Thermophysical modelling and parameter estimation of small Solar system bodies via data assimilation

M. Hamm,1,2 I. Pelivan,3 M. Grott2 and J. de Wiljes1

1Institute for Mathematics, University of Potsdam, Karl-Liebknecht-Str. 24-25, D-14476 Potsdam-Golm, Germany
2German Aerospace Center (DLR), Rutherfordstr. 2, D-12489 Berlin, Germany
3Fraunhofer Heinrich Hertz Institute (HHI), Einsteinufer 37, D-10587 Berlin, Germany

ABSTRACT
Deriving thermophysical properties such as thermal inertia from thermal infrared observations provides useful insights into the structure of the surface material on planetary bodies. The estimation of these properties is usually done by fitting temperature variations calculated by thermophysical models to infrared observations. For multiple free model parameters, traditional methods such as least-squares fitting or Markov chain Monte Carlo methods become computationally too expensive. Consequently, the simultaneous estimation of several thermophysical parameters, together with their corresponding uncertainties and correlations, is often not computationally feasible and the analysis is usually reduced to fitting one or two parameters. Data assimilation (DA) methods have been shown to be robust while sufficiently accurate and computationally affordable even for a large number of parameters. This paper will introduce a standard sequential DA method, the ensemble square root filter, for thermophysical modelling of asteroid surfaces. This method is used to re-analyse infrared observations of the MARA instrument, which measured the diurnal temperature variation of a single boulder on the surface of near-Earth asteroid (162173) Ryugu. The thermal inertia is estimated to be $295 \pm 18 \text{ J m}^{-2} \text{ K}^{-1} \text{s}^{-1/2}$, while all five free parameters of the initial analysis are varied and estimated simultaneously. Based on this thermal inertia estimate the thermal conductivity of the boulder is estimated to be between 0.07 and 0.12 W m$^{-1}$ K$^{-1}$ and the porosity to be between 0.30 and 0.52. For the first time in thermophysical parameter derivation, correlations and uncertainties of all free model parameters are incorporated in the estimation procedure that is more than 5000 times more efficient than a comparable parameter sweep.

Key words: radiation mechanisms: thermal – methods: data analysis – methods: statistical – minor planets, asteroids: individual: (162173) Ryugu.

1 INTRODUCTION
Thermal conditions on atmosphereless, small Solar system bodies are governed by the thermophysical properties of the surface material, e.g. thermal conductivity, heat capacity, and emissivity. The thermal conductivity is coupled to structural properties of the surface material such as grain size and porosity (Sakatani et al. 2017). Observing the surface in the thermal infrared wavelength range, typically 5–25 $\mu$m, provides direct insight into the thermal conditions on the surface. Thus, thermophysical and structural material properties can be derived from thermal infrared data.

The thermophysical properties of numerous Solar system bodies have been investigated using telescopes (e.g. Masiero et al. 2011; Müller, Hasegawa & Usui 2014; Harris & Drube 2016; Müller et al. 2017), or satellite remote sensing data (e.g. Chase 1969; Kieffer et al. 1972; Kührt et al. 1992; Mellon et al. 2000; Christensen et al. 2001; Fergason, Christensen & Kieffer 2006b; Paige et al. 2010), as well as close-up studies performed by rovers and landers (e.g. Fergason et al. 2006a; Gómez-Elvira et al. 2012; Hamilton et al. 2014; Spohn et al. 2015; Vasavada et al. 2017; Grott et al. 2019). Recently, the Japanese Hayabusa2 mission (Watanabe et al. 2017) investigated the C-type near-Earth asteroid (162173) Ryugu with four instruments, including a thermal infrared imaging system (Okada et al. 2017; Sugita et al. 2019; Watanabe et al. 2019; Okada et al. 2020). The mission included the MASCOT lander (Ho et al. 2017) that, among other instruments, carried a thermal infrared radiometer (Grott et al. 2017). MASCOT landed on the surface of Ryugu and investigated a single boulder on the surface of Ryugu for 2.5 asteroid rotations (Jaumann et al. 2019; Preusker et al. 2019).
been successfully applied to highly non-linear problems with large number of free parameters of the order of 10^7 and its accuracy and stability has been rigorously investigated in recent years (e.g. de Wiljes & Tong 2019; de Wiljes, Stannat & Reich 2019; Lange & Stannat 2019).

DA techniques are widely employed in the Earth sciences, in particular in meteorology, atmospheric physics, and oceanography. For other Solar system bodies, DA has been applied to atmospheric data sets from orbital Mars missions (e.g. Montabone et al. 2006; Wilson et al. 2008). However, so far, DA has not been applied to the thermal infrared data sets gathered from small Solar system bodies.

2 METHODS

The method described in this section is one of the standard approaches for non-linear high-dimensional state estimation. Here, ‘state’ denotes the variables that describe the time-dependent condition of the system, i.e. the surface and subsurface temperature, as opposed to ‘parameters’ that govern the state, i.e. the thermophysical properties of the surface material.

The DA technique is designed to infer states and parameters of a dynamical system of interest on the basis of two sources of information: a model (typically given by an evolution equation of a state of interest dependent on partially unknown parameters) and partial and noisy observations of the system. At first, we will discuss the considered model and the associated observations followed by an introduction of the ESRF (Tippett et al. 2003; Reich & Cotter 2015).

