atmosphere

MDPI

Article

Using an Ensemble Filter to Improve the Representation of Temporal Source Variations in a Volcanic Ash Forecasting System

Meelis J. Zidikheri

Bureau of Meteorology, Melbourne 3001, Australia; meelis.zidikheri@bom.gov.au

Abstract: The use of ensemble models to forecast the dispersion and transport of airborne volcanic ash in operational contexts is increasingly being explored. The ensemble members are usually constructed to represent a priori uncertainty estimates in meteorological fields and volcanic ash source parameters. Satellite data can be used to further filter ensemble members within an analysis time window by rejecting poorly performing members, leading to improved forecasts. In this study, the ensemble filtering technique is used to improve the representation of temporal source variations. Ensemble members are initially created by representing the source time variations as random functions of time that are modulated by crude initial estimates of the variations estimated from satellite imagery. Ensemble filtering is then used to remove members whose fields match poorly with observations within a specified analysis time window that are represented by satellite retrievals of volcanic ash properties such as mass load, effective radius, and cloud top height. The filtering process leads to an ensemble with statistics in closer agreement with the observations. It is shown in the context of the 30 May 2014 Sangeang Api eruption case study that this method leads to significantly enhanced forecasting skill beyond the analysis time window—about 20% improvement on average—when compared to a system that assumes constant emission rates for the duration of the eruption, as is the case in many operational volcanic ash forecasting systems.

Keywords: volcanic ash; dispersion modeling; inverse modeling; ensemble modeling; quantitative forecasts; satellite retrievals; data assimilation; ensemble filter

Citation: Zidikheri, M.J. Using an Ensemble Filter to Improve the Representation of Temporal Source Variations in a Volcanic Ash Forecasting System. *Atmosphere* **2022**, *13*, 1243. https://doi.org/10.3390/atmos13081243

Academic Editor: Eugene Rozanov

Received: 20 July 2022
Accepted: 4 August 2022
Published: 5 August 2022

1. Introduction

It has long been recognized that improved quantitative volcanic ash forecasting models are required to enable airline operators to better manage the risks of flying during eruption events [1]. To achieve this goal, volcanic ash forecasting models need to better represent the uncertainties in forecasted meteorological fields, such as winds and rainfall rates, volcanic ash source parameters (including eruption heights and mass emission rates), and other dispersion model uncertainties (e.g., parameterizations used to represent various physical processes, such as wet and dry deposition). In recent years, significant strides have been made in this regard. The advent of ensemble numerical weather prediction (NWP) models, such as the Bureau of Meteorology's ACCESS-GE model [2], have enabled the operational implementation of ensemble volcanic ash forecasting models such as the Dispersion Ensemble Prediction System (DEPS) used by the Darwin Volcanic Ash Advisory Centre (VAAC Darwin). This system can represent uncertainty in the aspects of winds and rainfall, thereby improving forecasting skill and reliability.

As well as developments in NWP models, there have also been developments in volcanic ash remote sensing capabilities, both in hardware, such as the Himawari-8 satellite sensor [3], and in software, with the advent of advanced retrieval algorithms capable of retrieving ash properties such as mass load, effective radii, and cloud top heights [4–8]. These remote sensing advancements have enabled better estimates of volcanic ash source parameters, such as the volcanic mass emission rate, as a function of height and time by

making use of inverse modeling techniques [9–14]. These inverse modeling approaches can be described as variational methods. The goal of the variational algorithms is to minimize a cost function, which usually comprises terms that seek to find a compromise between an a priori solution based on simple physical considerations, such as the Mastin empirical relationship [15], and a solution yielding ash mass loads closer to the retrieved values. An optimal solution is found by taking into consideration the estimated errors of the a priori solution and the retrievals, which are usually assumed to be normally distributed. In some versions, the cost function may also contain other terms to further constrain the solution.

Inverse modeling can also be performed by ensemble filtering. In this approach, different ensemble members are generated by sampling the source parameter uncertainty space and using different NWP members to create different realizations of the modeled ash fields. This ensemble of test simulations is then filtered by keeping only those members whose pattern correlations, with respect to observations, are high. The pattern correlation cut-off number is computed to maximize the probabilistic Brier score so that areas with high probability of exceeding a series of thresholds coincide, on average, with areas in which observations exceed the series of thresholds. The methodology has been applied using semi-qualitative detection observations (i.e., binary fields indicating the presence or absence of pollutants) with eruption height as the control variable [16–18]. It has also been applied using mass load satellite retrievals, with additional parameters (see Section 2.1) controlling vertical emission rates, particle size distribution, and time variation as control variables [19]. It was also shown to be useful in improving forecasting skill when the volcanic ash dispersion model is driven by an ensemble of NWP fields with excessive spread [20].

Because the ensemble filtering approach is intrinsically based on a brute-force parameter-space sampling algorithm, the computational cost rises exponentially with the number of control variables, which limits its use to low-dimensional representations of the ash source term. Since the computational task is very intensive, even with about ten control variables that would be required to completely described the simple, low-dimensional source term, training datasets comprising satellite retrievals from past eruptions within the Darwin VAAC region are used to represent various source parameters (such as those describing emission rates as a function of height and particle size) as functions of eruption height [21]. This enables a considerable reduction in the number of control variables when the method is applied to eruptions in real time, thereby making it computationally tractable in operational contexts. Furthermore, it was shown that the "curse of dimensionality" that afflicts all ensemble filtering schemes, in which ensemble size is at times too severely diminished by the algorithm, could be ameliorated by dividing the ensemble members into smaller batches and applying the algorithm to each batch rather than to the whole ensemble [22]. The ensemble filtering method can also be used to reject poor retrievals or detections as initial fields when a direct data insertion scheme is used to initialize the dispersion models, as well as to find optimal top and base ash cloud heights, which can further improve forecasting skill, in some situations [23].

As alluded to above, the source term representation in the ensemble filtering approach needs to be a simple, low-dimensional, representation of the true source term to limit the number of control variables. This can limit the ability of the scheme to model more complex situations. For example, the time variation of the source in this scheme is a simple top hat function, which means that emissions remain constant for the duration of the eruption. This implies that events with very significant fluctuations in eruption strength, such as the 30 May 2014 Sangeang Api eruption in Indonesia, are difficult to model with this scheme. Zidikheri et al. (2017) [19] attempted to solve this problem by introducing an ad hoc function to model the fluctuations of the eruption height with time. However, this comes at the cost of introducing new parameters to the source term, which increases the number of control variables, thereby increasing the computational cost and making it difficult to implement operationally. In this study, we approach this problem by introducing a new ensemble generation scheme that can capture more complex features of the source term,

without the exponential rise in computational cost entailed by the brute-force sampling approach.

In this study, the approach to solve the time variation problem assumes that reasonable a priori estimates of the temporal source fluctuations may be made and corresponding uncertainties specified by keeping track of the fluctuations in cloud top height near the source. This could be accomplished, for example, by noting the fluctuation of brightness temperatures near the source, or indeed, of the retrieved cloud top height fields. A priori ensemble members are then generated by random sampling of the top height uncertainty and posterior ensemble members obtained by filtering the prior ensemble with respect to satellite retrievals. This process removes members that do not agree with the retrievals due to incorrect spatiotemporal patterns and leaves only members with similar patterns of spatial and temporal variability as in the retrievals. The rest of this paper is focused on describing the new scheme in more detail, applying it to the 30 May 2014 Sangeang Api eruption, and evaluating the results.

2. Methodology

In Section 2.1, the ensemble generation method that was used in the studies referred to in Section 1 ([16–23]) is described in detail. This method employs a top hat function for the time dependence of the volcanic source, and it will be used to create ensembles for the control run in this study. The new ensemble generation scheme that is the focus of the experiments in this study is described in Section 2.2; this method uses random functions to represent the time dependence of the source. Finally, in Section 2.3, the method by which ensemble members generated by both approaches are filtered with respect to observations is described. The Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) **[24]** model is used as the forward model in this study, and it is driven by the 24-member ACCESS-GE NWP ensemble model fields [2]. The observations are derived from the National Oceanic and Atmospheric Administration (NOAA)'s VOLcanic Cloud Analysis Toolkit (VOLCAT) [25] remote sensing software and comprise gridded fields of mass load, effective radius, cloud top height, and ash detection retrieved from MTSAT [26] satellite data (Himawari-8 satellite data only became available in late 2015, which is after the period of interest in this study). Further details about the setup of the HYSPLIT model and the satellite retrievals may be found in the work of Zidikheri and Lucas [22,23].

