

Inverse Analysis of Heat Transfer in Blast Furnaces

Xiao Tian (23R51042)

Tokyo Institute of Technology
Japan 152-8550
tian.x.af@m.titech.ac.jp

Abstract

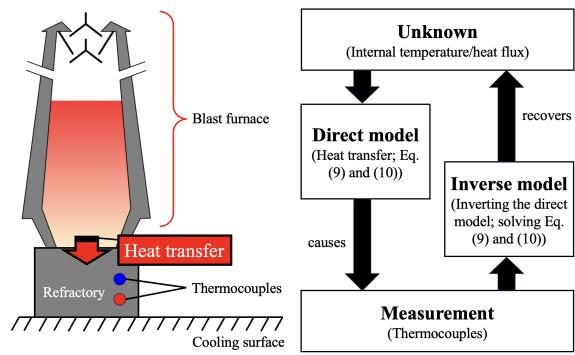
When direct measurement of quantity of interest is not feasible, one can resort to inverse analysis - measuring some accessible resultant quantities and recovering the unknown quantity by inverting the causal process. This applies to blast furnaces: One needs to gauge and control the internal temperature to ensure stable chemical reaction, but cannot measure it directly due to the intense temperature. In this report, we revisit and examine the inverse analysis approach to this problem: One can instead measure the relatively low temperature within a refractory beneath the blast furnace, and retrieve the internal temperature by inverting the heat transfer process. We reinstate the theoretical underpinning of this technique and verify its practicality using simulated data.

1 Introduction

Molten iron, an intermediate product in iron-making that can be subsequently used to generate iron and steel, is typically produced in *blast furnaces* using iron ores, coke and oxygen-enriched air [20]. The chemical reaction among the aforementioned ingredients takes place at around 1500°C and it is thus important to know and control the temperature inside blast furnaces to maintain a stable reaction. However, direct measurement is impracticable since the internal temperature of blast furnaces can reach around 2000°C [23]. Therefore, a surrogate technique to infer the internal phenomenon of blast furnaces from some indirect measurement is essential.

The phenomena of *heat transfer* has been well studied [17] and heat is typically modelled as a partial differential equation (PDE) of time and position. Therefore, we can alternatively measure the temperature at a separate position where heat can transfer from the internal of blast furnaces to that position, then recover the internal temperature by “inverting” the heat transfer process. As an example, we can place a pair of thermocouples within a refractory below the blast furnace (Fig. 1a).

The surrogate technique as described above is actually an *inverse problem* which contains three elements: **unknown**, **measurement** and **model** (Fig. 1b, we will discuss how to formulate the three elements in Sec. 2). To solve an inverse problem, we fit our measurements to a strategically chosen inverse model to find out the unknown quantities. In this project, we explore the possibility of the surrogate technique as described above - **to recover the internal phenomena**



- (a) Demonstration of the surrogate measurement. A pair of thermocouples are placed within a refractory below the blast furnace where the temperature is less high yet still dependent on the internal temperature and an inverse model to which through heat transfer.
- (b) Illustration of an inverse problem. Measurements of blast furnace temperature are made at a cooling surface. These measurements are used to fit the inverse model, which then recovers the unknown quantity of interest.

Figure 1: Physical and logical overview of the problem.

through an inverse problem approach. Instead of conducting real-life experiments using actual blast furnaces, we propose (in Sec. 3) to simulate the process by choosing an arbitrary input heat flux and calculating the temperatures at the positions of the thermocouples as a result of heat transfer. Such simulation would make our project more customizable, efficient and resource-saving. We then recover the input heat flux using the inverse problem approach and compare it with the actual heat flux (in Sec. 4).

2 Problem Formulation

In this section, we discuss how to formulate our goal as an inverse problem. To simplify our problem, we model the heat transfer within the refractory as a 2-dimensional (i.e., time and 1-dimensional displacement) boundary value problem, where the temperature u at any point is related to its displacement from the heat source (i.e., blast furnace) x (with domain $[0, L = 4 \text{ m}]$) and the time t (with domain $[0, T = 30 \text{ d}]$). The positions of thermocouples are shown

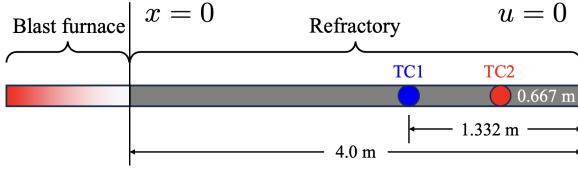


Figure 2: Heat transfer within the refractory is simplified to 1-dimensional in space¹. The boundary between the blast furnace and the refractory is set as $x = 0$, and the boundary between the refractory and the ground has to satisfy temperature $u = 0$ for the heat to successfully cool down.

in Fig. 2 as TC1 and TC2. The relationship between temperature u , time t and displacement x is governed by the well-known *Heat equation* [6]:

$$\frac{\partial u}{\partial t} = \frac{k}{\rho \cdot C_p} \frac{\partial^2 u}{\partial x^2}, \quad (1)$$

where $k = 21.20 \text{ W}/(\text{m K})$ denotes the *thermal conductivity* of the refractory, $\rho = 2300 \text{ kg}/\text{m}^3$ denotes the *density* of the refractory and $C_p = 712 \text{ J}/(\text{kg K})$ denotes the *specific heat capacity* of the refractory. For simplicity, let $\alpha = \frac{k}{\rho \cdot C_p} = 1.29 \times 10^{-5} \text{ W m}^2$ denote the constant coefficient.