2.1 Model

The thermophysical model used in this study is similar to the one used in Pelivan et al. (2017), Hamm et al. (2018), and Grott et al. (2019) and assumes the surface to be a semi-infinite and homogeneous half-space. The 1D-conduction equation is solved as follows:

\[
\frac{\partial T(x,t)}{\partial t} = \frac{\pi d^2}{\Omega} \frac{\partial^2 T(x,t)}{\partial x^2},
\]

where \( \Omega \) is the rotation period, \( T(x,t) \) is the time- and depth-dependent temperature, with \( x \) being the depth variable in the direction of the local surface normal and \( x = 0 \) at the surface. The depth is normalized to the diurnal skin depth \( d \), which is defined as

\[
d = \sqrt{\frac{k}{\rho c_p \pi}},
\]

where \( \rho \) is the density, \( c_p \) is the specific heat capacity, and \( k \) is the thermal conductivity. This normalization requires to assume \( k, c_p \), and \( \rho \) to be constants. At the lower boundary, the flux is set to zero. The upper boundary condition is given by the energy balance at the surface:

\[
(1 - \epsilon) I(t) = \epsilon \sigma_B T^4(x = 0, t) + \Gamma \sqrt{\frac{\pi d}{\Omega}} \frac{dT}{dx}(x = 0, t) + Q_\text{in}(t),
\]

where \( A \) is the surface Bond albedo, \( I(t) \) is the solar illumination, \( \epsilon \) is the thermal emissivity, and \( \sigma_B \) is the Stefan–Boltzmann constant. Further, \( Q_\text{in}(t) \) denotes the thermal radiation received from the surrounding terrain. The thermal inertia is defined as \( \Gamma = \sqrt{\rho c_p k} \) in units of J m^-2 K^-1 s^-1/2. This parameter is commonly used to describe the amplitude of the diurnal surface temperature variation.
and its phase shift with respect to maximum insolation. The higher a surface’s thermal inertia, the later it will reach its maximum temperature in the afternoon, and the smaller is the difference between day and night temperatures. This thermal model calculates the temperatures on a spatial grid. We chose this grid to consist of 41 points spread over eight diurnal skin depth with increasing distance, as described in Hamm et al. (2018) and Grott et al. (2019).

The aim now is to determine unknown parameters of interest of equation (3), e.g. the thermal inertia $\Gamma$, emissivity $\varepsilon$, etc., which result in a specific temperature profile. In this paper, we use a sequential DA algorithm to simultaneously estimate the temperatures on the 41 grid points, the state, as well as the model parameters. This is achieved by defining an ‘augmented’ state vector, which consists of the temperatures and model parameters and will be denoted $z(t) \in \mathbb{R}^{N_y}$ as follows:

$$z(t) = (T(0, t), T(x_1, t), ..., T(x_{41}, t), \Gamma, \varepsilon),$$

(4)

where the temperatures evolve according to the thermal model equation (1). Note that besides $\Gamma$, any parameter of the thermal model can be included in $z$.

While the model parameters are time independent, the DA requires some sort of evolution in time for a sequential improvement of the parameter’s estimate. Here, a Brownian motion is chosen to ensure that the parameter space is traversed sufficiently to converge into the true parameter value. Here, the forward model of the thermal inertia is defined as

$$\frac{d\Gamma}{d\tau} = \frac{dW_t}{d\tau},$$

(5)

where $W_t$ is a Wiener process, i.e. the mathematical description of the Brownian motion. This process is realized by adding, in each forecast step, a random number to the previous estimate:

$$\Gamma(\tau_0) = \Gamma(\tau_{n-1}) + \zeta_t(\tau_{n-1}),$$

(6)

where $\tau_0$ is the time at which the parameter is updated, $\tau_{n-1}$ is the time of the previous update, and $\zeta_t(\tau)$ is a random number drawn from a normal distribution $\mathcal{N}$ centered on zero and with a standard deviation of $\alpha_t(\tau_0)$. The update time $\tau$ in this study is equivalent to the time of the observations. This formalism can be applied to other parameters of the thermal model, each with their own choice of $\alpha$ as provided in Table 1.

### 2.2 Data

In order to employ DA techniques, observations that can be linked to the state of interest are required. The relationship between observations $y_{ob}$ and augmented state $z(t)$ can be written as

$$y_{obs}(\tau_0) = H z(\tau_0) + \nu(t),$$

(7)

where $\nu(t) \in \mathbb{R}^{N_y}$ is the observational noise that is assumed to be Gaussian distributed with zero mean and covariance matrix, $R \in \mathbb{R}^{N_y \times N_y}$, and $H \in \mathbb{R}^{N_y \times N_x}$ is the observation operator. Further note that the number of observed components $N_y$ is often significantly smaller than the dimension of the augmented state space $N_z$. In this study, only one component of $z(t)$ is observable. In the first part of the study, this is the surface temperature $T(0, \tau)$. In the second part, it is the radiance emitted by the surface and observed by the MASCOT radiometer. This means that $N_y = 1$. $R$ is a scalar corresponding to the measurement uncertainty, and $H$ is given by $H(1) = 1$ and $H(i) = 0$ for $i \neq 1$.

### 2.3 Sequential DA and the ESRF

In the following, we will briefly introduce basic concepts of sequential DA and the classic Kalman filter (KF) as an example for sequential DA. Then, we will introduce the ESRF, which is a prominent member of the family of ensemble Kalman filters (EnKFs; Evensen 2006).

#### 2.3.1 Classic KF

In sequential DA, the probability distribution of a system’s state is estimated by repeatedly applying two steps called forecast and analysis. Based on an initial estimate of the state, the state in the first time-step is predicted by applying a model, e.g. the thermophysical model described above. This prediction is corrected by incorporating an observation. The corrected prediction is referred to as analysis and utilized as the input of the model to predict the state in the next time-step. One iterates over this procedure for all subsequent time-steps. The forward model can be expressed by

$$z(\tau_0) = \Psi (z(\tau_{n-1})),$$

(9)

where $\Psi$ is the operator evolving the augmented state from time $\tau_{n-1}$ to $\tau_n$. For the temperature evolution, $\Psi$ is the solution to the PDE given in equation (1) evolving the temperatures $T(x, \tau_{n-1})$ to the temperature forecast $T(x, \tau_n)$. For the evolution of the model parameters, $\Psi$ is described in equation (6). Note that it is possible to add some noise in equation (9) to express uncertainties stemming from model or numerical errors.