2.1. Ensemble Generation in Control Experiments

Ensemble members are generated by running a dispersion model with different members of the ACCESS-GE ensemble and by perturbing the parameters describing a cylindrically shaped ash source term whose mass distribution is given by

$$\dot{m}(t, \underline{x}, z, r) = \kappa \epsilon f_t(t) f_x(\underline{x}) f_z(z) f_r(r) \dot{M} \tag{1}$$

where κ is a scaling coefficient, and ϵ is the fine ash fraction; the time (t) variation of the source is described by

$$f_t(t) = \begin{cases} 1, & t_0 + \Delta t_0 < t \leq t_0 + \Delta t_0 + \tau \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

which is modeled as a simple top hat function and depends on the eruption start time t_0, the start-time lag, Δt_0, which accounts for umbrella-cloud effects, and the eruption duration, τ; the horizontal mass distribution (with $\underline{x}$ representing longitude and latitude)

$$f_x(\underline{x}) = \begin{cases} {}^{1}\!/\!_{A(D)}, & |\underline{x} - \underline{x_0}| \leq {}^{D}\!/\!_{2} \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

is simply a circle of diameter D, with area A, centered at the volcano location $\underline{x_0}$; the vertical (z) mass distribution

$$f_z(z) = \frac{c + \exp\left[-\frac{(z-\mu_h)^2}{2{\sigma_h}^2}\right]}{\int_{H_0}^{H}\left\{c + \exp[-\frac{(z-\mu_h)^2}{2{\sigma_h}^2}]\right\}dz}, \tag{4}$$

(with $H_0 < z < H$) comprises a constant emission rate c augmented by a Gaussian profile centred at μ_h, and with standard deviation σ_h, bounded by the summit height H_0 and the eruption top height H; the particle radius (r) mass distribution

$$f_r(r) = \frac{r^2 \exp\left[-\frac{(\ln r - \ln \mu_r)^2}{2(\ln \sigma_r)^2}\right]}{\int_0^R \left\{r^2 \exp\left[-\frac{(\ln r - \ln \mu_r)^2}{2(\ln \sigma_r)^2}\right]\right\}dr}, \tag{5}$$

(with $0 < r < R$, R being the maximum particle radius, specified to be 50 microns) is log-normal and depends on the mean particle radius μ_r and the standard deviation σ_r, the latter being fixed to a value of 1.5 microns (see Section 6 of [19]); the volcanic mass emission rate is denoted $\dot{M}$ and computed from the eruption height H using the Mastin relationship $\dot{M} = aH^b$ with a and b being empirical coefficients [15].

As is evident from Equations (1)–(5), the source term is dependent on the parameter set ($\underline{p_o}$, κ, τ, H), where $\underline{p_o} = (\Delta t_0, D, \mu_h, \sigma_h, c, \epsilon, \mu_r)$ is a collection of parameters that have been fitted to power–law relationships with eruption height, H, by using a dataset of past eruptions to estimate the fitting coefficients [21,22]. This implies that the only free parameters in this scheme are κ, τ, and H. This source parameter set is perturbed N_s times for each NWP member, drawn from a set of N_m NWP members, to yield an ensemble with $N = N_s \times N_m$ members; each ensemble member source term is written as $\dot{m}_k\left(t, \underline{x}, z, r; \underline{p_k}\right)$, with the subscript $k \in \{1, 2, 3, \cdots N\}$ labelling the ensemble member. The vector $\underline{p_k} = \left(\underline{\widetilde{p}_k}, \kappa_k, \tau_k, H_k\right)$ represents a collection of perturbed parameters values with

$$\underline{\widetilde{p}_k} = \underline{\varepsilon_k} \circ \underline{p_o}(H_k) \tag{6}$$

being the perturbed parameter values corresponding to the set p_o, which is a function of eruption height. Here, the vector $\underline{\varepsilon_k}$ is a collection of random coefficients that multiply $\underline{p_o}$ element by element (the so-called Hadamard or Schur product) to yield the parameter perturbation for each ensemble member k. The random numbers comprising $\underline{\varepsilon_k}$ are normally distributed (centered on $\underline{p_o}$) and correlated (i.e., perturbations of different parameters are not completely independent). The error covariance matrix that controls the magnitudes and correlations of the parameter perturbations is also computed from the dataset of past eruptions [21]. The parameters τ_k and H_k are considered independent and therefore are based on uniform random sampling within the range of uncertainty of these parameters, which are supplied directly to the system by the user. It is assumed that these uncertainty ranges would be based on factors such as the inspection of satellite imagery and ground or aircraft reports. In previous studies, only H was perturbed, and τ was simply specified and held constant to minimize computational cost; this approach shall also be followed when constructing control simulations in this study. The scaling parameter κ_k could also, in principle, be randomly sampled, but this is not practical and is not in fact necessary. It is instead simply estimated as the ratio of the spatial means of the observed and ensemble member mass loads. This ensures that, in a spatially averaged sense, the ensemble mass loads are always in agreement with observed mass loads for the time window in which observations are available. In previous studies [21], the scaling factor was simply included in the definition of the fine ash fraction, ϵ; that is, an effective $\epsilon' = \kappa\epsilon$ was defined and computed as the ratio of spatial means, as described above. Here, however, κ and ϵ are kept separate to enable a clearer comparison between this formulation and the new formulation

described in Section 2.2, which also makes use of a scaling factor, but it is computed in a slightly different way than that described above.

This simple source-term formulation has been demonstrated to work extremely well in the context of several case studies [21–23]. As alluded to above, the simplicity of the formulation enables the computation of power–law relationships between different parameters and eruption height from a training dataset (these are contained in the vector $\underline{p_0}$), which means that good source-term representations can be obtained, even when satellite data are not available in real-time, provided that the eruption height can be estimated by other means. This makes it a powerful tool in the tropics where satellite retrievals are frequently not available due to the presence of convective clouds [21], which make automatic detection of ash clouds difficult. However, the simplicity of the formulation also means that the scheme cannot easily handle situations where there is significant temporal variation. In the next section a new source-term representation scheme is developed with the goal of addressing this limitation.

2.2. Accounting for Temporal Source Variations

To better model eruptions with substantial time variations, instead of Equation (1), the following form for the source term is employed:

$$\dot{m}(t, \underline{x}, z, r) = \kappa \epsilon(t) f_x(\underline{x},\ t) f_z(z,\ t) f_r(r,\ t) \dot{M}(t) \tag{7}$$

In this formulation there is no explicit time dependence term $f_t(t)$ as in Equation (1), but the time dependence appears through the fine ash fraction $\epsilon(t)$, the horizontal function $f_x(\underline{x},\ t)$, the vertical function $f_z(z,\ t)$, and the Mastin relationship $\dot{M}(t)$. It should be stressed these factors have the same forms described in Section 2.1. The only difference is that here, the factors acquire time dependence through their dependence on the eruption height, which in this section, is taken to be a function of time. Each ensemble member k is associated with an eruption height function

$$H_k(t) = H^a(t) + \rho_k(t) \Delta H \tag{8}$$

In Equation (8), $H^a(t)$ is an a priori guess function, which, as alluded to previously, may be estimated from satellite imagery; ΔH is the uncertainty in the estimates $H^a(t)$, which for simplicity, is taken to be independent of time, and $\rho_k(t)$ is a sequence of random numbers sampled from the interval [–1, 1]. In general, the random function $\rho_k(t)$ may be generated such that it is correlated, to a degree, with the a priori function $H^a(t)$; that is, a constraint is placed on $\rho_k(t)$ such that the correlation

$$\mathrm{cor}(\rho_k,\ H^a) > \beta \tag{9}$$

is greater than a specified cut-off correlation β for each ensemble member. Different values of β are trialed in this study, as further discussed in Sections 3 and 4. Different random sequences ρ_k are repeatedly generated until the constraint described by Equation (9) is satisfied for each ensemble member. The eruption height test values $H_k(t)$ are then used to compute $\epsilon(t)$, $f_x(\underline{x},\ t)$, $f_z(z,\ t)$, $f_r(r,\ t)$, and $\dot{M}(t)$ in Equation (7). The scaling factor κ is computed to minimize the root mean square error (RMSE) of the simulated ensemble mean mass load with respect to satellite retrievals. This approach was found to yield better results than simply computing the ratio of the spatial average of the retrieved mass load to the simulated spatially averaged mass load for each ensemble member, as described in Section 2.1.

2.3. Ensemble Filtering

Having generated ensemble members from a priori distributions of source parameters, the ensemble is then filtered with respect to observations. In other words, a subset $\hat{K} \subseteq K$ is found, with $K = \{1, 2, 3, \cdots N\}$, so that members labeled by $\hat{k} \in \hat{K}$ describe the filtered

ensemble. As discussed in Section 2.2, the observations comprise a set of four fields, namely mass load, effective radius, cloud top height, and detections, labeled by the index f, which are defined at different geographical locations, $\underline{x}$, and times, t, and denoted by $\phi_o^{tf}(\underline{x})$. Likewise, corresponding fields may be computed from ensemble model simulations, denoted by $\phi_k^{tf}(\underline{x})$. Spatial pattern correlations

$$r_k^{tf} = \mathrm{cor}\left(\phi_k^{tf}, \phi_o^{tf}\right) \tag{10}$$

between each ensemble member k and the observations, for each observation time and field, are then found and averaged over the different times and fields to yield the average correlation $\bar{r}_k$ for each ensemble member. Using the average correlation values, weights w_k are then to allocated to each ensemble member

$$w_k(\alpha) = \begin{cases} 1, & \bar{r}_k \geq \alpha R \\ 0, & \bar{r}_k < \alpha R \end{cases} \tag{11}$$

with normalized form

$$\widetilde{w}_k\ (\alpha) = \frac{w_k(\alpha)}{\sum_{k=1}^{N} w_k(\alpha)} \tag{12}$$

Here R is the maximum correlation value obtained within the ensemble, and α is a parameter that defines the cut-off pattern correlation values below which the ensemble member is rejected. The cut-off value is not set a priori, but it is computed to maximize the multicategory Brier score (i.e., the ranked probability score) of the ensemble forecasts. To define the multicategory score, we need the categorical indicator field

$$d^{tfc}(\underline{x}) = \begin{cases} 1, & \phi_o^{tf}(\underline{x}) > \varphi^{fc} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

which is binary valued and indicates whether a field with index f at time with index t exceeds a categorical threshold with index c at a location $\underline{x}$. The categorical thresholds for a field with index f are denoted by φ^{fc}. These are the values used to define the thresholds for each field and are listed in Table 1. The corresponding 'indicator' field for the ensemble members is

$$e_k^{tfc}(\underline{x};\ \alpha) = \begin{cases} \widetilde{w}_k\ (\alpha), & \phi_k^{tf}(\underline{x}) > \varphi^{fc} \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Table 1. A description of the object denoted φ^{fc} in Section 2.3. This object lists the field thresholds used to compute the multi-categorical Brier scores as a function of field index, f, and category index, c. These values are the same as those used in previous studies and were chosen to broadly represent the different field values that are possible, both in retrieved and simulated values.

Field Number (f)	Field Name and Units	Category Number (c)	Threshold Values
1	Mass load (g/m^2)	1	1.0
1	"	2	10.0
1	"	3	100.0
1	"	4	1000.0
2	Effective radius (μm)	1	0.0
2	"	2	5.0

Table 1. *Cont.*

Field Number (*f*)	Field Name and Units	Category Number (*c*)	Threshold Values
2	"	3	10.0
2	"	4	15.0
2	"	5	20.0
3	Cloud top height (km)	1	0.0
3	"	2	5.0
3	"	3	10.0
3	"	4	15.0
3	"	5	20.0
3	"	6	25.0
4	Detection (g/m^2)	1	0.1

The exceedance probability (i.e., the fraction of ensemble members exceeding a threshold) is defined as

$$P^{tfc}(\underline{x};\alpha) = \sum_{k=1}^{N} e_k^{tfc}(\underline{x};\alpha) \tag{15}$$

and finally, the multicategory Brier score is defined as

$$s(\alpha) = \frac{\sum_{t=1}^{N_t}\sum_{f=1}^{N_f}\sum_{c=1}^{N_c^f}\int\left[P^{tfc}(\underline{x};\alpha) - d^{tfc}(\underline{x})\right]^2 d\underline{x}}{N_t N_f N_c \int d\underline{x}} \tag{16}$$

In Equation (15), the spatial averaging is indicated by integrals, while averages of times, fields, and categories are indicated by sums. This is simply for notational convenience; in practice, the averaging is performed in the same way (as discrete sums, rather than integrals), although the number of spatial points tends to be much larger than the number of times, fields, and categories. To find the optimal value of the cut-off parameter α, a crude grid search algorithm is employed, which means the Brier score needs to be computed several times to find its minimum value. Once the optimal α is computed, the binary-valued weights may be computed from Equations (11) and (12). If the weight of an ensemble member is greater than zero, then it is accepted; otherwise, it is rejected. In this way, the accepted members will yield probabilistic scores that are in closer agreement to observations than the original ensemble.

3. Eruption Case Study, Experimental Setup, and Verification Methods

3.1. Case Study

Sangeang Api, a volcano on an island to the northeast of Sumbawa, Indonesia (8.2° S, 119.07° E, 1.95 km summit elevation) erupted on 30 May 2014. The first and strongest phase of the eruption was short and occurred at around 0800 UTC, leading to a plume well exceeding 15 km in height, possibly as high as 20 km from some reports at the time of the eruption. The second eruption phase was longer and started at around 1700 UTC. This phase comprised at least three distinct "pulses" (i.e., intensification of volcanic activity) at around 1700, 2000, and 2300 UTC. This eruptive behavior is clear in Figure 1a. This figure shows the minimum 11 μm satellite brightness temperature over a small region (about 1 degree $\times$ 1 degree in size) centered at the volcano. Because at that point in time, there were no meteorological clouds around the volcano, the sharp drops in brightness temperature seen at 0832, 1732, 2032, and 2332 UTC are clearly associated with the rapid emergence of high-altitude, optically thick volcanic plumes. In this study, we shall focus on the second eruption phase starting around 1700 UTC, as it provides an excellent case study to ascertain

the impact of strong fluctuations in source strength over the course of an eruption on the forecasting skill, therefore providing a good test for the new methodology described in Section 2.2. Based on the brightness temperature, the eruption height during the second phase was somewhat lower than in the earlier phase. Straightforward linear regression of brightness temperature against the ACCESS-GE control member vertical temperature profile at the closest grid point to the volcano suggests peak heights of 13–14 km (as indicated on the right-hand side of Figure 1a), assuming the cloud is very optically thick and behaves like a black body. The corresponding VOLCAT satellite retrievals of cloud top height shown in Figure 1b, on the other hand, suggest peak heights exceeding 16 km. Note that at 1732 UTC, VOLCAT did not provide a valid retrieval, most likely due to the ash cloud being too optically thick, and therefore the heights in this case are just negative fill values, which should be ignored. The difference of a few kilometers in the height estimates between straightforward brightness-temperature analysis and VOLCAT estimates is possibly due to different assumptions regarding the emissivity of the cloud and the use of different NWP fields in the VOLCAT algorithm.

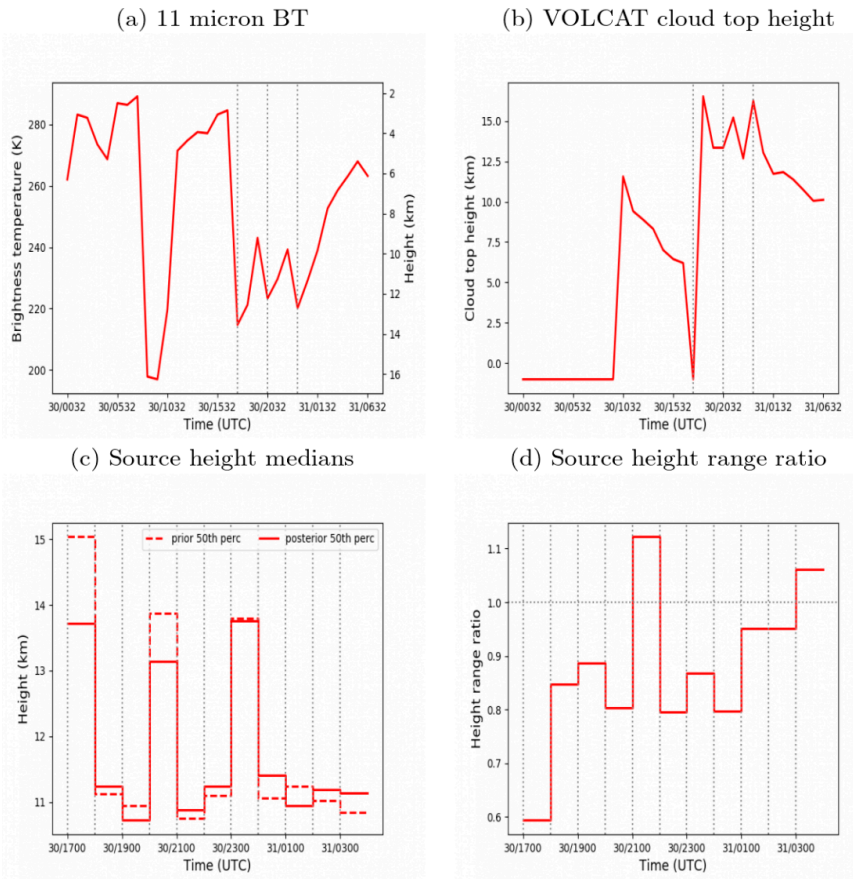

Figure 1. (**a**) Minimum satellite brightness temperature values in an approximately 1° × 1° box centered at the location of the Sangeang Api volcano. The vertical lines indicate times of interest, namely 1732, 2032, and 2332 UTC, at which significant eruptions occurred. (**b**) Same as in (**a**), but for maximum VOLCAT cloud top height field values. (**c**) Prior and posterior median (50th percentile) eruption heights as a function of time used in Experiment 1 of this study. (**d**) Ratio of posterior to prior source height ranges, defined as the difference between the 90th and 10th percentile heights in the respective distributions; ratios less than 1 indicate reduction in uncertainty.

Figure 2 shows the mass load retrievals from VOLCAT at various times during the eruption. Figure 2a shows the mass loads at 30/1930 UTC. At this point in time, the mass load structure is predominantly unimodal; the pattern is dominated by the high mass load area from the 1700 UTC eruption that has been swept to the southeast of the volcano. By

30/2130 UTC (Figure 2b), the ash from the 2000 UTC eruption has created a more bimodal pattern, with high mass loads appearing immediately to the southeast in the vicinity of the volcano. A trimodal pattern can be discerned by 30/2330 UTC (Figure 2c), with the eruption at around 30/2300 UTC also contributing to the mass load pattern. Figure 2d shows that by 31/0330 UTC, the eruption is continuing, but the emissions are more uniform, with less distinctive modal patterns discernable than during the early stage of the eruption. The corresponding cloud top height patterns are shown in Figure 3. The patterns of variability in these fields closely mirror the mass patterns. This suggests that the high mass load patterns are associated with the eruption being in a stronger phase and therefore reaching greater heights in the atmosphere.

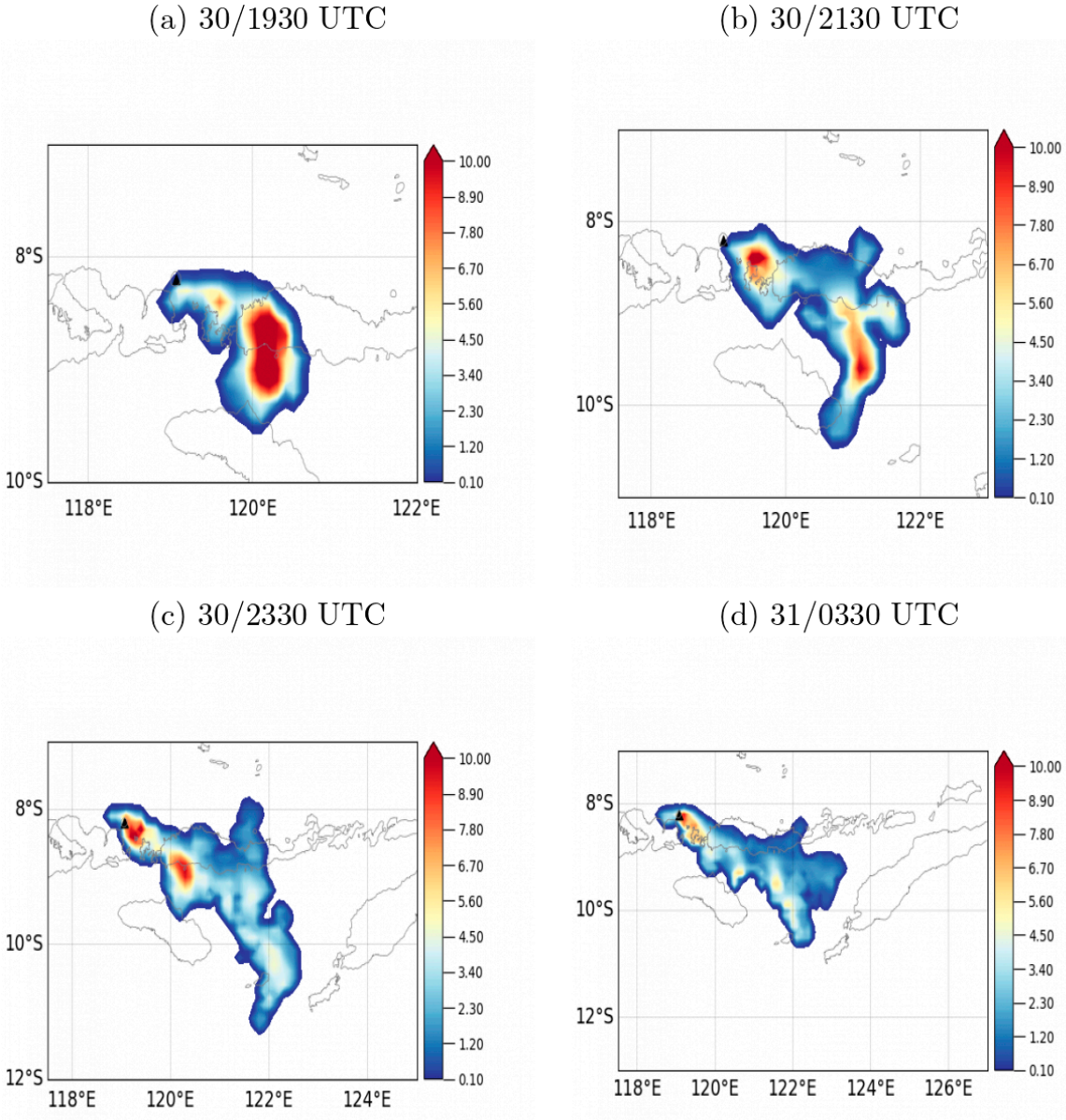

Figure 2. VOLCAT mass load retrievals (in g/m^2) corresponding to the 30 May Sangeang Api eruption at times indicated on the panels. The black triangle indicates the volcano location.

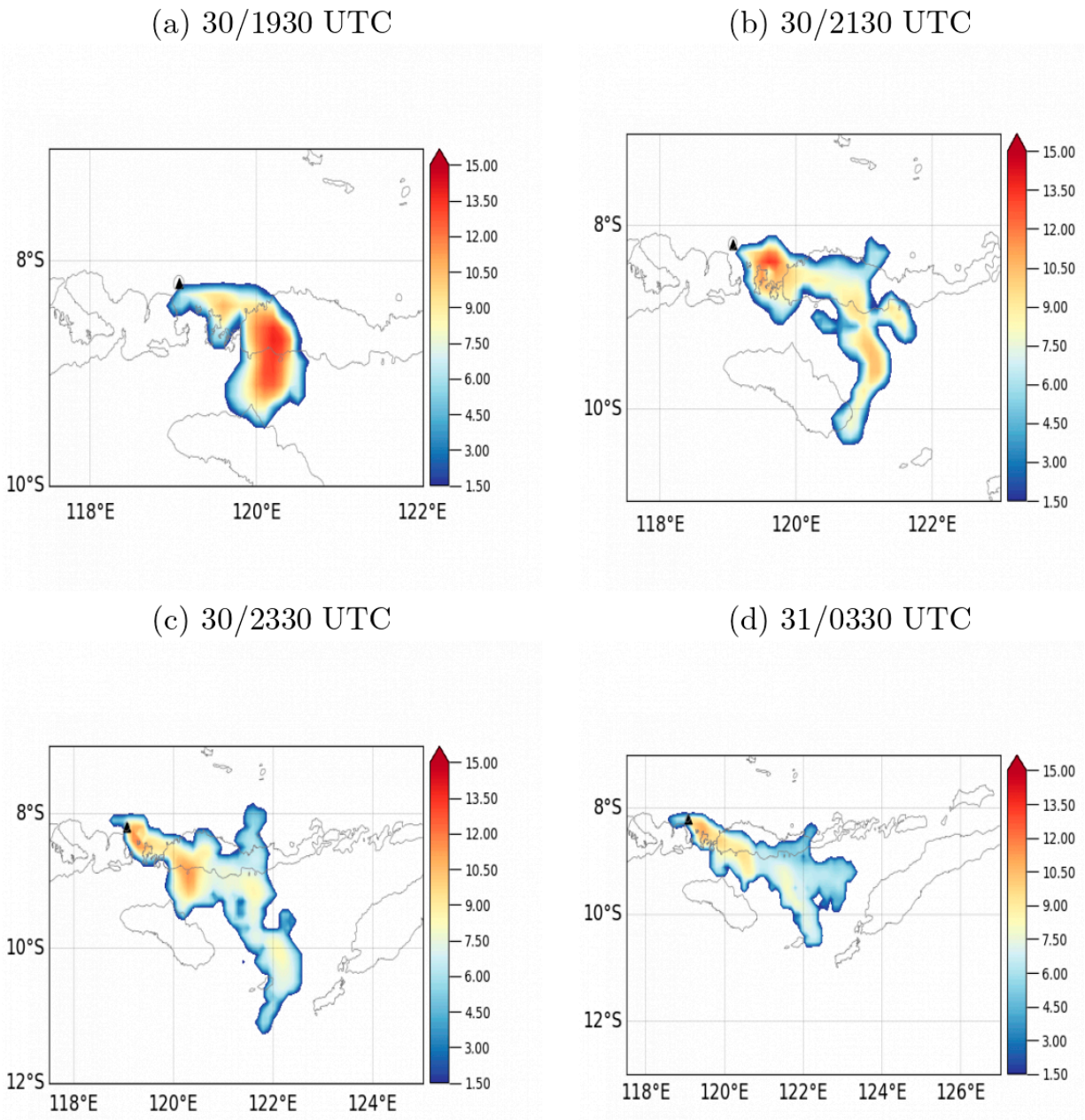

Figure 3. VOLCAT cloud top height retrievals (in km) corresponding to the 30 May Sangeang Api eruption at times indicated on the panels. The black triangle indicates the volcano location.

3.2. Experimental Setup

The HYSPLIT dispersion model, which is the forward model in this study, is set up in the same way as in previous studies [23]. The gridded wind fields used by HYSPLIT are provided by the Australian Bureau of Meteorology's ACCESS-GE ensemble NWP model, which in the version used in this study, comprises 24 members that represent uncertainty in the NWP fields.

The ACCESS-GE fields (in HYSPLIT ARL format) for the time period corresponding to the 30 May 2014 Sangeang Api eruption are archived and available for download at http://dx.doi.org/10.25914/6142ad42d977e (accessed on 1 December 2021). In the control run, the ash source term is time-invariant for the duration of the eruption, as described in Section 2.1. The new methodology described in Section 2.2 is used to construct the source term for all other experiments in this study. In the new approach, time variation is introduced by allowing the eruption in the model to vary with time in a random fashion, given prior

constraints on this variation. Both approaches involve the use of satellite retrievals of volcanic ash properties to estimate the source term by inverse modeling, with VOLCAT satellite retrievals being used for this purpose. VOLCAT data corresponding to the 30 May Sangeang Api eruption were downloaded from http://doi.org/10.5281/zenodo.3579613 (accessed on 1 December 2021) for use in this study.

In both the control run and experiments, simulations were performed in the period 30/1700 UTC to 31/0330 UTC. The simulation start time was chosen to correspond to the start of the second eruptive phase, and the stop time was chosen to correspond to the latest time at which good quality VOLCAT data were available to evaluate the results. This period was broken into a 6.5 h analysis period, lasting until 30/2330 UTC, in which inverse modeling was used to optimize the simulations, and a forecast period, 4 h in duration, in which the simulations were not coupled to VOLCAT data (the data was only used for evaluation in this period). The length of the analysis 'time window' was dictated by the need to simulate the fluctuations in eruptive activity, which lasted until at least 30/2330 UTC, as can be clearly seen in Figure 1a.

In the control experiment, the free source parameters are the eruption height, H, and the eruption duration, τ. Brightness temperature analysis and VOLCAT retrievals (e.g., Figure 1a,b) suggest peak eruption top heights over 16 km, with troughs just below 10 km over the course of the eruption. However, to account for the possibility of clouds being higher than these estimates, the 10–20 km range was chosen as a prior estimate from which random samples were drawn to generate source terms for the ensemble members. The start time was chosen to be 30/1700 UTC to correspond to the start of the second phase of the Sangeang Api eruption, and the eruption duration, τ, was chosen to be long enough to cover the whole simulation period, which was 10.5 h.

In all other experiments that used the new temporally varying formulation of Section 2.2, the median of the a priori height function $H^a(t)$ that was used is shown in Figure 1c; the uncertainty ΔH = 3 km was used in Equation (8) based on the discussion in Section 3.1. The simulation period was broken up into 11 1 h periods in which eruption heights (and therefore, emission rates) were uniform; at longer timescales, however, eruption heights were allowed to vary. In what follows, times associated with eruption heights refer to the start of each 1 h uniform emission period. The prior estimate of the eruption height at 1700 UTC is 15 km, which is roughly in between the maxima of 16.4 km, seen at 1832 UTC in the VOLCAT estimate (Figure 1b), and the value of about 13.5 km, obtained from brightness temperatures (Figure 1a). The peak values at 2000 and 2300 UTC are specified to be 1 km lower because the corresponding brightness temperature troughs are higher, indicating lower peak heights. The prior height of 11 km at 1800, 1900, 2100, and 2200 UTC is a compromise of the values of less than 10 km, obtained by brightness temperature analysis, and values of just over 12 km, obtained with VOLCAT.

It is assumed that information beyond 2330 UTC is unavailable at the time the analysis is performed to try to replicate a real-time operational scenario as closely as possible. Figure 1a shows that the eruption slowly reduced in strength over the next few hours, but this would have been hard to predict if the analysis were performed at around 2330 UTC, which is a key assumption in these experiments. Instead, it is assumed that the eruption continues in its 'neutral' mode (at about 11 km, which is the approximate height to which the eruption seemed to settle to in between the periods of explosive activity within the analysis time window) over the course of the forecast period (in this case, 4 h). Of course, other assumptions may be made about the emissions beyond the analysis time window. However, since the goal of this study is to compare the performance of the new temporal variation scheme against the control, if similar assumptions about future emission rates in both the control and experiment are used, then the details of the assumptions about future emission rates should not significantly affect the results. In both the control and the experiments, 10 random samples of the eruption height function were generated for each of the 24 NWP ensemble members, yielding 240 ensemble members in total. Hence, both control and test experiments had an equal number of members prior to filtering.

Four different time varying experiments were performed. Each experiment employed a different value of the correlation cutoff β and was labeled 1–4 in order of increasing β, as indicated in Table 2. Increasing the value of β creates temporal perturbations that are more strongly correlated with the prior height function $H^a(t)$. This can be a useful feature if the uncertainty ΔH is large, but there is a high degree of confidence in the temporal variation in $H^a(t)$. One of the goals of this study is to investigate appropriate values of β.

Table 2. Different experiments used in this study. The control run utilizes the simple, low-dimensional, time-invariant source term formulation described in Section 2.1 (and in previous studies). The other experiments employ the new time variable formulation of Section 2.2 and differ in the correlation cutoff β used to constrain the random temporal perturbations in the source height; $\beta = 0$ represents unconstrained perturbations while $\beta \to 1$ represents perturbations highly correlated with the a priori height function $H^a(t)$.

Experiment	β
0 (Control)	-
1	0.0
2	0.5
3	0.7
4	0.9

3.3. Verification Methods

The root mean square error (RMSE) metric was used to verify the ensemble mean forecasts relative to observations for each of the four fields used in this study, namely, mass load, effective radius, cloud top height, and detection. The ranked probability score (RPS), with 10 threshold values, ranging from the maximum observed field value to the minimum, was used to verify the ensemble forecasts. In addition, skill scores, which compare the values of these metrics for test experiments relative to control experiments, were computed as follows:

$$X_s = 1 - \frac{X}{X_c} \tag{17}$$

Here X_s is the skill score for a verification metric X, such as RMSE (RMSESS) or RPS (RPSS), that is zero for a perfect score and increases for poorer scores, and X_c is the metric value for the control experiment. The skill score is 1 for a perfect model (in which $X = 0$, $X_c > 0$), 0 for a model with no improvement ($X = X_c$), and $-\infty$ for an infinitely poorer model ($X/X_c \to \infty$). One of the advantages of a skill score, particularly for RMSE, is that scores for different fields can be meaningfully averaged to yield a single value, which is otherwise difficult to accomplish with a dimensional metric such as RMSE, as, in general, the different fields have different units.

4. Results

4.1. Source Term

As discussed in Section 3.2, in the control run, the source is time invariant for the duration of the eruption. However, in the experimental runs, the source varies with time due to the temporal variation of the eruption height, as formulated in Section 2.2. The height variation is based on multiple random sampling of the prior eruption height uncertainty space, whose median value is shown in Figure 1c; these heights are used to generate prior ensemble members. These members are then filtered with respect to observations (VOLCAT satellite retrievals) within an analysis time window, which is about 6 h, to yield posterior ensemble members. The posterior members are propagated forward in time for the duration of the forecast, which is 4 h in this case, but generally up to 24 h in operational contexts. As alluded to in Section 3, 4 experimental runs (labeled 1–4) were performed, each employing a different level of correlation between the prior height function and the random temporal perturbations that were employed to model the variation of the eruption height with time. Figure 1c also shows the median posterior source height function for

the experiment labeled as 'Experiment 1,' which employed the value of $\beta = 0$ (effectively meaning that perturbations were unconstrained in their temporal variability pattern). The biggest changes to the median posterior function, which are a result of ensemble filtering, include the drop in the peak heights at 1700 UTC and 2300 UTC. The changes at other times are smaller in comparison. These shifts result in a temporal pattern with two similarly sized peaks at 1700 and 2300 UTC, with a smaller peak at 2000 UTC. The relatively smaller size of the peak at 2000 UTC is consistent with the brightness temperature and VOLCAT retrievals shown in Figure 1a,b. Figure 1d shows the ratio of the posterior and prior distribution ranges, here defined as the difference between the 90th and 10th percentile height values in the distribution. Since this ratio is, for the most part, less than 1, this demonstrates that the posterior distribution is associated with lower height uncertainty values than the prior, which demonstrates that information from the observations has been assimilated into the forecasting system.

4.2. Ensemble Mean Field Patterns

Figure 4 shows the ensemble mean mass load (in g/m^2) obtained from the control experiment, which utilizes the simple time-invariant source formulation described by Equation (2). The mass load patterns are unimodal within the analysis time window (Figure 4a–c), with high loads in the vicinity of the volcano. These patterns are not in very good agreement with the VOLCAT mass load retrievals patterns in Figure 2. Similarly, the ensemble-mean cloud top height fields (in km) in Figure 5 do not replicate the fluctuations seen in the corresponding VOLCAT retrievals in Figure 3. This contrasts with the excellent performance of the algorithm in earlier studies (see references in Section 1), which did not involve eruptions with such strong temporal fluctuations. Hence, this case study demonstrates the shortcomings of the simple time-invariant source formulation when applied to eruptions with such characteristics.

In contrast to the control experiment, Figure 6, based on Experiment 1 (see Discussion in Section 4.1), demonstrates the utility of the new formulation when applied to eruptions with strong temporal variability. The peaks in mass load during the analysis phase (Figure 6a–c) correlate quite well with the peaks in the VOLCAT mass load retrievals in Figure 2a–c. Similarly, the ensemble mean cloud top height in Experiment 1 (Figure 7) is clearly better than the control during the analysis, with clearer patterns of highs and lows, which correlates quite well with the retrievals in Figure 3. On the other hand, the improvement during the forecast phase is less clear from a visual inspection of the mass load pattern (comparing Figure 6d to Figure 4d). This is also the case for the cloud top height pattern (comparing Figure 7d to Figure 5d). However, it does appear that the control run predicts a broad region of higher cloud tops further to the north than the results found in the retrievals. For example, the region of high clouds extends to nearly 8° S between 122° E and 124° E in Figure 5d. In Experiment 1, however, this region does not extend as far to the north—it is a more confined pattern moving to the southeast—and therefore, is in better agreement with the retrievals. One difficulty with this comparison is that as the cloud becomes thinner through dispersal, it probably falls beyond the detection threshold of the VOLCAT algorithm, and therefore, a significant portion of the ash cloud is not captured. This might explain why the simulated pattern extends so significantly beyond the retrieved cloud (in the southeast direction). The presence of meteorological clouds is also a possible contributor to this detection failure.

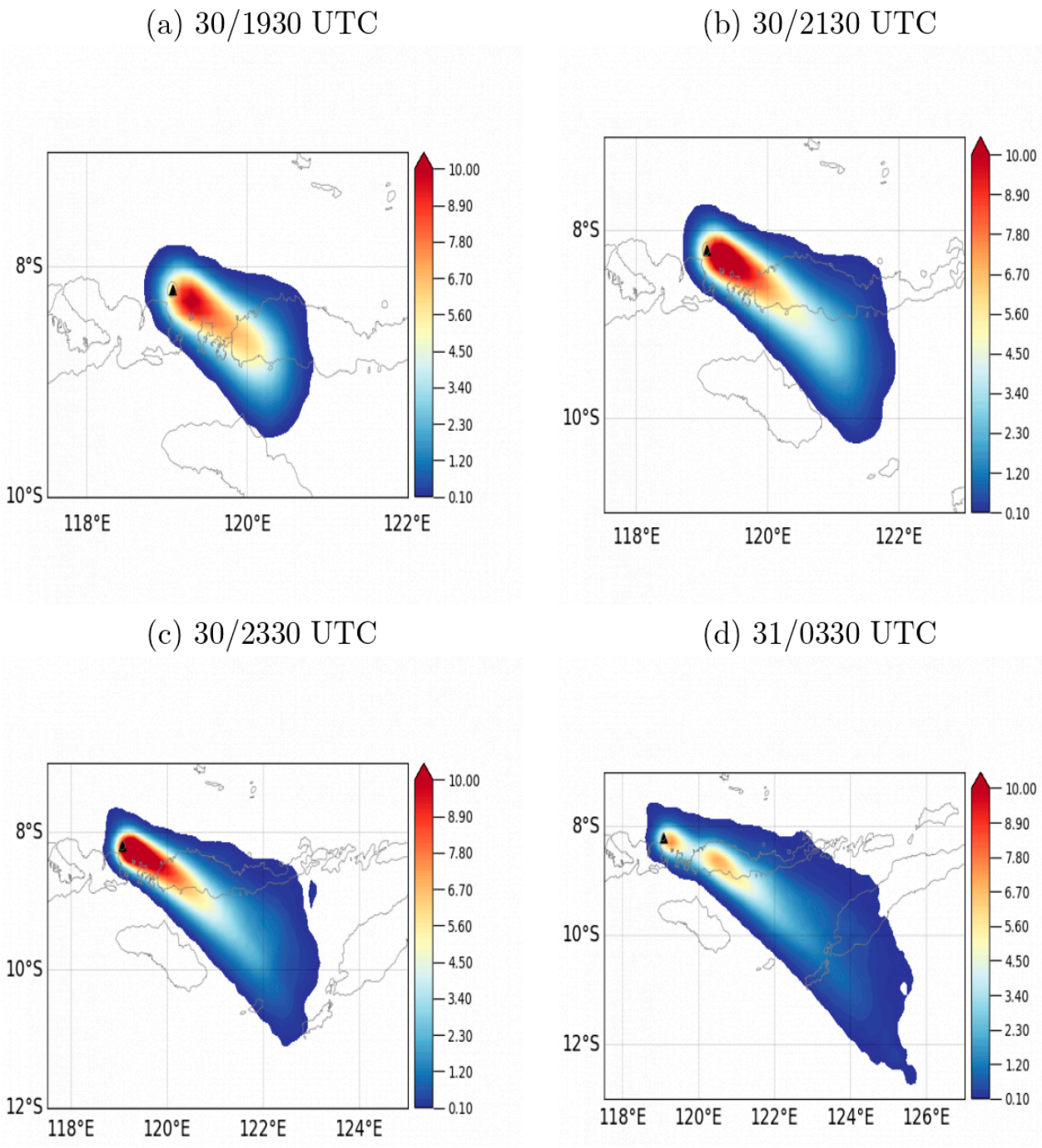

Figure 4. Ensemble mean mass load fields (in g/m^2) obtained from the control experiment in this study utilizing a simple top hat function for the source time variation. The black triangle indicates the volcano location.

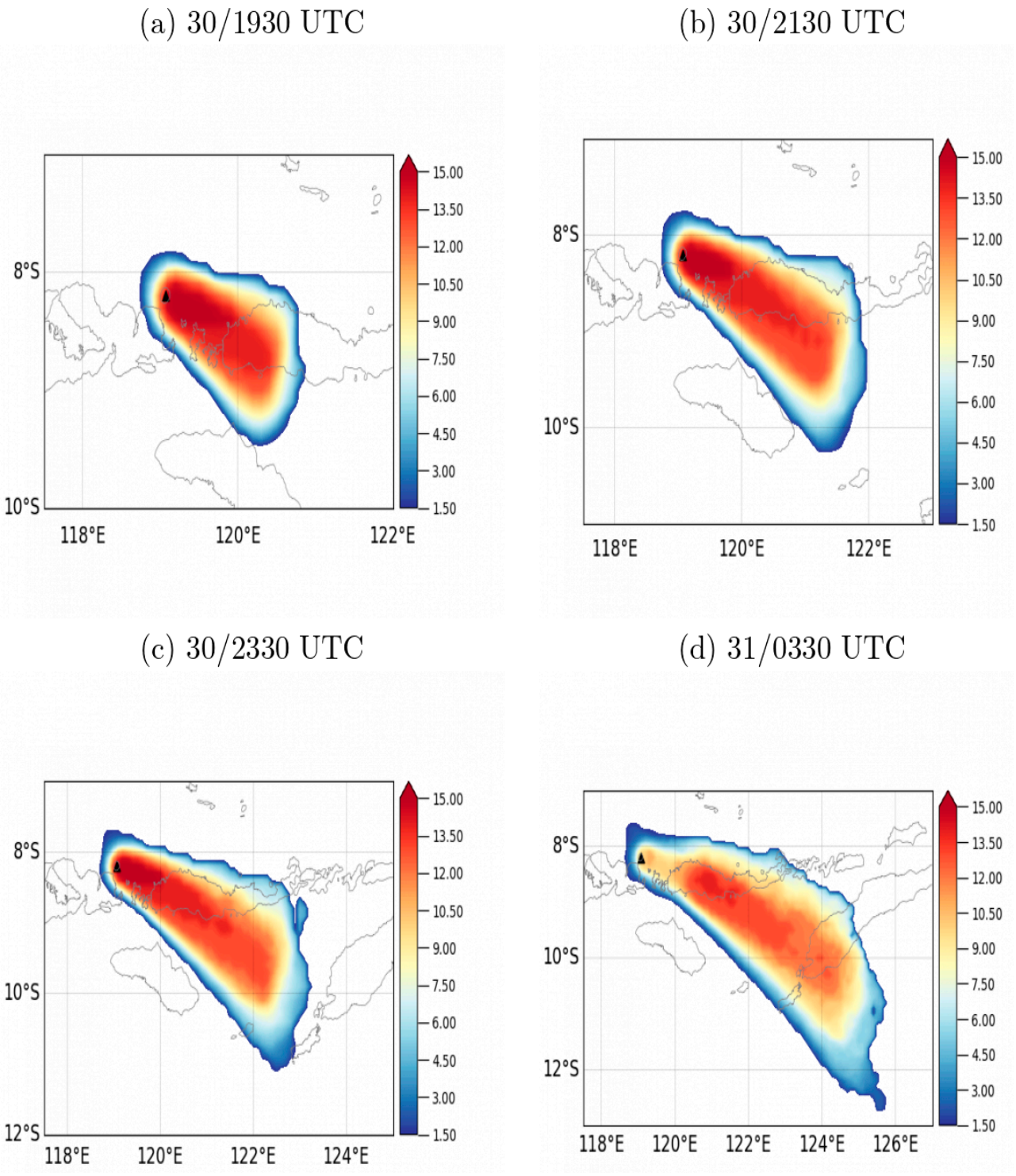

Figure 5. Ensemble mean cloud top height fields (in km) obtained from the control experiment in this study utilizing a simple top hat function for the source time variation. The black triangle indicates the volcano location.

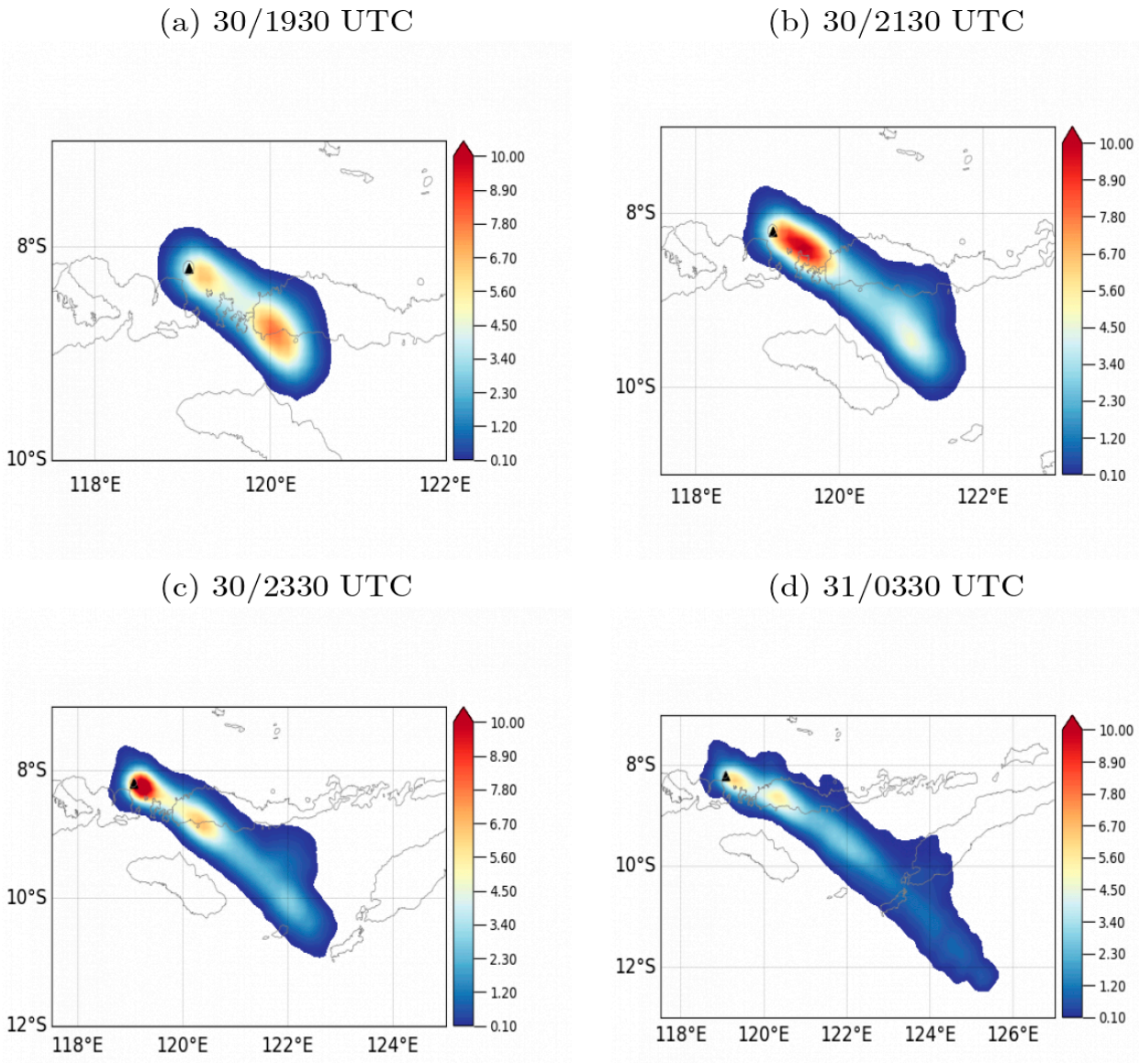

Figure 6. Ensemble mean mass load fields (in g/m^2) obtained from an experiment employing the time varying prior and posterior source height functions shown in Figure 1c,d. The black triangle indicates the volcano location.

