Probabilistic kernel machines for predictive monitoring of weld residual stress in energy systems

Miltiadis Alamaniotis a, *, Jino Mathew b, Alexander Chroneos b, Michael E. Fitzpatrick b, Lefteri H. Tsoukalas a

a School of Nuclear Engineering, Purdue University, West Lafayette, IN, USA
b Faculty of Engineering, Environment and Computing, Coventry University, Priory Street, Coventry CV1 5FB, UK

A R T I C L E   I N F O

Keywords:
Probabilistic kernel machines
Gaussian process regression
Machine learning
Welding
Residual stresses
Support vector regression

A B S T R A C T

Predictive monitoring supports the a priori scheduling of critical component maintenance and contributes significantly in attaining a safe yet economic operation and management of complex energy systems by mitigating the risk of accidents and minimizing the number of operational pauses. The current work studies the learning ability of probabilistic kernel machines, and more particularly of Gaussian Processes (GP) equipped with various kernels for the estimation of weld residual stress profiles of stainless steel pipe welds. The GP models are tested on experimentally-obtained data of axial and hoop residual stresses in two different stainless-steel pipes. The results exhibit the ability of GP to accurately predict the weld residual stress profile in the axial and hoop direction by providing a predictive distribution, i.e., mean and variance values. Furthermore, performance of GP is compared to a non-probabilistic kernel machine, such as support vector regression (SVR) equipped with the same kernels, and to multivariate linear regression (MLR). Comparison results exhibit the robustness of GP over SVR and MLR with respect to prediction accuracy of weld residual stress in terms of root mean square error. With respect to a second metric, namely, correlation coefficient between measured and predicted values, GP is superior to SVR and MLR in the majority of the cases.

1. Introduction

Satisfaction of the growing demand for electrical energy necessitates the continuous operation of power generation plants that provide “base-load” electricity generation (McCoy et al., 2013). Part of the overall power plant operational management is the predictive maintenance that encompasses the a priori scheduling of maintenance and replacement of close-to-failure mechanical components (Atoui et al., 2015), aiming to minimize the number of operational pauses and enhance the overall system safety (Lei et al., 2009; Ebersbach and Peng, 2008). To that end, estimation – within narrow uncertainty bounds – of the remaining life of vital components in energy systems (Liu et al., 2008) will contribute to the adoption of predictive monitoring techniques (Woottton et al., 2017). Such techniques may realize cost effective maintenance strategies that lead to enhanced system safety and performance (Hashemian and Bean, 2011; Alamaniotis et al., 2014; Mathias et al., 2012).

Essential part in the overall operation of every energy installation are pipes: in thermal power plants, these pipes carry steam at high temperatures and pressures. It has been identified that residual stresses introduced as a consequence of welding processes is a fundamental factor that can lead to the initiation of cracks whose growth may lead to component failure (Castellanos et al., 2011; Withers, 2007; Babu et al., 2009). The thermomechanical effects of the welding process result in plastic misfits that lead to the generation of elastic residual stresses (Bouchard, 2007; Withers and Bhadeshia, 2001; Francis et al., 2007). However, this can be further exacerbated by repair processes (Edwards et al., 2005).

Accurate prediction of residual stresses requires the development of analytical models encompassing a high number of interacting factors, detailed knowledge of the welding parameters and materials properties, such as cyclic hardening behavior and the thermal-dependence. It is the inherent complexity and, often, lack of necessary information in sufficient detail that makes quantification of the residual stress distribution a challenging problem (Stone et al., 2008). Experimental techniques such as neutron diffraction (Hutchings et al., 2005), deep hole drilling (George et al., 2002), and the contour method (Kartal et al., 2008) have

https://doi.org/10.1016/j.engappai.2018.02.009
Received 8 September 2017; Received in revised form 19 December 2017; Accepted 13 February 2018
0952-1976/ © 2018 Elsevier Ltd. All rights reserved.
been developed to measure residual stress distribution, but come at a high financial cost, cannot be deployed in situ, and are at least semi-
destructive of the component (Mahmoudi et al., 2009; Prime, 2001). In
addition, computational and data-driven methods have been applied in
predicting residual stress in energy systems. For example, finite element
models (Schwane et al., 2014; Xie et al., 2017) and artificial neural
networks (Mathew et al., 2014) are used despite their limitations in
modeling of highly complex non-linear and interacting processes that
introduce a high degree of uncertainty (Tsoukalas and Uhrig, 1997).