There are also three obvious conditions:

- (Initial condition) When there is no input heat, the temperature is equal to 0 at any distance, i.e.,

$$\forall x \in [0, L] [u(x, 0) = 0]; \quad (2)$$

- (Dirichlet condition) The temperature is always equal to 0 at the cooling surface, i.e.,

$$\forall t \in [0, T] [u(4, t) = 0]. \quad (3)$$

- (Neumann condition) The input heat flux at the source, h , changes over the time. By Fourier's law [3], we have

$$\forall t \in [0, T] \left[\frac{\partial u}{\partial x}(0, t) = -\frac{1}{k} \cdot h(t) \right]. \quad (4)$$

With the above modelling of heat transfer, we proceed to define the **unknown**, **measurement** and **model** of our inverse problem.

Unknown. The unknown quantity in our inverse problem is clearly the input heat flux at the source, $h(\cdot)$.

Measurement. The measurements in our inverse problem are the temperatures measured by thermocouples TC1 and TC2 every $\Delta t = 3$ hours over $T = 30$ days. For simplicity, we denote each timestamp when measurement takes place as $t_0 = 0, t_1, t_2, \dots, t_M = 30$. We further denote the temperatures measured by TC1 as a vector **TC1** with elements TC1_m ($m = 1, 2, \dots, M = \frac{24 \text{ h}}{3 \text{ h}} \cdot 30 \text{ d} = 240$)² and the temperatures measured by TC2 as a vector **TC2** with elements TC2_m . Specifically, the measurements **TC1** and **TC2** will be simulated according to the Heat equation (1). More details are given in Sec. 3.

¹Fig. 2 is a 90° anti-clockwise rotated view of a blast furnace.

²The first measurement at $t = 0$ is trivial and thus omitted.

Model. Let x_{TC1} and x_{TC2} denote the locations of thermocouples TC1 and TC2. Our measurements represent the temperature at these two locations, i.e., $u(x_{\text{TC1}}, \cdot)$ and $u(x_{\text{TC2}}, \cdot)$ (if not discretized). There exists a mapping L from each possible input heat flux $h(\cdot)$ to $[u(x_{\text{TC1}}, \cdot); u(x_{\text{TC2}}, \cdot)]$ by solving the Heat equation (1). Therefore, the (conceptual) direct model is

$$[u(x_{\text{TC1}}, t); u(x_{\text{TC2}}, t)] = L(h(t)), \quad (5)$$

and the (conceptual) inverse model is L^{-1} . More concrete forms (i.e., integral and matrix forms) of the direct and inverse models are derived in the following subsections.

2.1 Green's Function

The Green's function [22] is a useful tool in PDEs to analyze a localized response to a point source. Define $G(x, \xi; t, \tau)$ to be the Green's function that satisfies

$$\frac{\partial G}{\partial t} - \alpha \cdot \frac{\partial^2 G}{\partial x^2} = \delta(x - \xi)\delta(t - \tau), \quad (6)$$

where $\delta(\cdot)$ represents the Dirac delta function [13]. Then $G(x, \xi; t, \tau)$ represents the response at position x and time t caused by the heat source at position ξ and time τ . For example, $G(x_{\text{TC1}}, 0; t, 0)$ represents how TC1 at time t is affected by the initial heat flux $h(0)$ from the source.

By Superposition principle [11], the actual temperature $u(x, t)$ is the sum of individual responses caused by the heat source at every position (in our setting, only $x = 0$) and time. It is thus an integral of each response. For example, the temperature at TC1 at time t is

$$u(x_{\text{TC1}}, t) = \int_0^t G(x_{\text{TC1}}, 0; t, \tau)h(\tau) \, d\tau. \quad (7)$$

It is difficult to invert the integral equation (7) for us to recover $h(\cdot)$ from $u(x_{\text{TC1}}, \cdot)$. Therefore, we discretize it and convert it into a matrix equation in the following subsection.