The concept of sequential DA can be best illustrated by its standard version, the classic KF (Kalman 1960). It is valid for a linear dependence of the observation on the state $z$ and a linear forward model, i.e. a linear $\Psi$:

$$z(\tau_n) = z(\tau_{n-1}) + \delta \tau (A z(\tau_{n-1}) + b),$$

(10)

where $\delta \tau$ is the time-step, and $A$ and $b$ are model parameters. A common example is the estimation of the position of some vehicle based on measurements of velocity and position at given points in time (e.g. Hu et al. 2003). The KF has the underlying assumption that the posterior distribution,

$$p(z(\tau_n)|y_{obs}(\tau_1 : \tau_n)),$$

which describes the probability of the augmented state $z(\tau_n)$, given all the data $y_{obs}(\tau_1 : \tau_n)$ from time $\tau_1$ up to time $\tau_n$, is a Gaussian $N(\mu^{\nu}(\tau_n), P^{\nu}(\tau_n))$, where $\mu^{\nu} \in \mathbb{R}^{N_y}$ is the first and $P^{\nu} \in \mathbb{R}^{N_y \times N_y}$: the second moment of the associated normal distribution.

Bayes theorem connects this posterior distribution with the prior distribution that describes the probability of the augmented state $z(\tau_n)$, given all the data $y_{obs}(\tau_1 : \tau_n)$ from time $\tau_1$ up to time $\tau_n$. The prior distribution is assumed to be Gaussian $N(\mu^{\nu}(\tau_n), P^{\nu}(\tau_n))$ as well. This is the case when the model operator $\Psi$ and the observation operator $H$ are linear, and the initial value, the model, and

| $\alpha$ | $N_{int} = 1$ | $N_{int} = 2$ | $N_{int} = 3$ | $N_{int} = 4$ | $N_{int} > 4$
|---|---|---|---|---|---|
| $\Gamma$ | 10 | 5 | 5 | 2 | 2 |
| $\varepsilon$ | 0.01 | 0.005 | 0.005 | 0.002 | 0.001 |
| $f$ | 0.001 | 0.0005 | 0.0005 | 0.0002 | 0.0001 |
| $\theta$ | 10 | 5 | 2 | 2 | 1 |
| $\phi$ | 1 | 0.5 | 0.5 | 0.5 | 0.1 |

Note. Listed are the standard deviations of the Gaussian distributions from which the parameter forecasts are sampled according to equation (6).
The first part of the right-hand side of the expression is the likelihood, \( f(y_{\text{obs}}(\tau_n)|m^f(\tau_n)) \), which is maximal when the forecast observation \( H z^f \) is close to the observation \( y_{\text{obs}} \). The second part is the prior distribution that is maximal at its mean \( m^f \). One can rearrange the exponent to show that the posterior is a Gaussian with mean and covariance given by

\[
m^p(\tau_n) = m^f(\tau_n) - K(Hm^f(\tau_n) - y_{\text{obs}}(\tau_n))
\]

\[
P^p(\tau_n) = P^f(\tau_n) - KH(1/K)P^f(\tau_n),
\]

where \( K \) is the Kalman gain defined as

\[
K(\tau_n) = \frac{P^f(\tau_n)H^\top}{R + HP^f(\tau_n)H^\top}.
\]

The Kalman gain weights how much the analysis, i.e. the posterior distribution, is governed by the forecast produced by the model or the observation. Details of this derivation are given in chapter 6 of Reich & Cotter (2015). The smaller the observation error \( R \), the larger \( K \) becomes and the more is the observation weighted into the calculation of the analysis. Contrarily, if the observation error is very large, i.e.

\[
K \approx 0 \quad \text{(for } R \gg P^f)\]

and, consequently, \( m^p = m^f \) and \( P^p = P^f \). One can show that the updated mean \( m^p \) maximizes equation (12). One can furthermore show that \( m^p \) is the ‘best linear unbiased estimator’ of the state for a linear system, i.e. it minimizes the expectation value \( E[|m^p - x_{\text{true}}|^2] \). The mean and covariance of the updated posterior are used to forecast the state in the next time-step by applying the linear forward model (equation 10):

\[
m^f(\tau_{n+1}) = (I + \delta A) m^p(\tau_n) + \delta b
\]

\[
P^f(\tau_{n+1}) = (I + \delta A) P^p(\tau_n)(I + \delta A)^\top.
\]

Iterations over equations (13) and (15) provide a sequence of best linear unbiased estimators for a series of observations for systems with linear dynamics, which is illustrated in the lower panel of Fig. 1.

### 2.3.2 Ensemble square root filter

The classic KF can be extended to a non-linear model setting via an ensemble approach where an ensemble of \( M \) augmented state vectors \( z_i^f(\tau_n) \) and \( z_i^p(\tau_n) \) with \( i \in \{1, \ldots, M\} \) are generated in each time-step \( \tau_n \) to approximate the Gaussian prior and posterior distribution via the empirical posterior mean

\[
\hat{m}^p(\tau_n) = \frac{1}{M} \sum_{i=1}^{M} z_i^p(\tau_n)
\]

and covariance

\[
\hat{P}^p(\tau_n) = \frac{1}{M-1} \sum_{i=1}^{M} (z_i^p(\tau_n) - \hat{m}^p(\tau_n))(z_i^p(\tau_n) - \hat{m}^p(\tau_n))^\top
\]

for each \( \tau_n \) and analogously for the empirical prior distribution. This means that the EnKFs are Monte Carlo approximations of the classic KF. As in the classic KF, the forecast prior is updated incorporating the observation. The mean and covariance of the analysis have to fulfill the Kalman update in equation (13). However, while in the KF, it was sufficient to update mean and covariance, in the EnKF, each ‘ensemble member’, i.e. \( z_i^f \), has to be updated individually.

The iterative update procedure of the samples is described in Algorithm 1. The corresponding code will be made available.
upon request. At first, the initial ensemble of $M$ augmented state vectors $z_i(0)$ with $i \in \{1, \ldots, M\}$ are generated by sampling from Gaussian distributions $N(m(0), P(0))$, which are then, individually, sequentially updated by iterating over the forecast and analysis step.

**Algorithm 1** Ensemble Square Root Filter

Set variables $m(0)$, $P(0)$ and $M$

Initialise ensemble of augmented states $z_i(0) \sim N(m(0), P(0))$ with $i \in \{1, \ldots, M\}$ by means of $N(m(0), P(0))$

for $n = 1 : N$ do

Forecast:

$$z_i^f(\tau_n) = \Psi(z_i^f(\tau_{n-1})) \quad \forall i \in \{1, \ldots, M\}$$

Analysis step:

$$z_i^a(\tau_n) = \sum_{j=1}^{M} D_{ij} z_j^f(\tau_n)$$

with update $D_{ij}$ given in equation (20)

end for

The update matrix $D \in \mathbb{R}^{M \times M}$, which performs the update for each ensemble member, is constrained by the condition that, after the update, the mean and covariance of the ensemble members fulfill the Kalman update and the calculation of $D$ depends on which EnKF variant is used. The different versions can be divided into a stochastic branch and a deterministic one. For this study we choose the deterministic branch, the ESRF approach. The numerical success of the family of EnKFs has been documented for various applications (Evensen 2006) and there are rigorous accuracy and stability analyses available for the considered ESRF (de Wiljes & Tong 2019; de Wiljes et al. 2019). Further, the proposed deterministic branch of this family is preferable (Tippett et al. 2003; Næset et al. 2012) over the stochastic alternative (also known as perturbed EnKF), which is also very popular in the literature (Evensen 2003). The entries

$$D_{ij} = w_j - \frac{1}{M} + S_{ji}, \quad (18)$$

of $D$ depend on the the components $w_j$ of the weight vector,

$$w = \frac{1}{M} [1 - \frac{1}{M} S^T(E)^T H^{-1} R^{-1} (H h^f - y_{obs})] \in \mathbb{R}^M, \quad (19)$$

where $1$ is a column vector filled with ones and in $\mathbb{R}^M$. $E$ is a matrix in $\mathbb{R}^{N \times M}$. $E^T$ provides the distance between each ensemble member to the mean of the ensembles:

$$E^T = [z_1^f - \bar{h}^f, \ldots, z_M^f - \bar{h}^f]. \quad (20)$$

The matrix $S$ and its entries $S_{ij}$ that enter equation (18) are defined by

$$S = \left(1 + \frac{1}{M} (HE^T)^T R^{-1} HE^T\right)^{-1/2}. \quad (21)$$

where the matrix square root is defined as $B^{1/2}B^{1/2} = B$ for a matrix $B$. The name ‘ensemble square root filter’ refers to this matrix square root computation.

Note that the update occurs at the times $\tau$, i.e. the observation times. For the sake of readability, the time dependency is explicitly written in equations (18)–(21). The update matrix $D$ is constructed so that empirical mean and covariance are equal to the true mean and covariance of the posterior, $m^a = m^f$ and $P^a = P^f$, for a linear $\Psi$. In other words, the algorithm is designed to produce the same mean and covariance as the classic KF for a linear setting even for finite number of ensemble members $M$, whereas other EnKF versions only produce the KF mean and covariance in the ensemble limit $M \to \infty$, e.g. the perturbed EnKF (Evensen & van Leeuwen 2000). Further note that the update of each ensemble member depends on all other ensemble members, coupled through the empirical covariance matrix $P_0$ given in equation (17). For a more detailed derivation of the ESRF and its properties, see chapter 7 in Reich & Cotter (2015).

This form of update does not require the model to be linear, which is one of the key benefits of the ESRF compared to the classic KF. Furthermore, despite the underlying Gaussianity assumption, the non-linear evolution of the particles allows to capture the more complex behaviour of the system and thus leads to more realistic estimates.

### 3 NUMERICAL SIMULATION

The ESRF is tested for two cases. The first case is a proof of concept where the thermal inertia is derived in a controlled and simplified set-up with an artificial data set based on a reference solution of the thermophysical model. In the second case, it is employed to revisit the analysis of the radiometric data set retrieved by the MASCOT lander from the surface of the near-Earth asteroid (162173) Ryugu (Grott et al. 2019).

#### 3.1 Estimation of thermal inertia in a simplified model

The aim of this numerical example is to show how the technique performs in a controlled setting. This is achieved by generating an artificial reference temperature profile computed by means of a set of fixed reference parameters. In order to validate the performance of the proposed technique, the estimates obtained via the ESRF are compared to the reference temperature variation and reference thermal inertia.

##### 3.1.1 Reference solution

The reference temperature is simulated using the model given in equation (1) above with thermal inertia $I^\text{ref} = 300$ J m$^{-2}$ K$^{-1}$ s$^{-1/2}$, an albedo of 0.015, emissivity of 1, and 0 $Q_{th}$. The illumination is calculated by the simple assumption of a spherical asteroid, with equal length of day and night:

$$I(t) = I_{\text{max}} \cos \left(\frac{2\pi t}{\Omega}\right), \quad (22)$$

where $I(t) = 0$ if $\cos \frac{2\pi t}{\Omega} < 0$ and $I_{\text{max}} = 800$ W m$^{-2}$, similar to the illumination conditions on Ryugu. The rotation period $\Omega$ can be chosen arbitrarily and was set to the rotation period of Ryugu of $7.63262$ h (Watanabe et al. 2019).

##### 3.1.2 Initial ensemble

For each ESRF simulation, an initial ensemble of $M = 50$ members is generated and the thermal inertia values of these ensemble members are then drawn from a Gaussian distribution, $\Gamma_i = \Gamma^\text{mean} + \zeta_i$, with $\zeta_i \sim N(0, \sigma^2)$ and $\sigma = 20$ J m$^{-2}$ K$^{-1}$ s$^{-1/2}$. We repeat this procedure for 20 ESRF simulations, sampling $\Gamma^\text{mean} \sim N(250, 10^5)$ rather than running a single simulation of 1000 members with a standard deviation of $100$ J m$^{-2}$ K$^{-1}$ s$^{-1/2}$. We found that by doing so, we gain a more homogeneous sampling
of the initial distribution in parameter space. Furthermore, we save computation time as each of the 20 ESRF simulations converges quicker than a single simulation with a larger ensemble and the individual runs can be evaluated in parallel.