Machine learning (Liu et al., 2015) has been recently identified as
a domain that potentially offers solutions to a large range of predictive
problems in materials (Yildiz, 2013; Balachandran et al., 2015). Machine
learning tools are exposed to a set of known datasets, which consist of
experimental or simulated data, in order to evaluate their parameters: a
process known as learning or training (Alamaniotis et al., 2012). The
trained models are able to provide predictions for conditions that
are exposed to similar conditions (Bishop, 2006). Predictions may be
performed either in the form of interpolation or extrapolation depending
on the specifics of the application under study (Alamaniotis et al.,
2010; Babu et al., 2010). To that end, machine learning tools such as
artificial neural networks have been applied to predict weld residual
stress profiles where the predictions from ensemble networks were used
to develop a prediction interval (Mathew et al., 2013, 2017a). Further,
Dhas and Kumanan (2016) proposed an evolutionary fuzzy support
vector regression method that predicts weld residual stresses, while
a non-dominated moth flame optimization technique is proposed by
Savas and Tawhid (2017). In addition, a method that integrates neural
networks and evolutionary computing in weld stress prediction is intro-
duced by Dhas and Kumanan (2014), and a hybrid method that utilizes
experimental data and neural networks by Mathew et al. (2017b).
The adoption of advanced statistical learning methods for weld prediction
was presented in (Lewis et al., 2017 July), and the use of an iterative
substructure method (ISM) for weld stress prediction in pressurized
water reactor (PWR) in Maekawa et al. (2016). Several weld residual
stress prediction techniques are based on the finite elements method
in conjunction with various simulated or experimental conditions as
introduced in Wang et al. (2015), Mondal et al. (2017), Jiang et al.
(2015) and Afshari et al. (2016).

In this paper, probabilistic kernel machines (Fricke et al., 2001) are
utilized for predicting weld stress profiles of power plant components
(Pilania et al., 2013). In particular, the machine learning aspect of Gaus-
sian Processes (GP) (Rasmussen and Williams, 2006; Quinonero-Candela
and Rasmussen, 2005) modeled with a variety of kernel functions is
studied for predicting residual stress, and its performance is compared to
non-probabilistic kernel machines. Overall, the set of research objectives
contains:

(i) Study of the learning ability of various forms of GP to predict the
weld stress profiles,
(ii) Comparison of GP prediction performance to non-probabilistic
driven tools,
(iii) Application of GP on recent experimentally obtained datasets of
stainless steel pipe welds.

In addition, the practical implications of the current study involve
the implementation of predictive monitoring of weld stress in energy
systems, and the development of automated maintenance techniques.
Additionally, GP predictions of weld stress allow a timely yet low cost
replacement of critical system components.

In the next section, Gaussian processes are presented in the context of
probabilistic kernel machines, while Section 3 proposes the application
of Gaussian process for weld residual stress prediction and lastly,
Section 4 concludes the paper and summarizes its findings.

2. Probabilistic kernel machines

2.1. Kernel-based Gaussian process regression

With the exception of linear regression methods where the output
is a single value that is computed by a set of weighted inputs, there
exists a class of methods that make predictions in the function space
Gregorčič and Lightbody, 2009); such methods are the probabilistic
kernel machines. In the realm of machine learning, probabilistic kernel
machines are recognized as the Bayesian extension of simple kernel
machines (Bishop, 2006; Rasmussen and Williams, 2006). Here, we
implement the notion of a kernel-based Gaussian process with a joint
distribution that is modeled as a function of a kernel.

A kernel, which is denoted as \( k(x_1, x_2) \), is a valid mathematical
function that can be written as (Bishop, 2006):

\[
k(x_1, x_2) = f(x_1)^T f(x_2)
\]

where \( f(x) \) is any valid analytical function. Expressing an analytical
model as a function of a kernel is called the “kernel trick” (Bishop,
2006). A Gaussian process expressed as a function of Eq. (1) and utilized
for prediction making in regression problems is simply called Gaussian
process regression (GPR).

Derivation of GPR has an initial point analogous to the simple linear
regression, whose vector form is given below:

\[
y(x) = w^T \varphi(x)
\]

where \( w \) is the regression coefficient vector and \( \varphi(x) \) the vector contain-
ing the basis functions. The basis function may be nonlinear in relation
to \( x \), but still \( y \) remains a linear combination of basis functions. Next,
a prior distribution over regression coefficients \( w \) is set, as shown in
Eq. (3):

\[
p(w) = N(0, \sigma^2 I)
\]

where the mean value is zero (a convenient choice since we have little
or no prior information on weights), \( I \) is the identity matrix, and \( \sigma^2 \)
denotes the variance associated with the regression coefficients.

A Gaussian process is defined by two parameters, namely, its mean
\( m(x) \) and covariance \( C(x', x) \) function. Therefore, a Gaussian process is
given by:

\[
GP \sim N \left( m(x), C(x', x) \right)
\]

where \( m(x) \) is taken as equal to zero and \( C(x', x) \) is replaced by a kernel
function, i.e., kernels are inserted into the GP as covariance functions.
As a result, the prior distribution over the output \( y \) is:

\[
P(y) = N(0, K)
\]

with \( K \) being the so-called Gram matrix, whose elements are taken as:

\[
K_{ij} = k(x_i, x_j) = \sigma_k^2 \sum_p \varphi_p(x_i) \varphi_p(x_j).
\]

In general, a measured value contains the real value \( y \) added by
noise:

\[
t_n = y(x) + \epsilon_n,
\]

where \( t_n \) is the \( n \)th datapoint of \( y \) and \( \epsilon_n \) is uncorrelated Gaussian noise
of zero mean and variance \( \sigma^2 \). Hence, we obtain a Gaussian distribution
over the observed targets:

\[
P(t) = N(0, C) = N(0, K + \sigma^2 I),
\]

where the elements of the covariance matrix are:

\[
C_{mn} = k(x_m, x_n) + \sigma^2 \delta_{mn}
\]

with \( \delta_{mn} \) representing the Dirac delta function. Considering the population
of available datapoints being equal to \( N \), then Eq. (8) can be used to
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات