-trivial, and the machine must be capable of welding out-of-round copper cylinders. For these reasons, Nielsen focused heavily on weld depth and axial forces as process variables.

Nielsen developed three different empirical controllers from PRBS weld data. Two controllers were based upon decentralized axial force and depth models, and these in turn were used to tune PI and PID controllers. Using a PI controller, a few welds were performed with good success. Nielsen also developed a full model for non-linear MPC. Nielsen's MPC model is composed of three separate cross-connected models for depth, temperature, and torque. While this controller performed very well against other controllers in simulation, no MPC welds were actually performed. MPC is a proven control technique that has been successfully applied to a variety of fields [14–16] and several different welding processes such as GMAW [17] and GTAW [18–20] and is consequently expected to work for FSW as well.

This study seeks to build upon prior work by developing models for MPC that are able to account for thermal transients, and that consequently have better control during significant transients and disturbances. Physics based models are developed and then tuned using time-series FSW data. The models fit the data well, and it is expected that using both models for MPC will result in good control.

2. Model development

A sufficiently accurate model is necessary for MPC to make accurate move predictions over the control horizon. An accurate model leads to accurate predictions, and good control follows [15].

Attempts to model FSW have ranged from empirical approaches including modified Computational Fluid Dynamics (CFD) programs [21–24]. All attempts have been able to capture some trends, but have been unable to accurately model all important aspects of FSW. However, in order to control temperature, the MPC controller's model needs only to predict the temperature changes relative to changes in system input(s).

In this investigation, two heat transfer based models are considered. The first is a simple first-order plus dead time (FOPDT) model. Both Ross and Marshall noted that the dynamic temperature response of FSW to power step inputs is approximated well by a FOPDT system [8–10]. The second model is a modified heat source method coupled with a thermal 1D tool Finite Element Analysis (FEA). The heat source method uses a history of point heat sources to calculate the temperature at any point and time within a semi-infinite solid, and the hybrid model couples this with a 1D thermal FEA of the tool.

2.1. First-Order Plus Dead Time model

The FOPDT model is based on a simplified thermal view of FSW with different regions of the weld providing for the major heat transfer modes via conduction and advection, as shown in Fig. 1. While these regions do not precisely match reality, an approximate model is often adequate because model parameters are adjustable and can be used to compensate for inaccuracies. Table 1 lists all terms used in the development of the model. In this model, power is the input, temperature is the output, and velocity is used as feed forward variable.

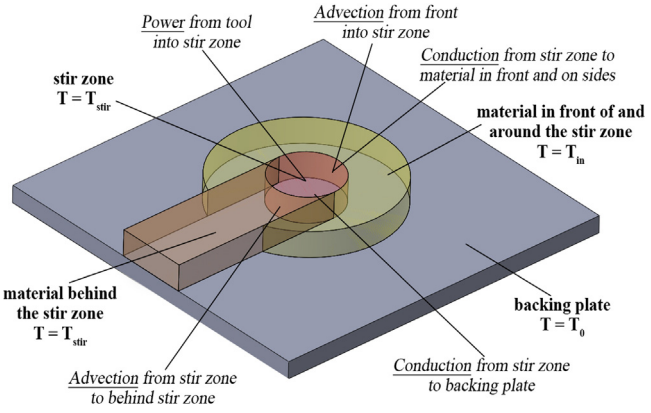

Fig. 1. Regions of the first-order model that interact with or are the stir zone (bold text), and modes and approximate locations of energy transfer between the stir zone and the other regions (underlined and italicized text).

Table 1
Definitions of terms in the FOPDT model.

Term	Definition
A	Area of metal advecting through stir zone (m^2)
E_{in}	Net energy into the stir zone (J)
P	Spindle/tool power (W)
Q_{adv}	Advection power into the stir zone (W)
Q_{cond}	Conduction power into the stir zone (W)
Q_{tool}	Power into the stir zone from tool (W)
T_{stir}	Temperature of the stir zone (°C)
T_0	Temperature of the room/backing plate (°C)
T_{in}	Temperature of metal entering stir zone (°C)
a, b, c, d	Weighting parameters
h_1, h_2	Convection coefficients (W/m^2 °C)
c_1	Power parameter (°C/kW)
c_2	Linear velocity parameter (°C s/m)
c_3	Quadratic velocity parameter (°C s^2/m^2)
c_4	Environment parameter (°C)
c_p	Heat capacity of metal (J/kg °C)
m	Mass of stir zone (kg)
$\dot{m}$	Mass flow rate through stir zone (kg/s)
ρ	Density of metal in the stir zone (kg/m^3)
τ	Time constant of the system/stir zone (s)

The stir zone is predicted for the purpose of control. The basic dynamic temperature equation of the stir zone is:

$$\frac{dT_{stir}}{dt} = \frac{E_{in}}{m\,c_p} \tag{1}$$

There are three separate modes of energy transport/generation in the stir zone: heat generation due to tool rotation, conduction, and advection. Substituting these three terms for E_{in}, and recognizing that the total thermal capacitance of the stir zone is proportional to the time constant of the system, τ, Eq. (1) is equivalent to:

$$\tau\frac{dT_{stir}}{dt} = aQ_{tool} + bQ_{cond} + cQ_{adv} \tag{2}$$

2.1.1. Assumptions

All regions are assumed to be isothermal. The backing plate and weld anvil reservoir temperature, T_0, is assumed to be a constant 25 °C. The backing plate and anvil system are of sufficient thermal mass as to be relatively unaffected by weld thermal energy and consequently heat transfer between them and the room is negligible for the time periods of a friction stir weld. Heat loss up the tool is either ignorable and/or can be lumped together with the backing plate. The geometry and thermal material properties of the regions do not change with temperature or travel speed, and thus the model parameters do not change with these. The measured temperature in the tip of the tool is representative of the temperature of the

stir zone. While these assumptions are over simplifications of FSW, if these simplifying assumptions have negligible effect on model predictions, or if their associated error can be lumped into aggregate empirical model parameters, then the model can useful for predictions in MPC.

2.1.2. Power term

Energy is transferred into the stir zone from the rotation of the tool through direct manipulation of the speed and torque (power) exerted on the metal. The metallurgical conditions in the stir zone mean that deformation of the material is happening at high strains and strain rates [24] due to the rotation of the tool. Under these conditions the flow stress saturates rapidly, causing energy which would be stored via work hardening to be liberated into heat just as quickly via mechanisms such as dynamic recovery and dynamic recrystallization. Therefore the entire energy associated with rotating the tool in the workpiece is assumed to convert completely to heat. The increase of thermal energy in the stir zone via the tool (Q_{tool}) is related to power expended by the motor (P) by the following equation:

$$Q_{tool} = P \tag{3}$$

2.1.3. Conduction term

Conduction occurs from the stir zone to the backing plate and material adjacent to the stir zone. Heat does not flow into the region immediately behind the tool because stir zone material is deposited in the region behind the tool as it advances, resulting in matching temperatures and a locally zero thermal gradient. Only about 3–5% of the energy during the weld is lost by conduction up the tool [25], and consequently this effect is minimal. Furthermore, the tool is cooled by the large FSW machine, and any minor loss through the tool can be lumped together with the backing plate (at T_0) to conduct heat away from the stir zone. The net conduction of heat into the stir zone is thus:

$$Q_{cond} = h_1A_1(T_{in} - T_{stir}) + h_2A_2(T_0 - T_{stir}) \tag{4}$$

2.1.4. Advection term

Advection occurs due to the transport of material in and out of the stir zone. Since the material behind the stir zone is assumed to be at the same temperature as the stir zone, the exit term is 0. The advection term is then:

$$Q_{adv} = \dot{m}c_p(T_{in} - T_{stir}) \tag{5}$$

When taking into account that mass flow rate is dependent upon the traverse speed, the area swept, and the density of the material, Eq. (5) becomes:

$$Q_{adv} = vA\rho c_p(T_{in} - T_{stir}) \tag{6}$$

2.1.5. The T_{in} term

The material in front of the stir zone is heated as a result of conduction from advancing the stir zone, and at steady state is between T_{stir} and T_0. It is not constant; as T_{stir} increases, T_{in} increases as well. As travel speed increases, conduction from the stir zone preheats the material less and T_{in} decreases. The relationship between T_{in}, T_{stir}, and v is modeled as a linear correlation. Eq. (7) is an approximation of the temperature of the colder material in front of the tool [25]. Certainly, Eq. (7) has inherent limits to its validity ,i.e. T_{in} could be negative with a high enough traverse speed. However, this linear approximation should be sufficiently valid as long as the same traverse speed range that the model is tuned to is representative of normal operating conditions.

$$T_{in} = T_{stir} - d \cdot v \tag{7}$$

2.1.6. Complete first-order model

Substitution of Eqs. (3), (4), and (6) into (2) results in:

$$\tau\frac{dT_{stir}}{dt} = aP + bvA\rho c_p(T_{in} - T_{stir}) + c[h_1A_1(T_{in} - T_{stir}) + h_2A_2(T_0 - T_{stir})] \tag{8}$$

Substituting Eq. (7) into Eq. (8), and condensing the parameters, material properties, and weighting terms, results in a model of the form:

$$\frac{dT_{stir}}{dt} = \frac{c_1P + c_2v + c_3v^2 + c_4T_0 - T_{stir}}{\tau} \tag{9}$$

The resulting model is first-order with respect to power which is in harmony with what others have observed [3,10]. Additionally, a time delay (θ) can be implemented via a simple time shift of the temperature results. Eq. (9) can also be formulated such that τ is not in the denominator and T_{stir} is instead multiplied by a parameter c_5. This has the advantage that the parameters to be determined by curve fitting do not interrelate (which is easier for an optimization routine), but the disadvantage that the time constant of the model is no longer explicitly stated. Both expressions are mathematically identical and produce the same dynamic response. Parameters for Eq. (9) are shown in Table 6.

2.2. Hybrid Heat Source FEA model

The Hybrid Heat Source model is created by first calculating the temperature of a point in the plate via the heat source method, and then combining this with a 1D thermal FEA of the tool. As with the FOPDT model, the power of the tool is the input, and temperature is the output.

2.2.1. The heat source method

The heat source model is based on the heat source method, which is a solution to the conduction of an instantaneous point heat source in an infinite solid [26,27]. This solution is:

$$\theta = \frac{Q_{pt}}{8c\rho(\pi a\tau')^{3/2}} \cdot e^{-R^2/4a\tau'} \tag{10}$$

Refer to Table 2 for the definitions of parameters in the heat source model. The parameter τ' is normally presented as τ, but is changed herein to avoid confusion with the FOPDT time delay parameter, τ. The temperature rise of a semi-infinite solid is twice the temperature rise of an infinite solid for the same heat input because it has half the thermal mass of the latter. Eq. (10) is transformed to a moving and variable heat source by using superposition and adding together a series of instantaneous point heat sources of potentially varying intensities and locations:

$$\theta = \frac{1}{4c\rho(\pi a)^{3/2}}\sum_{i=1}^{N}\frac{Q_i}{\tau_i^{3/2}} \cdot e^{-R_i^2/4a\tau_i} \tag{11}$$

For a sufficiently small time period, a moving and variable heat source is effectively stationary if the observation point is sufficiently far away. Thus, heat liberation over a short period of time ($d\tau_i$) may be approximated from the heat liberation rate (q_i) by an instantaneous heat source such that

$$Q_i = q_i d\tau_i \tag{12}$$

Combining (11) and (12) results in:

$$\theta = \frac{1}{4c\rho(\pi a)^{3/2}}\sum_{i=1}^{N}\frac{q_i}{\tau_i^{3/2}} \cdot e^{-R_i^2/4a\tau_i} \cdot d\tau_i \tag{13}$$

Table 2
Definitions of terms in the heat source method.

Term	Definition
Q_i	Amount of heat instantaneously entering the solid at the i^{th} point in time (J)
Q_{pt}	Amount of heat instantaneously entering the solid at a point in space and time (J)
q_i	Rate of heat entering the solid over the i^{th} time period (W)
$q(\tau')$	Rate of heat entering the solid τ' seconds before observation (W)
R	Distance between the observation point and the instantaneous point heat source (m)
R_i	Distance between the observation point and the i^{th} point heat source that occurred τ_i seconds before observation (m)
$R(\tau')$	Distance between the observation point and the point heat source at a time of τ' seconds before observation (m)
a	Thermal diffusivity of the solid; equal to $k/\rho c$ (m^2/s)
c	Heat capacitance of the solid (J/kg °C)
k	Thermal conductivity of the solid (W/m°C)
t	Time relative to observation; zero at the time of observation, negative for the past, and positive for events in the future of the observation (s)
t_{total}	Total time of integration; the minimum of $20a/v^2$ and the elapsed time since the application of a heat source to the solid (s)
θ	Temperature rise at the observation point and time (°C)
ρ	Density of material (kg/m^3)
τ'	Time elapsed between an instantaneous heat source being applied to the solid and the time of observation
τ_i	Time elapsed between the i^{th} differential heat source entering the solid and the time of observation

which can then be written as an integral, from a time of 0 to t_{total}, as:

$$\theta = \frac{1}{4c\rho(\pi a)^{3/2}} \int_0^{t_{total}} \frac{q(\tau')}{\tau'^{3/2}} \cdot e^{-R^2(\tau')/4a\tau'} \, d\tau' \tag{14}$$

For Eqs. (13) and (14), Q, q, and R vary with time and therefore cannot be brought outside the summation/integral. By evaluating either equation over the history, a total rise in temperature, θ, can be calculated at any past, present, or future time and at any point in the semi-infinite solid. By specifying a point in space and time that tracks the stir zone, temperature predictions can be made, which can be used as a basis for an MPC controller.

When evaluating Eq. (14):

- $\tau' = 0$ is the time of observation, or the time at which the change in temperature due to previous events is desired. τ' has a small positive value for events in the recent past, and has a large positive value for events in the distant past.
- Eq. (14) is indeterminate at $\tau' = 0$, so the integral is numerically evaluated by starting τ' at a very small number, such as .00001 s [27].
- Events that are distant both in time and space have an extremely small impact. For a quasi-steady state process with linear movement, events satisfying the criteria $\tau' > 20a/v^2$ can be excluded as their effect is negligible [27].
- The magnitude, derivative, and second derivative of temperature rise due to recent heat inputs is *much* higher in the recent past than in the distant past. In order to accurately calculate the rise in temperature, the recent past must be discretized much finer than the distant past in order to obtain accuracy but still maintain a reasonable numeric evaluation time.

Parameters in this fundamental model could be set to appropriate theoretical values because this model attempts to capture the fundamental physics of heat. However, the heat source method assumes that there is no material flow and that material properties are not a function of temperature, which is *not* the case in FSW. Consequently, some parameters may be altered from their theoretical values to fit the model to FSW process data; this is discussed in greater depth in Section 2.2.5.

Table 3
Definitions of additional terms in the Hybrid Heat Source Model.

Term	Definition
A_j	Area of tool element "*j*" that is in contact with the plate (m^2)
A_x	Area at the boundary of the element and a neighbor/interacting element, element "*x*", usually *n*, *s*, 0, or *c* (m^2)
T_{plate}	Initial temperature of plate/solid (°C)
T_x	Temperature of element "*x*", usually *n*, *s*, 0, or *c* (°C)
V	Volume of material in a tool element (m^3)
c	Heat capacitance of the tool at the node in question (J/kg °C)
h_0	Convection coefficient between the plate and tool nodes (W/m^2 °C)
hc	Convection coefficient between the chilled collar and tool nodes (W/m^2 °C)
k_x	Thermal conductivity of the tool at the boundary of the element and a neighbor/interacting element "*x*", usually *n*, *s*, 0, or *c* (W/m°C)
n	The number of elements in the discretized tool (unitless)
$q(\tau')$	Power dissipated into the *plate* at a given time (W)
$q'(\tau')$	Total power dissipated into the *tool*/*plate system* at a given time (W)
$q_{tool}(\tau')$	Power dissipated into the *tool* at a given time (W)
z_{pt}	Depth of the point of observation in the plate; this affects the value of R (m)
η_p	Power adjustment factor to help fit the model to experimental data by multiplicatively scaling power values (unitless)
δ_x	Distance between centers of neighbor/interacting node, node "*x*", usually n, s, 0, or c (s)
Δt	Total time duration of the current step (s)
Θ	Same as θ (°C)
Θ_0	Rise of temperature at point of observation due to known events from the distant past up until the somewhat recent past (°C)
Θ_1	Rise of temperature at point of observation due to events from the recent past due to energy dissipated into the plate (°C)
Θ'_1	Rise of temperature at point of observation due to events from the recent past *if* 100% of that energy were dissipated into the plate instead of some into the plate and some into the tool (°C)
ρ	Density of tool at the node in question (kg/m^3)

2.2.2. Motivation for developing the hybrid model

The hybrid model is the heat source model with a 1D discretized tool that interfaces with the plate which allows heat transfer between the plate and the tool. Additional terms are needed in excess of the standard heat source model and are defined in Table 3.

The hybrid model addresses the heating of the tool in addition to the plate. This is particularly important at the start of a weld when the tool is not close to a steady state thermal condition. A standard H13 FSW tool can take about 45 kJ to reach a steady state temperature – approximately the same amount of energy that is spent during the entirety of a slow plunge. Ignoring the tool thermal capacitance can result in calculated temperature peaks that are up to 250 °C higher than actual measured temperatures [25].

2.2.3. Assumptions

The heat source method assumes that material properties are temperature independent. There is no material flow anywhere in the plate – energy transfer only happens by convection and never advection. No heat flows up the tool. The tool does not occupy any

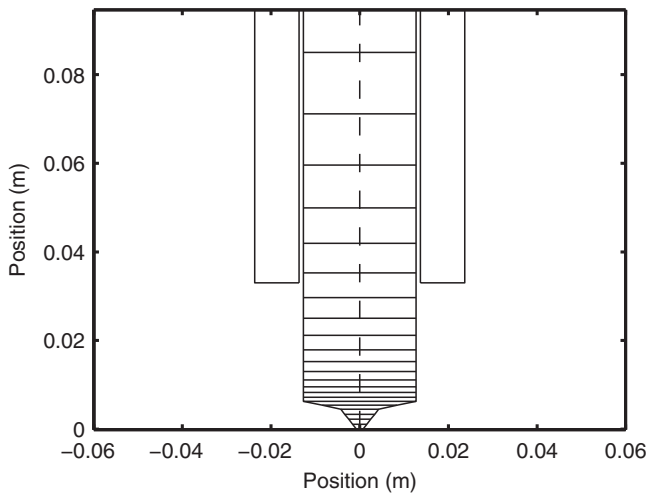

Fig. 2. Meshed tool for transient FEA heat transfer. The elements near the tip are the "south" elements, and the elements toward the top of the tool are the "north" elements. Also pictured on both sides of the tool is the chilled tool collar which conducts heat away from the tool.

space within the plate. The point of observation is equivalent to the point at which the thermocouple is measuring temperature.

The hybrid heat source model has the same assumptions as the regular heat source model, except that heat flows through the tool. Also, all of the tool elements that are in contact with the plate thermally contact the point of observation.

2.2.4. Heat transfer equations for FEA of tool elements

The tool is first discretized into elements. The tool discretion needs to be fine at the tool tip to capture quick changes and larger thermal gradients. The back of the tool does not undergo such dramatic and quick temperature swings and can be accurately modeled with larger elements. The meshing of the tool is shown in Fig. 2.

Eq. (15) shows the heat transfer equations for an individual element. Equations for all elements can be written in numerical heat transfer standard form [28] in preparation for simultaneous matrix solution as:

$$\left(a_p^0 + a_s + a_n + a_0 + a_c\right) T_p - a_s T_s + a_n T_n + a_0 T_0 = a_p^0 T_p^0 + a_c T_c \quad (15)$$

where the temperatures T_p, T_s, T_n, and T_0 on the left hand side of the equation are solved in the matrix solution, and everything else is a constant that goes into either the matrix A or vector b. In Eq. (15), "p" is the principal element, "n" and "s" are the element to the north and south of the principal element, "0" is the pseudo-element of the plate, "c" is the chiller collar, and "p^0" is the previous value of the principal element. Table 4 shows a summary of the a_x values in Eq. (15). Because individual elements may have different geometries, the a_x values may be different for different elements. For example, the value of a_s for element #2 may be 1.4, but might be 1.8 for element #3. If a element is not in contact with the plate and/or the collar, then a_0 and/or a_c is 0, respectively.

Table 4
Values of parameters in Eq. (15) for an individual element.

a_p^0	$\rho cV/\Delta t$
a_s	$k_s A_s/\delta_s$
a_n	$k_n A_n/\delta_n$
a_0	$h_0 A_0$
a_c	$h_c A_c$

2.2.5. Modification of heat source equation and coupling with tool FEA equations

The heat source equation is slightly modified in order to properly couple with the tool equations. This is because the plate loses energy to the tool and consequently this energy does not heat the plate. To more easily discuss terms, Eq. (14) is simplified such that:

$$\int_a^b \psi(\tau)\,d\tau' \triangleq \frac{1}{4c\rho(\pi a)^{3/2}} \int_a^b \frac{q(\tau')}{\tau'^{3/2}} \cdot e^{-R^2(\tau')/4a\tau'}\,d\tau' \quad (16)$$

In this and all other equations, q is the rate of energy into the *plate*. The term q' is defined as the total rate of energy put into the combined plate-tool system at a given time (in other words, the spindle power), and q_{tool} is the rate of energy that goes into the tool:

$$q'(\tau') \triangleq q(\tau') + q_{tool}(\tau') \quad (17)$$

If all of the energy from the spindle enters the plate (instead of some entering the plate and some the tool) the temperature rise over that period of time is:

$$\int_a^b \psi'(\tau')\,d\tau' = \frac{1}{4c\rho(\pi a)^{3/2}} \int_a^b \frac{q'(\tau')}{\tau'^{3/2}} \cdot e^{-R^2(\tau')/4a\tau'}\,d\tau' \quad (18)$$

Based upon Eqs. (14) and (16) the temperature rise at a given point in the plate up to the current time is thus:

$$\theta = \int_0^{t_{total}} \psi(\tau')\,d\tau' \quad (19)$$

which is split into two separate integrals:

$$\theta = \int_0^{\delta t} \psi(\tau')\,d\tau' + \int_{\delta t}^{t_{total}} \psi(\tau')\,d\tau' \quad (20)$$

For past values up until the most recent time step (δt), the amount of energy q into the plate vs into the tool is already known, and thus the first integral of Eq. (20) can be completely calculated. However, θ is unknown for the most recent time step (0 to δt) until it is calculated. If q_{tool} and q' were known for the most recent time step, then q could be calculated.

Three more terms are defined for simplicity of discussion: Θ_0, Θ_1, and Θ_1'. The first is the rise of temperature of the plate due to events in the distant past, the second is the rise of temperature due to events in the recent past, and the third is what the rise of temperature due to events in the recent past would be *if* the entire spindle energy were transferred only into the plate.

$$\Theta_0 \triangleq \int_{\delta t}^{t_{total}} \psi(\tau')\,d\tau' \quad (21)$$

$$\Theta_1 \triangleq \int_0^{\delta t} \psi(\tau')\,d\tau' \quad (22)$$

$$\Theta_1' \triangleq \int_0^{\delta t} \psi'(\tau')\,d\tau' \quad (23)$$

Eq. (20) is rewritten with substitution from Eqs. (21) and (22) to more explicitly show that each of the parts directly constitutes a discrete change in temperature of the plate:

$$\Theta_{plate} = \Theta_0 + \Theta_1 \quad (24)$$

Since θ is linearly proportional to Q as shown by Eq. (10), the following is true:

$$\Theta_1 = \Theta_1' \left(1 - \frac{q_{tool}}{q'}\right) \quad (25)$$

Thus, if the entire amount of power going into the system goes into the tool, then Eq. (25) evaluates to $\Theta_1 = 0$. On the other hand, if

Table 5
Values for parameters in Eq. (30).

α_0	$1 + \Theta_1^{'}/q^{'} \cdot \sum_{j=1}^{n} h_0 A_j$
α_1	$-h_0 A_1 \cdot \Theta_1^{'}/q^{'}$
α_2	$-h_0 A_2 \cdot \Theta_1^{'}/q^{'}$
α_n	$h_0 A_n \cdot \Theta_1^{'}/q^{'}$
β	$\Theta_0 + \Theta_1^{'} + T_{plate}$

no power goes into the tool, then Eq. (25) evaluates to $\Theta_1 = \Theta_1^{'}$. Substituting Eq. (25) into Eq. (24) gives:

$$\Theta = \Theta_0 + \Theta_1^{'}\left(1 - \frac{q_{tool}}{q^{'}}\right) \tag{26}$$

The T_0 element in the FEA equations is also the plate temperature, Θ_{plate}. $T_{1:n}$ are elements in the tool which may or may not be in contact with the plate. Therefore, the energy flowing into a given tool element from the plate is:

$$q_{tool} = \sum_{j=1}^{n} h_0 A_j (T_0 - T_j) \tag{27}$$

Any tool element that is not in contact with the plate will have a contact area (A_j) of 0 and consequently its contribution to the summation is also 0. Substitution of Eq. (27) into Eq. (26) results in:

$$\Theta = \Theta_0 + \Theta_1^{'} \cdot \left[1 - \left(\sum_{j=1}^{n} h_0 A_j (T_0 - T_j)\right)/q_i^{'}\right] \tag{28}$$

If the entire plate is at nonzero initial temperature, T_{plate}, then:

$$\Theta = T_0 - T_{plate} \tag{29}$$

Substituting Eq. (29) into Eq. (28) and rearranging so that each of the temperature terms only appears once in the equation results in:

$$\alpha_0 T_0 - \sum_{j=1}^{n} \alpha_j T_j = \beta \tag{30}$$

where Table 5 lists values for Eq. (30).

With Eqs. (15) and (30), the matrix A and vector b can then be assembled to allow simultaneous solving of the new temperatures in the plate and tool. The values for each individual line come from Tables 4 and 5.

The A matrix is:

$$A = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & & & & \\ -a_0 & a_p & -a_n & 0 & & & & \\ & & & & \dots & & 0 & \\ -a_0 & -a_s & a_p & -a_n & & & & \\ -a_0 & 0 & -a_s & a_p & & & & \\ & \vdots & & & \ddots & & \vdots & \\ & & & & & a_p & -a_n & 0 \\ -a_c & 0 & 0 & 0 & \dots & -a_s & a_p & -a_n \\ & & & & & 0 & -a_s & a_p \end{bmatrix} \tag{31}$$

In Eq. (31), the α_n term on the top row continues all the way to the right side, but equals 0 when there is no contact between a particular tool element and the plate (as is also the case for the elements at the back of the tool). The same is true of the a_0 and a_c terms in the first column of the matrix in Eq. (31).

The b vector is:

$$\vec{b} = \begin{bmatrix} \beta \\ a_p^0 T_p^0 + a_c T_c \\ a_p^0 T_p^0 + a_c T_c \\ \vdots \\ a_p^0 T_p^0 + a_c T_c \end{bmatrix} \tag{32}$$

Similarly to the a_0 term in Eq. (31), if there is no contact between a element and the tool holder/collar, then the a_c term in Eq. (32) equals 0. Alternately, the last row of Eqs. (31) and (32) could be altered to constrain the back of the tool to a fixed temperature.

Similar to Eq. (15), the terms in each row of Eqs. (31) and (32) are respective to each row's principal element. Therefore, the a_p, a_s, a_n, $a_p{}^0$, a_0, and a_c values may vary from one row to the next.

The solution to the system of equations $A \cdot \vec{T} = \vec{b}$ returns the temperatures at the end of a given time step, where the first value in $\vec{T}$ is the plate temperature at the point of observation (T_0), the second value in $\vec{T}$ is the first element in the tool (at the tip), the third value in $\vec{T}$ is the second element in the tool, etc. Solving for $\vec{T}$ can be accomplished by matrix inversion or any other matrix solution method.

Some of the variables and material properties in this model are changed from their theoretical values to account for errors in the model, and others are changed because the actual physical values are unknown. These variables are (1) z_{pt}, the depth of the point in the plate at which the heat source method is used and which also couples the tool to the plate; (2) h_0, the convection coefficient between the tool and the plate; and (3) η_p, the power adjustment multiplicative factor to help fit the model to experimental data. The variable z_{pt} is incorporated in the calculation of R^2 such that $R^2 = z_{pt}^2 + \Delta x^2 + \Delta y^2$, and η_p is incorporated into the power input to the model such that $p_{tool} = \eta_p p_{actual}$.

3. Model parameter estimation by fitting to process data

3.1. Materials, experimental setup, and methods

All welds were performed with a CS4 scroll style tool with a thermocouple placed approximately 2.5 mm from the end of the tool tip. All welds were performed in 6.35 mm (.25″) thick Al 7075-T7 plate. To reduce variability, two separate plates were not welded together; rather welds were performed in the middle of the plate. Thus all welds are technically friction stir processing, not friction stir welding.

Room temperature was approximately 24 °C. All welds were run with active water cooling for the tool holder which chilled the tool to approximately 10 °C before welding. After the plunge, the traverse speed was ramped from 0 m/s (0 in/min) up to 3.8 mm/s (9 in/min) over a distance of 5.1 cm (2 in).

Matlab was used for all computations. The *fmincon* function utilizing an interior point method was utilized to perform optimization-based parameter estimation. An explicit Eulerian ODE solver was used to evaluate the FOPDT dynamic models in all cases. An Eulerian approach was chosen because of its lower computation cost. The Hybrid Heat Source model innately uses an implicit Euler scheme inherent in the above FEA equations to evolve the ODE. Appropriate grid-independence was confirmed for both models.

3.2. Quasi-PRBS for parameter estimation

In order to determine parameters for the models, several welds were performed in 2.44 m (8′) long material. After the start-up

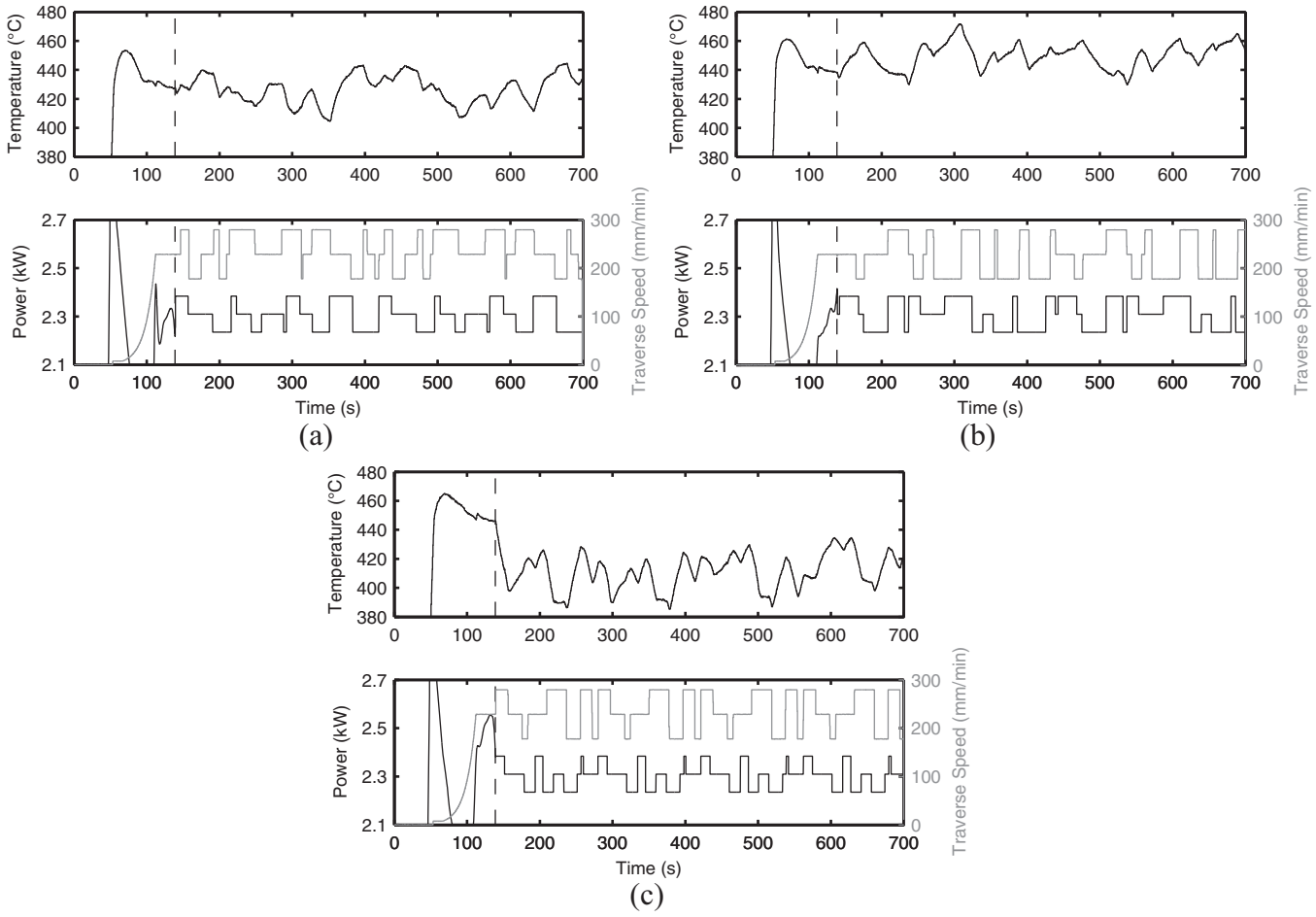

Fig. 3. Temperature, power, and traverse speed for the welds (a) PRBS 002, (b) PRBS 003, and (c) PRBS 004. The vertical dashed line indicates the start of the PRBS section of the weld.

sequence, the heat input and traverse speed were abruptly and semi-randomly changed to one of three possible levels and were held at that level for a random length of time. This approach is similar to a Pseudo Random Binary Sequence (PRBS), but three levels of heat input and traverse speed were used. A standard PRBS uses two levels of the input to obtain maximum system excitation, but because there are only two levels, it is difficult to detect model nonlinearity from the resulting output. While adding a middle level reduces the magnitude of some of the excitations, it enables the detection of nonlinearity. Three different PRBS welds were run for system identification and are shown in Fig. 3. Power and traverse speed sequences were repeated twice in welds 002 and 003, and four times in weld 004.

3.3. Determining model parameters

Parameters for both the FOPDT model and the Hybrid Heat Source model were determined by optimization-based and manual curve fitting. The PRBS 003 and 004 data sets were used to train the models. PRBS 002 was reserved for the testing of the model parameters.

Almost tautologically, a properly functioning optimizer returns the best possible results. An often overlooked yet very important assumption is in the value of the objective function itself. If a poor objective function is chosen, then the results of an optimizer may not give the best results in process control.

A common choice of objective function for an optimizer tasked to determine model parameters via curve fitting is the sum of squared error (SSE) of the output. This is a very useful metric when all the data can be modeled via a single set of model parameters. In this case, both the SSE and average absolute error of the derivative converge to 0; results from the optimizer cannot be improved upon and are in fact the best possible model parameters. However, if the same inputs result in non-trivially different outputs, then a single set of parameters does not exist that unifies all data. In this case, the optimizer cannot drive the SSE close to 0 and must compromise on the different data sets.

Using the SSE as the objective function prioritizes minimizing the model vs actual offset instead of matching model dynamics and gains, this is shown schematically in Fig. 4 for a nearly first-order system. MPC controllers however need correctly modeled system dynamics (derivative information) so that the MPC controller can predict the appropriate amount of corrective action. Furthermore, during MPC control, a constant offset between model predictions and measured values is quickly compensated for with biasing and therefore a constant offset of the model becomes a non-issue. Consequently, optimizing the SSE at the cost of other system dynamics (particularly derivative information) can in some cases produce model parameters that are non-optimal for MPC.

Given this problem, other optimization approaches could be attempted, but each has its own drawbacks. A multi-objective optimization can be employed to minimize both SSE and derivative error. This however returns a Pareto front of values and not a single value, and the ultimate decision is left to a decision maker. Alternately, a weighting factor can be used to combine both objectives and the optimizer will find a single best parameter set. However

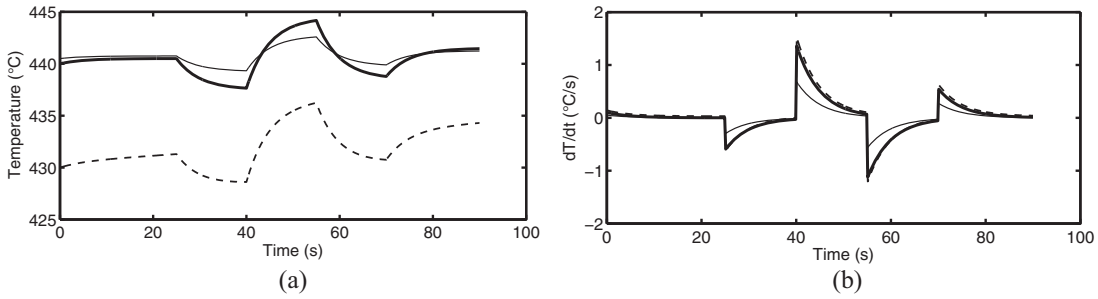

Fig. 4. Schematics of two different models and process data showing (a) temperaure vs time, and (b) the derivative of temperature vs time. An optimizer trying to minimize SSE against data (thick solid line) will pick model 1 (thin solid line) over model 2 (dashed line) as the average error is much lower for model 1, as clearly seen by (a). However, model 2 captures system dynamics much better than model 1 as can seen by the derivative information shown in (b); model 2 is almost perfect while model 1 has errors of ~50%. Model 1 would be expected to perform better than model 2 as the basis for an MPC controller.

Table 6
Parameters for the first-order model.

	c_1 (°C/kW)	c_2 (°C·s/m)	c_3 (°C s^2/m^2)	c_4	τ (s)	θ (s)
Optimizer	215.8	−4.591e4	−4.773e6	7.146	48.09	1.5
Manual	158.9	−1.663e4	−5.397e5	4.918	16.39	1.5

Table 7
Parameters for the Hybrid Heat Source model.

	z_{pt} (mm)	h_0 (W/m^2 K)	η_p (unitless)
Optimizer	0.840	4690	0.567
Manual	4.57	600	1.40

Table 8
Other values and parameters needed for the Hybrid Heat Source model.

h_c (W/m^2 K)	ρ (kg/m^3)	c (J/kg K)	k (W/m°C)
200	2800	1100 [29,30]	140 [29,30]

the weighting factor is manually chosen and the returned optimal value is dependent upon the weighting factor. Since there is no easy way to completely avoid subjectivity with these methods, the SSE was used because it is impartial and can produce acceptable results in many circumstances.

Manual parameters were determined by trying to balance (1) minimizing the SSE, and (2) matching system dynamics such as the temperature-time derivative. A set of parameters was first chosen which resulted in an output of a roughly correct level of output. Parameters were individually modified in order to better capture the observed trends, and then the next parameter was modified. This was performed recursively until additional changes to any parameters did not improve the output results.

Two metrics are used in order to quantify how well the manual and optimization methods predict the system response. For the first metric, the mean of actual and predicted temperatures is subtracted from the predicted temperatures, and then the sum of squared errors is calculated to determine a "mean subtracted sum of squared error" (MSSSE). This quantitative metric is used to determine how well the prediction matches the observed data if the offset was zero. A second metric is the average absolute error of the derivatives of the modeled vs predicted temperatures.

3.3.1. FOPDT model

For the FOPDT model, the parameters c_1, c_2, c_3, c_4, and τ from Eq. (9) were determined by fitting the model to the PRBS data and are presented in Table 6. Additionally, a time delay of 1.5 s was manually determined in both cases and added to the model to account for the observed time delay; the model output is relatively insensitive to this parameter. The FOPDT model is not of the correct form to predict the transient trends seen immediately after the weld's plunge: velocity increases and power decreases, but temperature remains relatively stable and does not drop dramatically. Consequently only data after the startup transient is used for curve fitting of the FOPDT model.

The temperature predictions resulting from both sets of parameters are shown against the actual temperature for the PRBS 002 weld in Fig. 5. The FOPDT model with manually determined parameters matches the system dynamics much better than the FOPDT model with optimizer determined parameters. Both the MSSSE and average absolute error of the derivative indicate that the FOPDT model with manual parameters is more representative of the FSW process and is consequently likely to perform better in MPC applications.

3.3.2. Hybrid Heat Source model

The parameters to be determined for the Hybrid Heat Source model are z_{pt}, h_0, and η_p, and are shown in Table 7. Several additional parameters are necessary in order to implement the model but are not changed to fit the model to weld data. These are either known (such as material properties of aluminum), or are unknown yet moderate changes do not significantly affect model predictions (such as h_c). These parameters are presented in Table 8.

Unlike the FOPDT model, the Hybrid Heat Source model predictions include input data during the startup transient (the plunge and feedrate ramp). This is necessary because the Hybrid Heat Source model captures plate heating during the plunge, which in turn affects the weld temperature for a period of time after the plunge. While the events during the plunge are important for the Hybrid Heat Source model's predictions of tool and workpiece heating, the model's output during the startup is not used to train the model for post-startup control. If it is desired to use this model for control during the startup transient, then the model should be trained to startup data such as the feedrate ramp.

The temperature vs time predictions for the optimizer determined and manually determined parameters for PRBS 002 are shown in Fig. 6. Again, the model with manually determined parameters is better than the model with optimizer determined parameters. Again, both the MSSSE and average absolute error of the derivative indicate that the manual parameters are more representative of the system and consequently are likely to perform better in MPC applications.

4. Discussion

Both the FOPDT and Hybrid Heat Source models best capture the dynamics of the system – including the magnitudes of step changes – using the manual parameter sets. For both models, both the MSSSE and average absolute error of the derivative are lower for

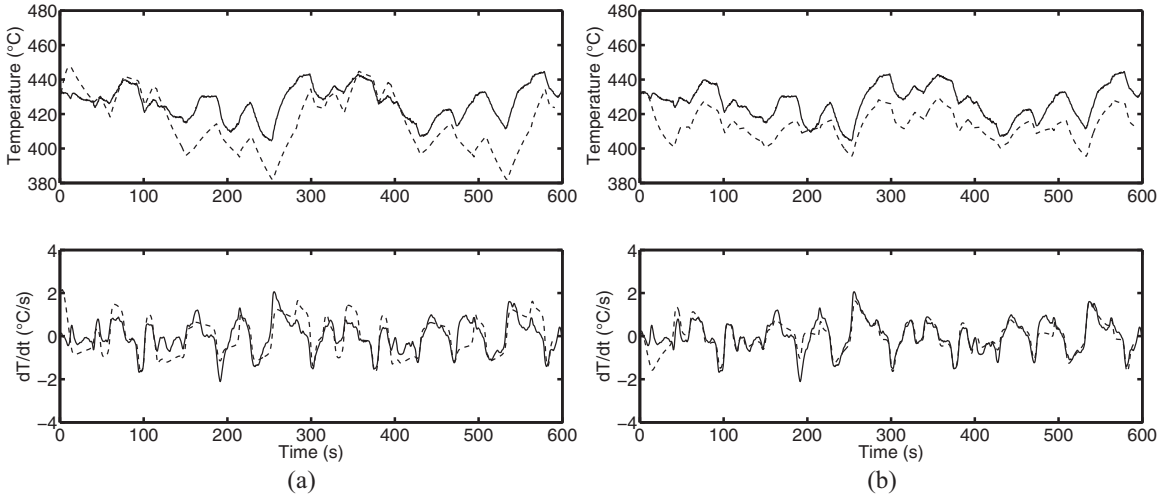

Fig. 5. Simulated temperature and temperature derivative based upon the FOPDT model (dashed line) against PRBS 002 weld data (solid line). (a) Simulation with optimizer determined parameters. (b) Simulation with manually determined parameters. The MSSSE for the optimizer and manually determined models' predictions is 91,400 °C^2 s and 13,600 °C^2 s, respectively. The average absolute error of the derivative is 0.93 °C/s and 0.59 °C/s, respectively. Temperatures were filtered before taking derivatives because the derivatives of the original signals had a high noise to signal ratio and were practically unusable.

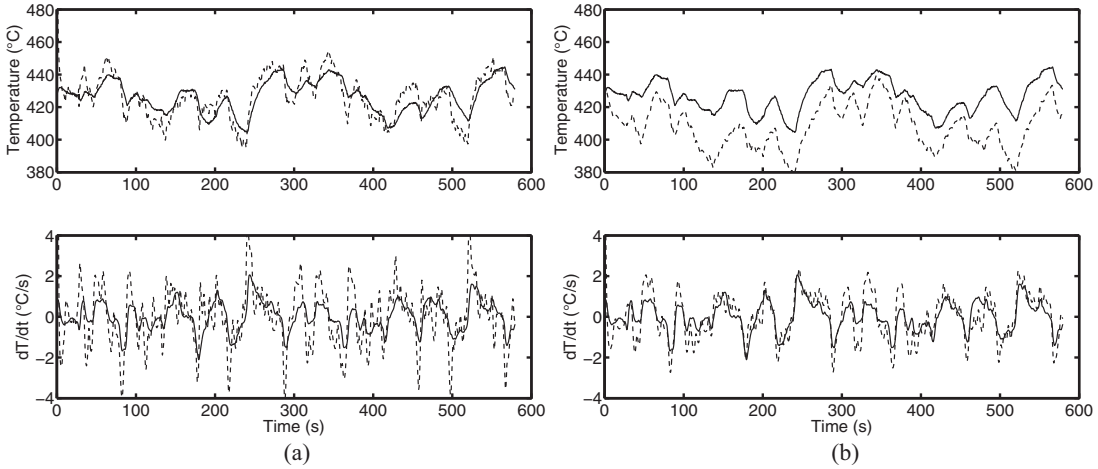

Fig. 6. Simulated temperature and temperature derivative based upon the Hybrid Heat Source model (dashed line) against PRBS 002 weld data (solid line). (a) Simulation with optimizer determined parameters. (b) Simulation with manually determined parameters. The MSSSE for the optimizer and manually determined models' predictions is 29,800 °C^2 s and 25,100 °C^2 s, respectively. The average absolute error of the derivative is 1.78 °C/s and 1.10 °C/s, respectively. Temperatures were filtered before taking derivatives because the derivatives of the original signals had a high noise to signal ratio and were practically unusable.

the manually determined parameters than for the optimizer determined parameter. Furthermore, the manual parameter set predicts a temperature with many less spurious fluctuations for the Hybrid Heat Source model. For both models, a MPC controller is expected to perform better using the manual parameters than with the optimizer parameters.

Comparing the two models' predictions to each other, the FOPDT predictions exhibit less spurious temperature variations, and matches more features of the actual temperature profile. The MSSSE and average absolute error of the derivative for the FOPDT model with manual parameters are the lowest of all models tested. After the startup transient, a MPC controller based upon the FOPDT model with manually determined parameters is expected to have the best control temperature.

Neither model predicts the startup transient region exceptionally well. The FOPDT model predicts large temperature decreases and is incapable of predicting the relatively steady or increasing temperatures which are observed in the process data. The FOPDT model is of the wrong form to predict these trends and is consequently ill-suited for control in this region of the weld. The Hybrid Heat Source model natively accounts for heat transfer within the plate and predicts thermal trends in this region much better than the FOPDT model. For best control, the model may be trained specifically for the starting transient conditions. Temperature control in this region using the Hybrid Heat Source MPC controller is likely inferior to the FOPDT MPC controller after the startup transient, but is expected to be better than a FOPDT based MPC controller or an absence of temperature control during the startup transient.

5. Conclusions

Both FOPDT and Hybrid Heat Source models capture most of the temperature trends resulting from changes in traverse speed and power. Therefore, both may be viable models for MPC temperature control. The FOPDT model is generally the better of the two for mid-weld control. The Hybrid Heat Source model on the other hand is better formulated to predict the dynamics observed at the beginning of the weld.

Using a manual curve fitting approach to fit the system resulted in a lower MSSSE and a lower average absolute error of the

derivative for both models. This quantitative analysis of the output of the two methods confirms that minimizing the SSE does not necessarily result in a model which best matches the dynamic responses required for good MPC control. Thus, minimizing the SSE appears not to be the best method for time-series modeling, and should be coupled with a derivative-based metric if curve fitting is to be performed via optimization.

Acknowledgements

This work was funded by the Center for Friction Stir Processing, an NSF Industry/University Cooperative Research Center.

References

[1] Lakshminarayanan A, Balasubramanian V, Elangovan K. Effect of welding processes on tensile properties of AA6061 aluminium alloy joints. Int J Adv Manuf Technol 2009;40(3–4):286–96.
[2] Mahoney M, Rhodes C, Flintoff J, Bingel W, Spurling R. Properties of friction-stir-welded 7075 T651 aluminum. Metall Mater Trans A 1998;29(7):1955–64.
[3] Ross KA. Investigation and implementation of a robust temperature control algorithm for friction stir welding. Brigham Young University; 2012.
[4] Chimbli SK, Medlin DJ, Arbegast WJ. Minimizing lack of consolidation defects in friction stir welds. In: Proc. 136th TMS annual meeting friction stir welding & processing IV symposia TMS. 2007.
[5] Cederqvist L, Garpinger O, Hagglund T, Robertsson A. Reliable sealing of copper canisters through cascaded control of power input and tool temperature. In: Mishra R, Mahoney MW, Sato Y, Hovanski Y, Verma R, editors. Friction stir welding and processing VI. Hoboken, NJ: John Wiley & Sons, Inc; 2011.
[6] Hattingh DG, van Niekerk TI, Pothier R. Closed-loop temperature control of material plasticisation during friction stir welding. Advanced Materials Research Trans Tech Publ; 2014.
[7] Fehrenbacher A, Smith CB, Duffie NA, Ferrier NJ, Pfefferkorn FE, Zinn MR. Combined temperature and force control for robotic friction stir welding. J Manuf Sci Eng 2014;136(2):021007.
[8] Ross K, Sorensen C. Advances in temperature control for FSP. John Wiley & Sons, Inc; 2013.
[9] Marshall DJ. An alternative system identification method for friction stir processing. Brigham Young University; 2013.
[10] Marshall D, Sorensen C. System parameter identification for friction stir processing. In: Friction stir welding and processing VII. John Wiley & Sons; 2013.
[11] O'Dwyer A. Handbook of PI and PID controller tuning rules. World Scientific; 2009.
[12] Nielsen I. Modeling and control of friction stir welding in 5 cm thick copper canisters. Linköpings Universitet; 2012.
[13] Cederqvist L, Garpinger O, Cervin A, Nielsen I. Improved temperature and depth control during FSW of copper canisters using feedforward compensation. In: Friction stir welding and processing VIII; 2015. p. 69–76.
[14] Cortinovis A, Ferreau H, Lewandowski D, Mercangöz M. Experimental evaluation of MPC-based anti-surge and process control for electric driven centrifugal gas compressors. J Process Control 2015;34:13–25.
[15] Garcia CE, Prett DM, Morari M. Model predictive control: theory and practice—a survey. Automatica 1989;25(3):335–48.
[16] Ma J, Qin SJ, Salsbury T. Application of economic MPC to the energy and demand minimization of a commercial building. J Process Control 2014;24(8):1282–91.
[17] Anzehaee MM, Haeri M. Welding current and arc voltage control in a GMAW process using ARMarkov based MPC. Control Eng Pract 2011;19(12):1408–22.
[18] Zhang YM, Kovacevic R. Neurofuzzy model-based predictive control of weld fusion zone geometry. IEEE Trans Fuzzy Syst 1998;6(3):389–401.
[19] Liu YK, Zhang YM. Control of 3D weld pool surface. Control Eng Pract 2013;21(11):1469–80.
[20] Liu YK, Zhang YM. Model-based predictive control of weld penetration in gas tungsten arc welding. IEEE Trans Control Syst Technol 2014;22(3):955–66.
[21] Posada M. Comparison of 3-D friction stir welding viscoplastic finite element model with weld data and physically-simulated data. Brigham Young University; 2012.
[22] Upadhyay P, Reynolds AP. Thermal management in friction-stir welding of precipitation-hardened aluminum alloys. JOM 2015;67(5):1022–31.
[23] Shrivastava A, Pfefferkorn FE, Duffie NA, Ferrier NJ, Smith CB, Malukhin K, et al. Physics-based process model approach for detecting discontinuity during friction stir welding. Int J Adv Manuf Technol 2015;79(1–4):605–14.
[24] Kuykendall KL. An evaluation of constitutive laws and their ability to predict flow stress over large variations in temperature, strain, and strain rate characteristic of friction stir welding. Brigham Young University; 2011.
[25] Taysom BS. Temperature control in friction stir welding using model predictive control. Brigham Young University; 2015.
[26] Carslaw HS, Jaeger JC. Conduction of heat in solids. 2nd ed. New York: Oxford University Press; 1959.
[27] Hou Z, Komanduri R. General solutions for stationary/moving plane heat source problems in manufacturing and tribology. Int J Heat Mass Transfer 2000;43(10):1679–98.
[28] Verstegg HK, Malalasekera W. An introduction to computational fluid dynamics: the finite volume method. 2nd ed. Pearson Prentice Hall; 2007. ISBN 978-0-13-127498-3.
[29] Incropera FP, Dewitt DP, Bergman TL, Lavine AS. Fundamentals of heat and mass transfer. 6th ed. John Wiley and Sons; 2006.
[30] Taylor R, Groot H, Goerz T, Ferrier J, Taylor D. Thermophysical properties of molten aluminium alloys. HTHP Abstr 1998;30(3):269–75.

Brandon (Scott) Taysom is a Ph.D. student at Brigham Young University and is currently researching bi-metal rotary friction welding of next-generation superalloys. His other research areas include friction stir welding control theory and application, and materials modeling in moderate-to-high temperature and strain conditions.

Carl Sorensen is a Professor of Mechanical Engineering at Brigham Young University. He received his Ph.D. in Materials Science from MIT and has worked as a consultant in manufacturing processes for General Electric and Chrysler, as well as numerous smaller companies. He actively researches friction stir welding, friction welding, and control of joining processes.

John Hedengren is an Assistant Professor in the Department of Chemical Engineering at Brigham Young University. He received his Ph.D. in Chemical Engineering from the University of Texas at Austin in process dynamics and control. He developed the APMonitor Optimization Suite and worked with ExxonMobil on advanced process control for oil and gas processing. His other research interests include drilling automation, fiber optic monitoring for Intelli-fields, reservoir optimization, unmanned aerial vehicles, and model predictive control.