In order to save more computation time, a number of temperature profiles are pre-calculated assuming the parameters given above and varying the thermal inertia between 100 and 500 J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) in steps of 50 J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\). The ensemble member’s initial temperature profiles are than initialized by interpolating from the pre-calculated temperature profiles. These provided a more realistic initial guess for the temperature solution, accelerating the convergence of the PDE solver.

For each ensemble member, the temperature profile is sampled from a Gaussian distribution centred on the interpolated temperature profile with a standard deviation of 1 K. This method ensures that the ensemble is spread sufficiently to evolve through the parameter space while, at the same time, keeping the temperature profiles close enough to a physical solution to ensure convergence of the differential equation solver.

3.1.3 DA settings

The partial differential equation is solved using the MATLAB\(^{®}\) `pdpe`-solver for a total of 30 000 time-steps per simulated rotation, i.e. diurnal cycle. This corresponds to a time-step \(\Delta t = 0.91\) s.

For the Kalman update, 15 observation points are placed equidistantly in time from noon \((\tau = 0)\) to noon. The thermal model is run between these observations for 2000 \(\Delta t\) to forecast the temperature profile at the next observation, using the thermal inertia from the last Kalman update. The observation error is set to 1 K, which corresponds to setting the associated covariance matrix to \(\mathbf{R} = 1\).

The augmented state vector is then given by

\[
\mathbf{z}(\tau) = (T(0, \tau), T(x_1, \tau), \ldots, T(x_{15}, \tau), \Gamma(\tau)).
\] (23)

The thermal inertia evolves as described in equations (5) and (6), where the parameter \(\alpha(\tau)\) is reduced from one simulated rotation to the next. The width is varied from \(\alpha = 10\) J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) in the first rotation to \(\alpha = 5\) J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) in the second, \(\alpha = 1\) J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) in the third, \(\alpha = 0.5\) J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) in the fourth, and \(\alpha = 0.2\) J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) for the remaining 16 rotations. This gradual decrease of \(\alpha\) is in line with classic simulated annealing schedules often employed in the context of Monte Carlo methods. The key idea however is very intuitive, i.e. big \(\alpha\) help to traverse the parameter space more quickly, while they also prevent convergence of the ensemble members. Thus, lower \(\alpha\) values are chosen as the estimation procedure progresses in order to allow the posterior distribution to converge.

3.1.4 Results

The 20 ESRF simulations were run for 20 asteroid rotations starting with a randomly chosen thermal inertia each. Fig. 2 shows the histogram of the thermal inertia after 20 rotations. During the first few simulated cycles, the parameters spread out wide before converging. The last thermal inertia estimates of all ensemble members over all 20 simulations were combined into the histogram, i.e. 1000 thermal inertia estimates make up the final result of \(\Gamma = 299 \pm 4\) J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) with the uncertainty given by the 2\(\sigma\) bound, where \(\sigma\) is the standard deviation over the thermal inertia set. The thermal inertia of the reference temperature was 300 J m\(^{-2}\) K\(^{-1}\) s\(^{-1/2}\) and could thus be successfully retrieved.

Furthermore, the reference temperature could be retrieved well within the assumed 1-K uncertainty. The bottom panel of Fig. 2 shows the temperature estimates at the 15 observation points, where, at each point, the mean and standard deviation was taken over the last diurnal cycle. The error bar indicates the 2\(\sigma\) uncertainty.

This study demonstrates the working principle of the considered DA algorithm for the retrieval of thermophysical parameters from temperature observations. In the next step, the ESRF will be used to revisit the radiometric data obtained on the surface of asteroid (162173) Ryugu by the MARA instrument (Grott et al. 2017, 2019).

3.2 Thermal inertia estimation of Ryugu

The MARA instrument onboard the MASCOT lander observed the infrared flux emitted by the surface of a single, irregularly shaped boulder on Ryugu for a full diurnal cycle. The instrument consists of six infrared bolometers that are placed behind different infrared filters. The 8–12 μm (W10) filter was the instrument channel with the highest fidelity and was used for the initial analysis reported in (Grott et al. 2019). In that work, the night-time data were fitted by minimising the \(\chi^2\) value that measured the misfit between observed flux and the one predicted by a thermal model.

This thermal model included as free parameters the thermal inertia \(\Gamma\), emissivity \(\epsilon\), the orientation of the surface observed by MARA in terms of azimuth \(\theta\) and elevation \(\phi\) of the surface normal, and the view factor to the surrounding terrain \(f\) that parametrizes...
where the temperature of the surrounding is assumed to be equal to the brightness temperatures observed by MARA, $T_{\text{obs}}$, as described in Grott et al. (2019). It should be noted here that the roughness correction applied to day-time observation in Grott et al. (2019) did not influence the night-time temperatures, and is therefore omitted in this study.

The surface orientation had to be included as a free parameter as the observed spot on the irregular boulder showed a rugged texture with various parts of unknown orientation, and thus unknown illumination condition. The parameters $\theta$ and $\phi$ therefore represent an averaged surface orientation within the field of view of MARA.

The illumination is determined by the scalar product of surface orientation $\mathbf{n}$ and (time-dependent) solar vector $\mathbf{s}$:

$$I(t) = \frac{I_0}{r_h^2} \mathbf{n}(\theta, \phi) \cdot \mathbf{s}(t),$$

where $I_0$ is the solar constant and $r_h$ is the heliocentric distance.

In Grott et al. (2019), the parameter space was sampled by a grid search, where the thermal inertia was varied in steps of 1 J m$^{-2}$ K$^{-1}$ s$^{-1/2}$. However, due to the high computational cost of the thermal model, the other parameters were varied in significantly coarser steps, i.e. only three emissivity values were considered ($0.9, 0.95, 1$) along with only two values for $f (0$ and $0.08$). To test the efficiency of our new approach, this analysis was repeated using the ESRF.

### 3.2.1 Forecast settings

The forecasts of parameters and temperatures are again calculated using equation (1) and a forward model of the free parameters as in equation (5). The free parameters of the model were chosen analogous to the analysis of Grott et al. (2019): $\Gamma$, $\epsilon$, $\phi$, $\theta$, $f$. Also, the same grid settings were applied, i.e. we calculate the temperature profiles for a 1D grid with 41 grid points spread over eight diurnal skin depths. The augmented state is then given by

$$z(t) = [F_{W10}(t), T(0, t), T(x_1, t), ... T(x_{41}, t)]$$

where $\lambda$ is the wavelength, $B$ is the Planck function, and $q$ is the MARA filter throughput (Grott et al. 2017). The radiance observed by MARA is calculated from the reported brightness temperatures (Grott et al. 2019), i.e. $T$ in equation (27) is set to the brightness temperature and $\epsilon = 1$. Note, that the MARA signal depends linearly on the net flux between MARA and the surface. The calculation of the brightness temperature from the signal incorporates the temperature of the MARA sensor, the instrument field of view, sensitivity, etc., which does not need to be repeated in this study. As in the simplified case, the forecast of the parameters is performed according to equations (5) and (6). The value is again sampled from a Gaussian distribution where the standard deviation $\sigma(\tau)$ is stepwise reduced. An overview of $\alpha$ for the respective model parameters is provided in Table 1.

### 3.2.2 Data settings

The Kalman updates are performed at nine points of the night-time data equally distanced in time starting from 17:45 local time, which corresponds to the first data point considered in (Grott et al. 2019). The distance between the points is similar to the one in the first part of this study, except for the last night-time data point where the step to the first point in the next simulated night encompasses the entire asteroid day. The temperature is forecast by the thermal model, and converted into radiance received by the MARA W10 filter ($F_{W10}$), based on the instrument calibration (Grott et al. 2017, 2019), and the ensemble member emissivity.

The illumination is calculated for each ensemble member based on the angle between the sun vector and the surface normal according to equation (25), while azimuth $\theta$ and elevation $\phi$ of the surface normal are updated from observation to observation. Likewise, surface emissivity, the view factor $f$, and the surface thermal inertia are updated.

### 3.2.3 Initialization

The ensemble states are initialized similar to the first case. In total, 20 ESRF simulations are performed and for each simulation a thermal inertia is randomly picked from $\Gamma_{\text{total}} \sim N(300, 100^2)$. In each simulation, an ensemble with 50 members is initialized, and for each ensemble member, a thermal inertia is sampled from $\Gamma_t \sim N(\Gamma_{\text{total}}, 100^2)$. The thermal inertia is confined to an interval of $150-450$ J m$^{-2}$ K$^{-1}$ s$^{-1/2}$, a range that is larger than given by conservative estimates for Ryugu’s thermal inertia (Wada et al. 2018). For a sample of $\Gamma_t > 450$, the thermal inertia is set to 450, and if $\Gamma_t < 150$, it is set to 150. The emissivity of the ensemble members is sampled from a Gaussian distribution $\epsilon \sim N(1, 0.02^2)$ and confined to the interval of 0 and 1. Thereby, emissivity values larger than 1 are folded back into the interval, i.e. an $\epsilon = 1.05$ is set to 0.95, etc. The view factor to the surrounding terrain $f$ is sampled from $f \sim N(0.048, 0.007^2)$ based on the topography of the landing site as described in the methods section of Grott et al. (2019).

Azimuth and elevation of the surface normal in the best-fitting case of the initial MARA data analysis were found to be 20$^\circ$ and 80$^\circ$, respectively, where an azimuth of 0$^\circ$ is defined by the local east and an elevation 90$^\circ$ corresponding to 314:207 east and 34:599 south in Ryugu’s body fixed frame. For each ensemble member, elevation and azimuth are sampled from $\theta \sim N(20, 360^2)$ where values are confined to 0 and 360$^\circ$, e.g. $\theta = 361^\circ = 1^\circ$ or $\theta = -1^\circ = 359^\circ$, and $\phi \sim N(80,10^2)$ where $\phi > 90^\circ$ are flipped back, e.g. $\phi = 95^\circ$ is set to $\phi = 85^\circ$ as an elevation larger than 90$^\circ$ is not defined.

The temperatures of the ensemble members are initialized by interpolating the temperature profile from pre-calculated simulations. These pre-calculated simulations were performed for $\epsilon = 1$, $\theta = 20$, $\phi = 80$, and $f = 0$, while the thermal inertia was varied between 150 and 450 J m$^{-2}$ K$^{-1}$ s$^{-1/2}$ in steps of 50 J m$^{-2}$ K$^{-1}$ s$^{-1/2}$. It should be noted here that the resulting initial temperature curves are not consistent with the initial parameter sets of the ensemble members. However, this is not problematic as the ensembles are given enough time to produce consistent solutions. Rather, this initial temperature profile serves as a better first guess for a solution of the 1D-heat conduction equation, e.g. compared to assuming a constant temperature as in (Grott et al. 2019), and results in a quicker convergence of the temperature solution.
Figure 3. Parameter estimation from MARA data set. Panel (a): histogram of thermal inertia estimates of ensemble members. The solid red line indicates the mean, and the dotted red line indicates $2\sigma$, with $\sigma$ the standard deviation. Panel (b): deviation between the estimated radiance and the radiance emitted by the surface and observed by MARA as a function of time given in hours since the first MARA observation point (UTC 08:02:31) (black); the error bars indicate $2\sigma$ with $\sigma$ the standard deviation over all ensemble members and simulations. The estimated radiance is derived from the ensemble member’s emissivity, surface temperature, and the instrument calibration. The red area indicates the $2\sigma$ uncertainty of the observed radiance based on the instrument calibration. The estimates of the other free parameters are shown in the following four histograms: (panel c) emissivity, (panel d) integrated view factor, (panel e) surface normal elevation, and (panel f) surface normal azimuth. As in panel (a), all histograms are based on all ensemble members and simulations for the last Kalman analysis step.

3.2.4 Results

The DA method allowed for a much finer variation of the free parameters, resulting in a more thorough estimate of the thermal inertia.

Fig. 3 shows parameter estimates of the ensembles at the last Kalman analysis step combining 20 simulations with randomly chosen starting thermal inertia. Mean and uncertainty of the estimates are displayed by solid and dashed red lines respectively with the uncertainty given as two standard deviations ($2\sigma$). The histograms displaying the posterior distributions of the various, simultaneously estimated model parameters show a major advantage of this method, which can account for non-Gaussian distributions.

The thermal inertia was found to be $295 \pm 18 \text{ J m}^{-2}\text{ K}^{-1}\text{ s}^{-1/2}$. This result lies within the range of the former estimate but with lower uncertainty. The thermal inertia estimate is roughly Gaussian-distributed, with a slight tilt towards lower values, accounting for the fact that the effect of thermal inertia on the temperature variation decreases with increasing thermal inertia.

The other parameter distributions contain important information about the observed boulder. The estimated emissivity is very high and estimated to be between $0.95 \pm 0.05$. This is in line with the extremely dark appearance of Ryugu and the fact that roughness, as observed on the boulder in front of MASCOT, tends to increase the emissivity even further. The estimates for $f$, $0.08 \pm 0.02$, show that this parameter might have been underestimated in Grott et al. (2019), i.e. that the view factor of the observed spot towards the surrounding terrain is up to 10 per cent.

The elevation of the average surface orientation within the MARA field of view is estimated to be within $81^\circ \pm 2^\circ$, which is consistent with the camera images of the boulder surface in the field of view (Jaumann et al. 2019; Scholten et al. 2019b). The azimuth distribution shows that the most likely values lie within $304^\circ \pm 68^\circ$, i.e. oriented towards south-east. This is also the direction of the MARA boresight, which is consistent with the fact that those surface parts oriented towards MARA will contribute most to the signal. Also, due to the roughness effect, such an orientation would result in a systematically lower noon temperature as reported by (Grott et al. 2019). The former best-fitting azimuth of $20^\circ$ is less likely as the very flat peak in the posterior distribution indicates. However, many of the fitting models reported in Grott et al. (2019) showed an azimuth similar to the one retrieved in this work. Note that the dashed line in Fig. 3(f) at about $12^\circ$ represents the upper limit of the estimate as an azimuth angle of $372^\circ$ is equivalent to $12^\circ$.

The figure also shows that the estimated surface radiance matches the observed one very well. The major improvement of this analysis over the initial one is the full correlation among the parameter estimates and a smooth, simultaneous, and statistically thorough estimation rather than a rough parameter sweep. Despite the fact that 47 parameters are estimated simultaneously, including the 40 subsurface temperatures, one simulation run requires only 30 min on...
The considered ESRF allows for a simultaneous estimation of the state, i.e. surface and subsurface temperatures, as well as model parameters, i.e. thermal inertia, emissivity, surface orientation, etc., based on observed thermal infrared flux. Ensembles generated by the ESRF form a distribution that represents the uncertainties of state and parameters, while automatically including their respective correlations. As the performed forecast step is done on the basis of the thermal model without a linearization, the ensemble is able to better capture the non-linear relations better than commonly employed techniques such as the least-squares method.

The observations of the MARA instrument onboard the MASCOT lander were revisited in this work applying the ESRF. The results are consistent with the initial analysis of (Grott et al. 2019) but narrow the range of the thermal inertia estimate to $295 \pm 18 \, \text{J m}^{-2} \, \text{K}^{-1} \, \text{s}^{-1/2}$. At the same time, the emissivity could be constrained to $0.95 \pm 0.05$. The average surface orientation of around $81^\circ \pm 2^\circ$ elevation and $304^\circ \pm 68^\circ$ azimuth indicate that a significant fraction of the boulder in the MARA field of view is orientated towards the instrument. As these parts of the boulder face away from the sun during day, this result is consistent with reduced day-time temperature and the roughness effect reported in Grott et al. (2019).

In the first analysis of the MARA data, thermal conductivity $k$ and porosity $\phi$ of the boulder on Ryugu was estimated based on two empirical relations of $k(\phi)$ (Henke, Gail & Triloff 2016; Flynn et al. 2018). Repeating the calculation (Grott et al. 2019) for the thermal inertia estimate of this study, $295 \pm 18 \, \text{J m}^{-2} \, \text{K}^{-1} \, \text{s}^{-1/2}$, results in $k = 0.11 \pm 0.01 \, \text{W m}^{-1} \, \text{K}^{-1}$ and $\phi = 0.50 \pm 0.02$ using Flynn et al. (2018), and $k = 0.08 \pm 0.01 \, \text{W m}^{-1} \, \text{K}^{-1}$ and $\phi = 0.32 \pm 0.02$ using Henke et al. (2016).

An advantage of the ESRF scheme is its computational design to cope with large dimensions of state and parameter spaces, while being robust for non-linear systems. The possibility of estimating many parameters simultaneously enhances the scientific output of the remote sensing data. The parameters retrieved from the MARA observations are more accurate than previous estimates, as the ESRF discards unlikely parameter combinations and incorporates their correlations. At the same time, the parameters were sampled from a wide section of the parameter space and allowed to vary freely, limited only by basic physical limits, which should ensure that the parameter space was sufficiently covered.

The features of the family of EnKFs, i.e. coping with large dimensions of state and parameter spaces, provide great flexibility. The efficiency in handling many model parameters simultaneously sets the ESRF method apart from other Bayesian Monte Carlo methods such as the MCMC (Cambioni et al. 2019) or Particle filters (also ABC methods; see Ogawa et al. 2019). Further applications may include other thermal models and sets of parameters, e.g. such as modelling multiple ground layers with different thermal conductivity, or including temperature-dependent models of thermal conductivity and heat capacity (e.g. Vasavada et al. 2017). This will allow for an improved estimation of thermal properties on many other objects in the Solar system such as Mars, the Moon, or Comets.

**ACKNOWLEDGEMENTS**

MH was financially supported by Geo.X, the Research Network for Geosciences in Berlin and Potsdam – grant SQ_087,GeoX. The research of JdW and IP has been partially funded by Deutsche Forschungsgemeinschaft (DFG) – SFB1294/1 and 318763901. JdW was also supported by European Research Council (ERC) Advanced Grant ‘ACRCC’ (grant 339390) and by the Simons CRM.
Scholar-in-Residence Program. We thank Dr. Andrew Ryan for his constructive and very helpful review.