2.2 Discretization

By considering the discretized timestamps t_0, t_1, \dots, t_M defined earlier, the integral equation (7) is transformed into the following equation for any $m = 1, \dots, M$:

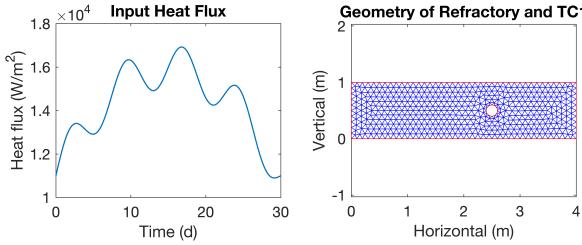
$$u(x_{\text{TC1}}, t_m) = \sum_{i=1}^m \left[\int_{t_{i-1}}^{t_i} G(x_{\text{TC1}}, 0; t_m, \tau) \, d\tau \right] \cdot h_i, \quad (8)$$

where $h_i = h\left(\frac{t_{i-1}+t_i}{2}\right)$ denotes the value of h at the midpoint of each time interval. Let **h** be the vector (h_1, h_2, \dots, h_M) . Let **G1** be an $M \times M$ matrix whose (m, i) -th element $\text{G1}_{m,i} = \int_{t_{i-1}}^{t_i} G(x_{\text{TC1}}, 0; t_m, \tau) \, d\tau$ when $m \geq i$ and 0 otherwise. We can now replace Eq. (8) with the matrix equation

$$\text{TC1} = \text{G1}\text{h}. \quad (9)$$

In the same manner, we can define a matrix equation for TC2 as follows (the notations are similar to those of TC1):

$$\text{TC2} = \text{G2}\text{h}. \quad (10)$$



(a) Hyperbolic input heat flux function $h(\cdot)$ with periodic fluctuations on a weekly basis.
(b) Mesh according to which the whole refractory is split into finite elements.

Figure 3: Customization of boundary value problem.

Eq. (9) and (10) serve as the **direct model** in our problem, whereas solving Eq. (9) and (10) serves as the **inverse model** in our problem. For example, if $\mathbf{G}1$ is invertible, we can simply recover \mathbf{h} from $\mathbf{TC}1$ by

$$\mathbf{h} = \mathbf{G}1^{-1} \mathbf{TC}1. \quad (11)$$

We will discuss further on how to obtain $\mathbf{G}1$ and $\mathbf{G}2$ and solve the linear system (9) and (10) in Sec. 4

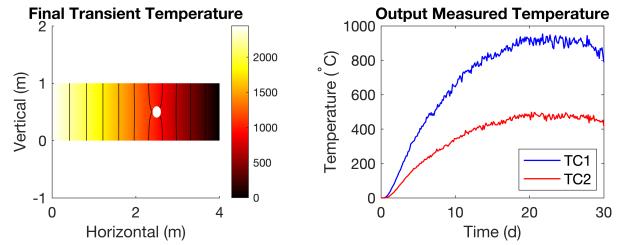
3 Simulation of Heat Transfer

In this section, we will explain how we simulate the heat transfer process within blast furnaces and generate the phantom data that corresponds to the measurements at thermocouples $\mathbf{TC}1$ and $\mathbf{TC}2$ (i.e., vectors $\mathbf{TC}1$ and $\mathbf{TC}2$).

We first arbitrarily choose the form of input heat flux $h(\cdot)$. Fig. 3a gives an example of a customized $h(\cdot)$. With this the boundary condition (4) is set and the boundary value problem is complete and solvable.

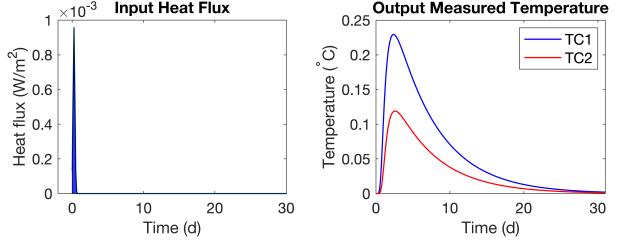
We then proceed to solving the boundary value problem as defined by Eq. (1), (2), (3) and (4). While an analytical solution is hard to find, we employ numerical methods, specifically finite element method (FEM) [21], to obtain an approximate solution. FEM discretizes the refractory into smaller *elements* by constructing a mesh of it (Fig. 3b), approximates the temperature function u over each element and recombines the element equations to form the whole approximated function. In this project, we utilize the Partial Differential Equation Toolbox in MATLAB as an implementation of FEM. Fig. 4a shows the final temperature after $T = 30$ days solved with the toolbox.

To reflect the real-life situation where there might be measurement errors, imperfect heat transfer or external interferences that affect the temperature measurements, we inject artificial random noises to the simulated temperature measurements at $\mathbf{TC}1$ and $\mathbf{TC}2$. Existing literatures diverge on the choice of suitable noise level (e.g., [25] chooses less than 3%, while [16] prefers more than 5%). Therefore, we moderately adopt Gaussian noises with mean 0 and standard deviation 1.94% of each measured temperature such that over 87% of the temperatures have noise level less than 3% and 99% of the temperatures have noise level less than 5%. The finalized temperature measurements with such noises are shown in Fig. 4b



(a) Final transient temperature distribution within the refractory after $T = 30$ days..
(b) Simulated temperature measurements at $\mathbf{TC}1$ and $\mathbf{TC}2$ with artificial random noises.

Figure 4: Simulated results.



(a) Gaussian approximation to an impulse at $t = 0$. We shade the region under the curve for better visualization.
(b) Simulated temperature measurements at $\mathbf{TC}1$ and $\mathbf{TC}2$ (i.e., local responses) without artificial noises.

Figure 5: Simulated impulse and response used to approximate Green's function.

4 Solving the Inverse Problem

4.1 Approximation of $\mathbf{G}1$ and $\mathbf{G}2$

Before solving the linear systems (9) and (10), we have to quantify matrices $\mathbf{G}1$ and $\mathbf{G}2$ as defined by Eq. (6) and (8). Recall that $G(x_{\mathbf{TC}1}, 0; t, 0)$ represents how $\mathbf{TC}1$ at time t is affected by the initial heat flux $h(0)$ from the source. Therefore, it is possible to approximate $G(x_{\mathbf{TC}1}, 0; t, 0)$ at different time t by simply replacing the input heat flux $h(\cdot)$ as a Dirac delta function of position and time, $\delta(x - 0)\delta(t - 0)$ (Fig. 5a). We then repeat the simulation procedures in Sec. 3 without adding the noises³. Fig. 5b shows the simulated temperature measurements, which are also graphs of $G(x_{\mathbf{TC}1}, 0; t, 0)$ and $G(x_{\mathbf{TC}2}, 0; t, 0)$.

The approximation of $\mathbf{G}1$ and $\mathbf{G}2$ is possible with the discrete values of G . We follow the Trapezium rule [5], e.g.,

$$G1_{M,i} = \int_{t_{i-1}}^{t_i} G(x_{\mathbf{TC}1}, 0; t_M, \tau) d\tau \approx \widehat{\Delta t} \cdot \frac{G(x_{\mathbf{TC}1}, 0; t_M, t_{i-1}) + G(x_{\mathbf{TC}1}, 0; t_M, t_i)}{2}, \quad (12)$$

³Refrain from setting the mean of Gaussian function at exactly $t = 0$ where the integral is reduced by half and the approximation would be inaccurate.

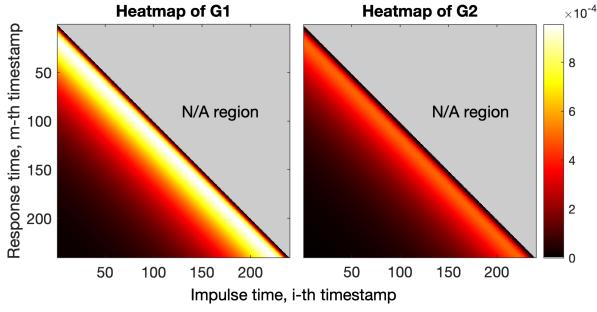


Figure 6: Heatmaps of **G1** and **G2**. Note that the “N/A region” corresponds to the 0 entries in the matrices, because an impulse happening at t_i later than t_m will not trigger a response at t_m .

where $\widehat{\Delta t}$ is the normalized time interval with regards to total time.

It is also straightforward that for any $\tau \in [0, t]$, $G(x, 0; t, \tau)$ is simply a translation of $G(x, 0; t, 0)$, i.e.,

$$G(x_{TC1}, 0; t, \tau) = G(x_{TC1}, 0; t - \tau, 0); \quad (13)$$

$$G(x_{TC2}, 0; t, \tau) = G(x_{TC2}, 0; t - \tau, 0). \quad (14)$$

Therefore, we can set other $G1_{m,i}$ ($1 \leq m < M$) by iteratively setting

$$G1_{m,i} = \begin{cases} 0 & , \text{if } i = m; \\ G1_{m+1,i+1} & , \text{otherwise.} \end{cases} \quad (15)$$

Fig. 6 shows the heatmaps of **G1** and **G2**. A brighter region indicates a stronger effect from the impulse at the heat source at time t_i to the response at TC1 and TC2 at time t_m . There are two intuitive trends:

- When $i < m$, the brightest region lies close to the $i = m$ line. This corresponds to the peak in Green’s function G (Fig. 5b).
- The heatmap of **G1** is in general brighter than that of **G2** because TC1 is closer to the heat source than TC2.

4.2 Solving the Linear System

We first assemble the two matrix equations (9) and (10) into one:

$$\begin{bmatrix} \mathbf{TC1} \\ \mathbf{TC2} \end{bmatrix} = \begin{bmatrix} \mathbf{G1} \\ \mathbf{G2} \end{bmatrix} \mathbf{h}. \quad (16)$$

For simplicity, let $\mathbf{TC} = \begin{bmatrix} \mathbf{TC1} \\ \mathbf{TC2} \end{bmatrix}$ and $\mathbf{G} = \begin{bmatrix} \mathbf{G1} \\ \mathbf{G2} \end{bmatrix}$. \mathbf{TC} is a $2M \times 1$ vector while \mathbf{G} is a $2M \times M$ matrix.

Vanilla Least Squares Method (LSM). The value of \mathbf{h} can be approximated through LSM:

$$\hat{\mathbf{h}} = (\mathbf{G}^\top \mathbf{G})^{-1} \mathbf{G}^\top \mathbf{TC}, \quad (17)$$

where $\hat{\mathbf{h}}$ denotes the least squares approximation of \mathbf{h} . However, the matrix $\mathbf{G}^\top \mathbf{G}$ has a very large condition number⁴

⁴A tool to diagnose matrix singularity.

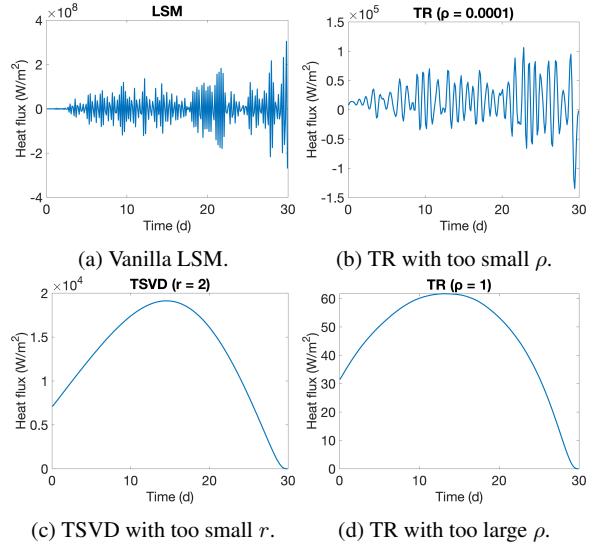


Figure 7: Effect of numerical errors. (7a) and (7b) demonstrate overfitting while (7c) and (7d) demonstrate underfitting. (7b), (7c) and (7d) implies poorly chosen regularization parameters.

4.506 $\times 10^{21}$ and is thus nearly singular. The inverse has significant numerical errors and in our case, makes the result \mathbf{h} almost a zero vector (Fig. 7a). Therefore, we need to adopt suitable regularization techniques, which are discussed further in the next section.

4.3 Regularization

In this section, we discuss and compare a few common regularization techniques: **truncated singular value decomposition (TSVD)** [12], **Tikonov regularization (TR)** [8] and **maximum likelihood estimation (MLE)** [18].

TSVD. Singular value decomposition decomposes the matrix \mathbf{G} into the product of three matrices:

$$\mathbf{G} = \mathbf{U} \mathbf{S} \mathbf{V}, \quad (18)$$

where \mathbf{S} is a diagonal matrix whose diagonal element represents the importance of each singular vector in \mathbf{U} and \mathbf{V} . In TSVD, we drop the smaller values in \mathbf{S} which corresponds to unimportant vectors causing numerical issues, and only keep the largest r values. Denote the truncated matrix \mathbf{S} as $\mathbf{S}_{\text{trunc}}$. The solution to the linear system \mathbf{h} is then approximated as

$$\hat{\mathbf{h}} = \mathbf{V} \mathbf{S}_{\text{trunc}} \mathbf{U}^\top \mathbf{TC}. \quad (19)$$

Note that $\mathbf{G}\hat{\mathbf{h}} = \mathbf{G}\mathbf{V}\mathbf{S}_{\text{trunc}}\mathbf{U}^\top\mathbf{TC}$ gives the predicted temperature measurements. Define $\mathbf{Q} = \mathbf{G}\mathbf{V}\mathbf{S}_{\text{trunc}}\mathbf{U}^\top$ to be the *smoother matrix* that maps the actual temperature measurements to the predicted ones. The trace of \mathbf{Q} , $\text{tr}(\mathbf{Q})$, represents the *effective degree of freedom* (i.e., effective number of parameters or model complexity) of the regularized model. For example, a larger r for TSVD would result in a less general model and thus larger $\text{tr}(\mathbf{Q})$. It is also worth noting that $\text{tr}(\mathbf{Q}) = r$ for TSVD.

TR. TR is another regularization technique whose central idea is that large solutions are results of overfitting the noises. Therefore, it adds a penalty term to the objective function to prevent large solutions, i.e.,

$$\begin{aligned}\hat{\mathbf{h}} &= \arg \min_{\mathbf{h}} (\|\mathbf{G}\mathbf{h} - \mathbf{TC}\|_2^2 + \rho^2 \|\mathbf{h}\|_2^2) \\ &= (\mathbf{G}^\top \mathbf{G} + \rho^2 \mathbf{I})^{-1} \mathbf{G}^\top \mathbf{TC}.\end{aligned}\quad (20)$$

Similar to TSVD, the smoother matrix for TR is $\mathbf{Q} = \mathbf{G}(\mathbf{G}^\top \mathbf{G} + \rho^2 \mathbf{I})^{-1} \mathbf{G}^\top$.

MLE. We are also able to avoid overfitting of noises by explicitly considering the noises as a probability distribution. We convert Eq. [16] to the following probabilistic model:

$$\mathbf{TC} = \mathbf{G}\mathbf{h} + \boldsymbol{\varepsilon}, \quad (21)$$

where $\boldsymbol{\varepsilon}$ denotes the random noises associated with each temperature measurement. As we are not able to imply the covariance of $\boldsymbol{\varepsilon}$ from the measurements alone, we make the assumption that for any $m = 1, 2, \dots, 2M$

$$\boldsymbol{\varepsilon}_m \sim \mathcal{N}(0, (1.94\% \text{TC}_m)^2). \quad (22)$$

This gives a valid estimation of $\boldsymbol{\varepsilon}$ since the signal-to-noise ratio in this problem is rather high⁵. As a result, the covariance of $\boldsymbol{\varepsilon}$, Σ , is a $2M \times 2M$ diagonal matrix with diagonal element $\Sigma_{mm} = (1.94\% \text{TC}_m)^2$.

MLE requires us to look for the most probable \mathbf{h} given the measurements \mathbf{TC} and random noises $\boldsymbol{\varepsilon}$, i.e.,

$$\begin{aligned}\hat{\mathbf{h}} &= \arg \max_{\mathbf{h}} \Pr[\mathbf{h} | \mathbf{TC}] \\ &= \arg \min_{\mathbf{h}} ((\mathbf{G}\mathbf{h} - \mathbf{TC})^\top \Sigma^{-1} (\mathbf{G}\mathbf{h} - \mathbf{TC})).\end{aligned}\quad (23)$$

By differentiation, $\hat{\mathbf{h}}$ is the solution of the linear system

$$(\mathbf{G}^\top \Sigma \mathbf{G}) \hat{\mathbf{h}} = \mathbf{G}^\top \Sigma \mathbf{TC}, \quad (24)$$

which can again be regularized using TSVD or TR.

4.4 Tuning of Regularization Parameters

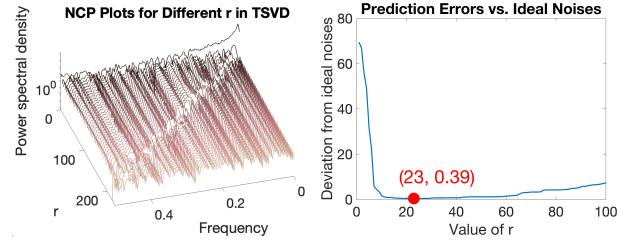
The number of diagonal elements kept r for TSVD and the scale of penalty ρ for TR are called *regularization parameters*. An unsuitable regularization parameter is not able to address the numerical issues well (see Fig. [7] for a demonstration of underfitting and overfitting with poorly chosen regularization parameters). Therefore, we present several criteria below to tune regularization parameters.

Discrepancy Principle (DP) [10]. From Eq. (22), we can approximate the expected squared sum of noises as

$$\sum_{m=1}^{2M} (1.94\% \text{TC}_m)^2 \approx 6.33 \times 10^4. \quad (25)$$

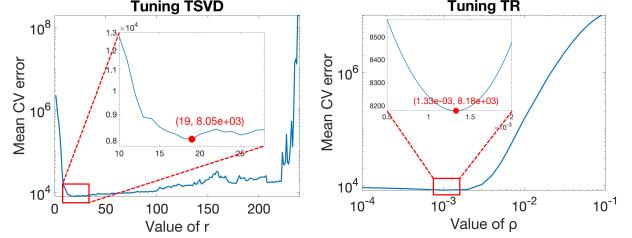
We follow DP to select the parameters r for TSVD and ρ for TR such that gives a squared error closest to Eq. (25) through iterative grid search.

⁵Nevertheless, the accurate estimation of noise distribution is important to probabilistic modelling. Real-life application may adopt advanced techniques such as [19].



(a) Periodograms for different regularization parameters.
(b) Parameter giving the smallest deviation is selected.

Figure 8: Demonstration of tuning with NCP.



(a) Bias-variance tradeoff of TSVD.
(b) Bias-variance tradeoff of TR.

Figure 9: Demonstration of tuning with CV.

Normalized Cumulative Periodogram (NCP) [1]. NCP uses the spectrum information of noises (i.e., the power accumulated across each frequency) (Fig. [8a]) to find the regularization parameter at which the prediction errors deviate the least from the noises (i.e., the nature prediction errors shifts from “signals” to “noises”) (Fig. [8b]).

Cross Validation (CV) [4]. CV is a powerful tool to diagnose overfitting. We adopt k -fold CV with $k = 10$ following the steps below:

- We first randomly split the dataset into $k = 10$ partitions.
- In each of the 10 iterations, we perform the inverse analysis with 9 partitions and validate the trained model with the remaining 1 partition (i.e., calculate the mean squared error between predicted temperatures and actual temperature measurements).
- Find the regularization parameter that gives the smallest mean squared error averaged across all iterations (i.e., mean CV error) through iterative grid search.

Fig. [9] shows the (empirical) bias-variance tradeoffs when we adjust the regulation parameters.

Generalized Cross Validation (GCV) [9]. As a generalization of CV, GCV aims to find the regularization parameter that minimizes the following objective:

$$\text{GCV} = \frac{\|\hat{\mathbf{G}}\mathbf{h} - \mathbf{TC}\|_2^2}{(2M - \text{tr}(\mathbf{Q}))^2}, \quad (26)$$

where \mathbf{Q} represents the smoother matrix as defined in Sec. [4.3]. In other words, the GCV objective represents the “goodness of fit” penalized by (effective) model complexity.

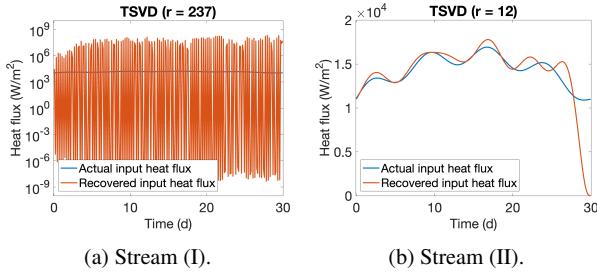


Figure 10: Comparison of Stream (I) (weak regularization) and (II) (strong regularization) in the final results. Note that Fig. 10a uses logarithmic-scale y -axis.

Akaike Information Criteria (AIC) [2]. AIC works by estimating the Kullback–Leibler divergence [24] between the actual data distribution and modelled data distribution (i.e., the information loss) as the following objective:

$$AIC = 2\text{tr}(\mathbf{Q}) + 2M \log \|\mathbf{G}\hat{\mathbf{h}} - \mathbf{T}\mathbf{C}\|_2^2. \quad (27)$$

Therefore, the aim is to find the regularization parameter that minimizes the AIC objective.

Bayesian Information Criteria (BIC) [15]. BIC works similarly to AIC but estimates the Bayes factor [14] instead:

$$BIC = \log 2M \cdot \text{tr}(\mathbf{Q}) + 2M \log \|\mathbf{G}\hat{\mathbf{h}} - \mathbf{T}\mathbf{C}\|_2^2. \quad (28)$$

Models with higher BICs have a higher posterior probability to be true. By comparing Eq. (27) and (28), we can also see that BIC places a higher penalty for complex models than AIC.

4.5 Results and Discussion

The best regularization technique(s) and parameter(s) diagnosed by each tuning criterion are shown in Tab. 1

The results diverge into two main streams: (I) DP, GCV and AIC that favour weak regularization (e.g., $r = 237$ for TSVD); and (II) NCP, CV and BIC that favour strong regularization (e.g., $r = 12$ for TSVD). To find out which stream performs better and reduces overfitting of noises, we plot the recovered input heat flux and compare it with the actual input heat flux in Fig. 10. Clearly, Stream (II) yields a better recovered input heat flux whereas Stream (I) still suffers from numerical issues. This is because tuning criteria in Stream (I) places insufficient penalty on model complexity (e.g., the coefficients of $\text{tr}(\mathbf{Q})$ is constantly 2 for AIC but $\log 2M$ for BIC). When the number of observations is large (e.g. $2M = 480$ in our setting), the penalty should also be adjusted accordingly.

On the other hand, for each tuning criterion, the differences among the four regularization techniques seem trivial. Therefore, all the mentioned regularization techniques are suitable for the problem.

5 Conclusion

In this report, we have reinstated the theoretical foundations and demonstrated the practicality of recovering the input heat flux in blast furnaces through inverse analysis. In

Parameter	Optimal Value	Objective
DP: Closest SSE to the noises		
TSVD, r	59	6.35×10^4
TR, ρ	3.05×10^{-4}	6.33×10^4
NCP: Deviation from ideal noises		
TSVD, r	23	0.39
TR, ρ	2.38×10^{-3}	0.39
CV: Minimized mean CV error		
TSVD, r	19	8.05×10^3
TR, ρ	1.33×10^{-3}	8.18×10^3
MLE (TSVD), r	26	8.60×10^3
MLE (TR), ρ	2.50×10^{-4}	8.61×10^3
GCV: Minimized GCV objective		
TSVD, r	237	0.34
TR, ρ	8.98×10^{-8}	0.31
MLE (TSVD), ρ	210	0.32
MLE (TR), ρ	1.39×10^{-7}	0.33
AIC: Minimized AIC objective		
TSVD, r	239	5.22×10^3
TR, ρ	5.24×10^{-8}	5.21×10^3
MLE (TSVD), r	220	5.24×10^3
MLE (TR), ρ	1.39×10^{-7}	5.38×10^3
BIC: Minimized BIC objective		
TSVD, r	12	5.48×10^3
TR, ρ	2.85×10^{-3}	5.51×10^3
MLE (TSVD), r	13	5.48×10^3
MLE (TR), ρ	2.29×10^{-3}	5.49×10^3

Table 1: Results of tuning. The best regularization technique(s) and parameter(s) diagnosed by each tuning criterion are marked red.

fact, such techniques have already been deployed in the iron-making industry (e.g., [23]).

This report may serve as a beginner’s guide for further research on related techniques. Possible advancements that can be made include:

- modelling heat transfer as a more complex boundary value problem. The 1-dimensional setting used in this report might over-simplify the problem;
- better estimation of noises;
- using more advanced statistical models, regularization techniques and tuning criteria.

Acknowledgement

This project serves as the final project of the SCE.A405 *Inverse Problems and Data Assimilation* course conducted by Prof. Amaya Kenji at Tokyo Institute of Technology.

All programmes used for this project can be found at the GitHub repository <https://github.com/snoidetx/inverse-analysis>.

References

- [1] Aburidi, M. J.; and Salman, A. 2018. Normalized Cumulative Periodogram Method For Tuning A Regularization Parameter of The EEG Inverse Problem. *The Journal of Middle East and North Africa Sciences*, 10(5725): 1–10.
- [2] Akaike, H. 2011. Akaike's information criterion. *International encyclopedia of statistical science*, 25–25.
- [3] Bonetto, F.; Lebowitz, J. L.; and Rey-Bellet, L. 2000. Fourier's law: a challenge to theorists. In *Mathematical physics 2000*, 128–150. World Scientific.
- [4] Browne, M. W. 2000. Cross-validation methods. *Journal of mathematical psychology*, 44(1): 108–132.
- [5] Buckwell, G.; and Buckwell, G. 1997. Curves, functions and equations. *Mastering Mathematics*, 300–321.
- [6] Cannon, J. R. 1984. *The one-dimensional heat equation*. 23. Cambridge University Press.
- [7] Cline, A. K.; Moler, C. B.; Stewart, G. W.; and Wilkinson, J. H. 1979. An estimate for the condition number of a matrix. *SIAM Journal on Numerical Analysis*, 16(2): 368–375.
- [8] Crépey, S. 2010. Tikhonov regularization. *Encyclopedia of Quantitative Finance*, 1807–1812.
- [9] Golub, G. H.; Heath, M.; and Wahba, G. 1979. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2): 215–223.
- [10] Goncharskii, A.; Leonov, A. S.; and Yagola, A. G. 1973. A generalized discrepancy principle. *USSR Computational Mathematics and Mathematical Physics*, 13(2): 25–37.
- [11] Gudder, S. P. 1970. A superposition principle in physics. *Journal of Mathematical Physics*, 11(3): 1037–1040.
- [12] Hansen, P. C. 1990. Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank. *SIAM Journal on Scientific and Statistical Computing*, 11(3): 503–518.
- [13] Hassani, S.; and Hassani, S. 2009. Dirac delta function. *Mathematical Methods: For Students of Physics and Related Fields*, 139–170.
- [14] Kass, R. E.; and Raftery, A. E. 1995. Bayes factors. *Journal of the american statistical association*, 773–795.
- [15] Konishi, S.; and Kitagawa, G. 2008. Bayesian information criteria. *Information criteria and statistical modeling*, 211–237.
- [16] Magnago, L. C. S.; Catabriga, L.; and Santo, E. 2014. Comparison of methods for solving inverse problems to estimate the wear line in a blast furnace hearth. *Blucher mechanical engineering proceedings*, 1(1).
- [17] Nellis, G.; and Klein, S. 2008. *Heat transfer*. Cambridge university press.
- [18] Pan, J.-X.; Fang, K.-T.; Pan, J.-X.; and Fang, K.-T. 2002. Maximum likelihood estimation. *Growth curve models and statistical diagnostics*, 77–158.
- [19] Pastor, D.; and Socheleau, F.-X. 2012. Robust estimation of noise standard deviation in presence of signals with unknown distributions and occurrences. *IEEE transactions on Signal Processing*, 60(4): 1545–1555.
- [20] Peacey, J. G.; and Davenport, W. G. 2016. *The iron blast furnace: theory and practice*. Elsevier.
- [21] Šolín, P. 2005. *Partial differential equations and the finite element method*. John Wiley & Sons.
- [22] Stakgold, I.; and Holst, M. J. 2011. *Green's functions and boundary value problems*. John Wiley & Sons.
- [23] Steel, N. 2011. Using Inverse Problems Methodology to Optimize Blast Furnace Operations. *Nippon Steel News No. 385*.
- [24] Van Erven, T.; and Harremos, P. 2014. Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory*, 60(7): 3797–3820.
- [25] Wang, Y.; Cheng, J.; Nakagawa, J.; and Yamamoto, M. 2010. A numerical method for solving the inverse heat conduction problem without initial value. *Inverse Problems in Science and Engineering*, 18(5): 655–671.