**CODE AND DATA AVAILABILITY**

The code and data used in this study will be made available by the corresponding authors upon request.

**REFERENCES**

Cambioni S., Delbo M., Ryan A. J., Furfaro R., Asphaug E., 2019, *Icarus*, 325, 16

Chase S., Jr., 1969, *Appl. Opt.*, 8, 639

Christensen P. et al., 2001, *J. Geophys. Res.*, 106, 23823

Christensen P. R. et al., 2018, *Space Sci. Rev.*, 214, 87

DellaGiustina D. et al., 2019, *Nat. Astron.*, 3, 341

de Wiljes J., Tong X., 2019, *Nonlinearity*, in press, https://arxiv.org/abs/1908.10580

de Wiljes J., Stannat W., Reich S., 2019, *SIAM J. Appl. Dyn. Syst.*, 17, 1152

Evensen G., 2003, *Ocean Dyn.*, 53, 343

Evensen G., 2006, *Data Assimilation: The Ensemble Kalman Filter*. Springer, Berlin, Heidelberg

Evensen G., van Leeuwen P. J., 2000, *Mon. Weather Rev.*, 128, 1852

Fergason R. L., Christensen P. R., Bell J. F. III, Golombeck M. P., Herkenhoff K. E., Kieffer H. H., 2006a, *J. Geophys. Res.*, 111, E12S21

Fergason R. L., Christensen P. R., Kieffer H. H., 2006b, *J. Geophys. Res.*, 111, E12004

Flynn G. J., Consolmagno G. J., Brown P., Macke R. J., 2018, *Geochemistry, 78*, 269

Gómez-Elvira J. et al., 2012, *Space Sci. Rev.*, 170, 583

Grott M., Knollenberg J., Borgs B., Hänischke F., Kessler E., Helbert J., Maturilli A., Müller N., 2017, *Space Sci. Rev.*, 208, 413

Grott M. et al., 2019, *Nat. Astron.*, 3, 971

Hamilton V. E. et al., 2014, *J. Geophys. Res.*, 119, 745

Hamm M., Grott M., Kürti E., Pelivan I., Knollenberg J., 2018, *Planet. Space Sci.*, 159, 1

Harris A. W., Drube L., 2016, *ApJ*, 827, 127

Henke S., Gail H.-P., Trieloff M., 2016, *A&A*, 589, A41

Ho T.-M. et al., 2017, *Space Sci. Rev.*, 208, 339

Hu C., Chen W., Chen Y., Liu D., et al., 2003, *J. Glob. Positioning Syst.*, 2, 42

Jaumann R. et al., 2019, *Science*, 365, 817

Kalman R. E., 1960, *J. Basic Eng.*, 82, 35

Kieffer H. H., Neugebauer G., Munch G., Chase S., Miner E., 1972, *Icarus*, 16, 47

Kührt E., Giese B., Keller H. U., Ksanfomality L., 1992, *Icarus*, 96, 213

Lange T., Stannat W., 2019, *TU Berlin, Technical Report*: https://arxiv.org/abs/1910.12493

Lauretta D. et al., 2019, *Nature*, 568, 55

Law K., Stuart A., Zygalakis K., 2015, *Data Assimilation: A Mathematical Introduction*. Springer-Verlag, New York

Masiero J. R. et al., 2011, *ApJ*, 741, 68

Mellon M., Jakosky B., Kieffer H., Christensen P., 2000, *Icarus*, 148, 437

Montabone L., Lewis S. R., Read P. L., Hinson D. P., 2006, *Icarus*, 185, 113

Müller T. G., Hasegawa S., Usui F., 2014, *PASJ*, 66, 52

Müller T. G. et al., 2017, *A&A*, 599, A103

Nerger L., Janjic T., Schroeter J., Hiller W., 2012, *Mon. Weather Rev.*, 140, 2335

Nowicki S. A., Christensen P. R., 2007, *J. Geophys. Res.*, 112, E05007

Ogawa K., Hammen M., Grott M., Sakatani N., Knollenberg J., Biele J., 2019, *Icarus*, 333, 318

Okada T. et al., 2017, *Space Sci. Rev.*, 208, 255

Okada T. et al., 2020, *Nature*, 579, 518–522

Paige D. A. et al., 2010, *Space Sci. Rev.*, 150, 125

Pelivan I., 2018, *MNRAS*, 478, 386

Pelivan I., Drube L., Kürti E., Helbert J., Biele J., Maibaum M., Cozzoni B., Lommatzsch V., 2017, *Adv. Space Res.*, 59, 1936

Preusker F. et al., 2019, *A&A*, 632, L4

Reich S., Cotter C., 2015, *Probabilistic Forecasting and Bayesian Data Assimilation*. Cambridge Univ. Press, Cambridge

Sakatani N., Ogawa K., Iijima Y., Arakawa R., Honda R., Tanaka S., 2017, *AIP Adv.*, 7, 015310

Scholten F. et al., 2019a, *A&A*, 632, L3

Scholten F. et al., 2019b, *A&A*, 632, L5

Spohn T. et al., 2015, *Science*, 349, 464

Sugita S. et al., 2019, *Science*, 364, 422

Tippett M., Anderson J., Bishop G., Hamill T., Whitaker J., 2003, *Mon. Weather Rev.*, 131, 1485

Vasavada A. R., Piqueux S., Lewis K. W., Lemmon M. T., Smith M. D., 2017, *Icarus*, 284, 372

Wada K. et al., 2018, *Prog. Earth Planet. Sci.*, 5, 82

Watanabe S. et al., 2019, *Science*, 364, 268

Watanabe S.-i., Tsuda Y., Yoshikawa M., Tanaka S., Saiki T., Nakazawa S., 2017, *Space Sci. Rev.*, 208, 3

Wilson R. J., Lewis S. R., Montabone L., Smith M. D., 2008, *Geophys. Res. Lett.*, 35, L07202

This paper has been typeset from a TeX/LaTeX file prepared by the author.