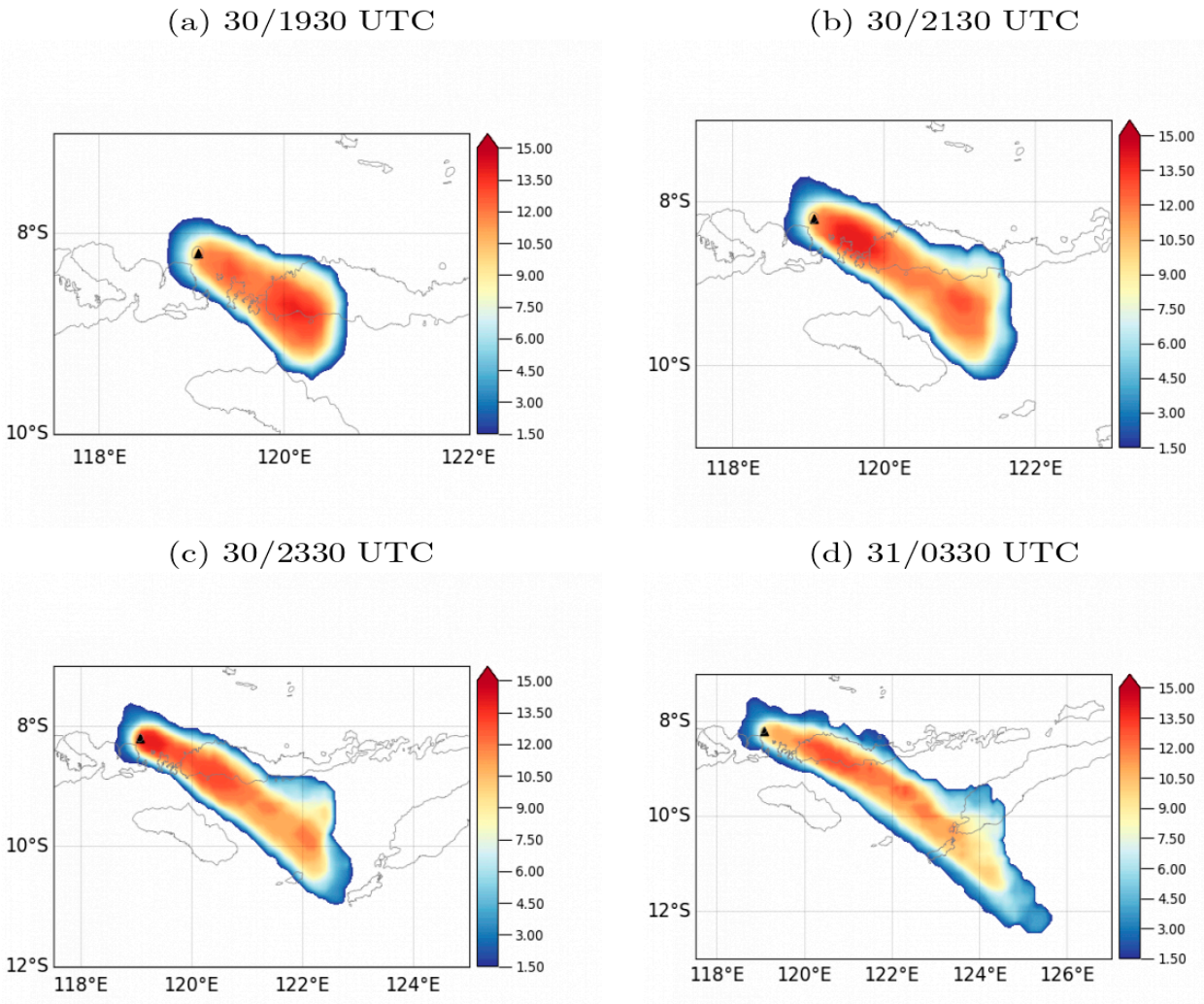

Figure 7. Ensemble mean cloud top height fields (in km) obtained from an experiment employing the time varying prior and posterior source height functions shown in Figure 1c,d. The black triangle indicates the volcano location.

4.3. Verification Study

As discussed in Section 3.3, various metrics were employed to evaluate the simulated mass load, effective radius, cloud top height, and detection fields with reference to VOLCAT satellite retrievals over the course of the simulation. The RMSE score was used to evaluate the performance of the ensemble mean fields (i.e., the forecasting skill) and the RPS was used to evaluate the ensemble field exceedance probabilities (which is a measure of the forecasting spread). Finally, verification scores for different fields were combined by computing relative scores (i.e., skill scores) with reference to the control experiment and averaging over different fields as indicated by Equation (17).

Figure 8 shows the ensemble mean RMSE scores for the control run and the 4 experiments numbered 1–4. These experiments employ different values of the correlation cut-off parameter β, as listed in Table 2. Lower RMSE values indicate smaller errors and therefore, better forecasts. The experimental RMSE values are in most cases, smaller than the corresponding control run values for all fields and for most times, which indicates superior forecasting skill. There is some variability in the RMSE scores between Experiments 1–4; however, all experiments show some degree of improvement in the mass load and cloud top height fields. However, Experiment 4 (with the highest β value) is, at best, neutral in the effective radius and detection fields. Except for the mass load field, in which it is difficult to discern a clear pattern trend, the RMSE values for all fields seem to indicate that high values

of β lead to somewhat poorer RMSE values. The effective radius field RMSE shows the most variability in scores. The mass load and height fields seem to have the most consistent improvement, as better scores relative to the control run are demonstrated irrespective of the β value. The improvements in the detection fields are small in comparison.

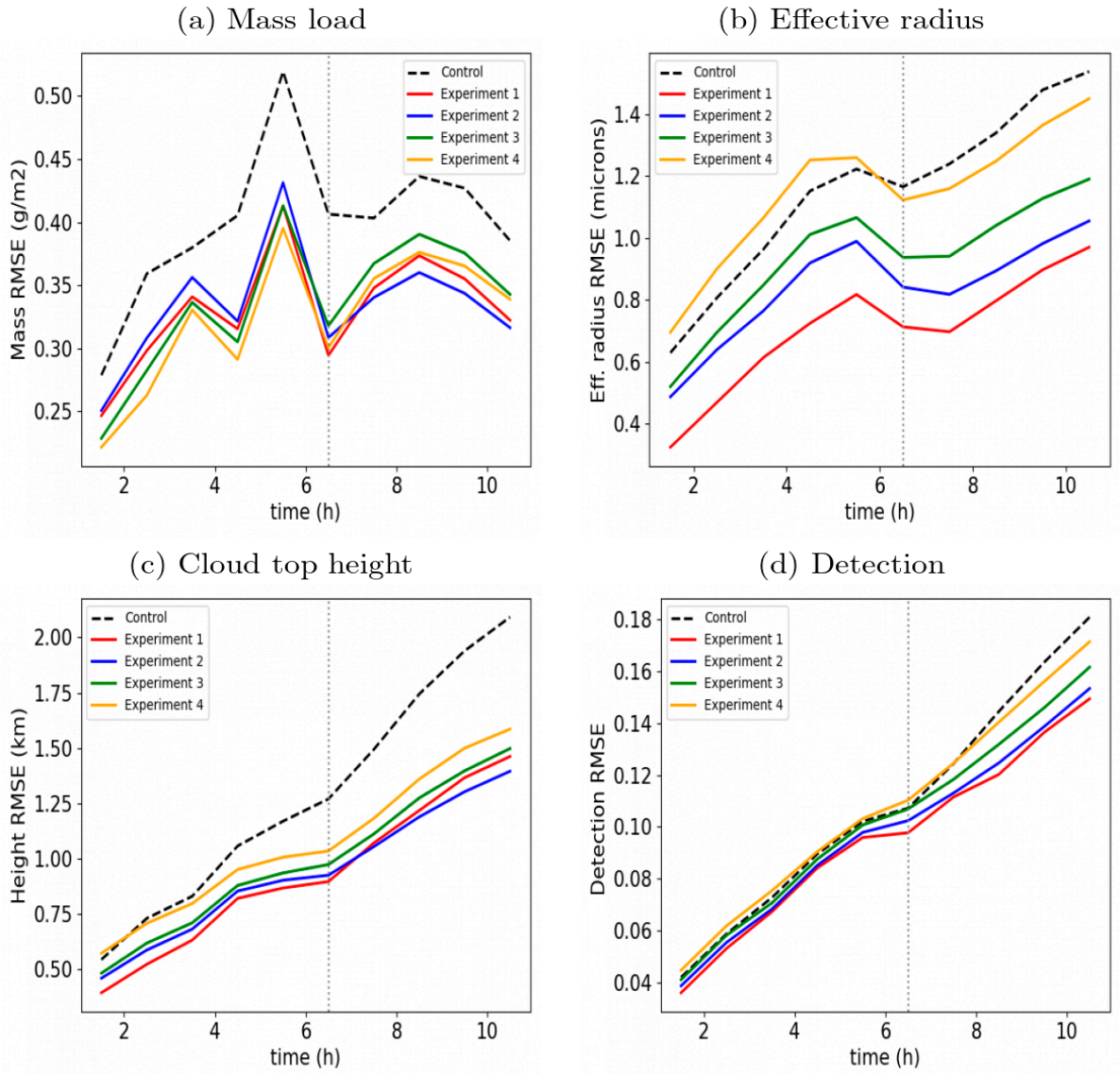

Figure 8. Ensemble mean root mean square error (RMSE) scores for different fields as a function of time for the control and different experiments in this study. The experiments utilize different values of β (see Table 2). Lower RMSE values indicate better ensemble mean skill. The dotted vertical line demarcates the analysis and forecast periods.

Figure 9 shows the RPS, with smaller RPS values indicating better probabilistic forecasting skill. All four fields in most of the experiments demonstrate some improvement over the control run, with the cloud top height field showing the most improvement. Interestingly, the improvement is rather modest in the analysis phase, but more pronounced in the forecast phase, in most cases. Although the trend is not as clear as in the RMSE scores, it does appear that experiments with low β values (Experiments 1 and 2) have some advantages over experiments with higher β values (Experiments 3 and 4), especially in the forecasting phase.

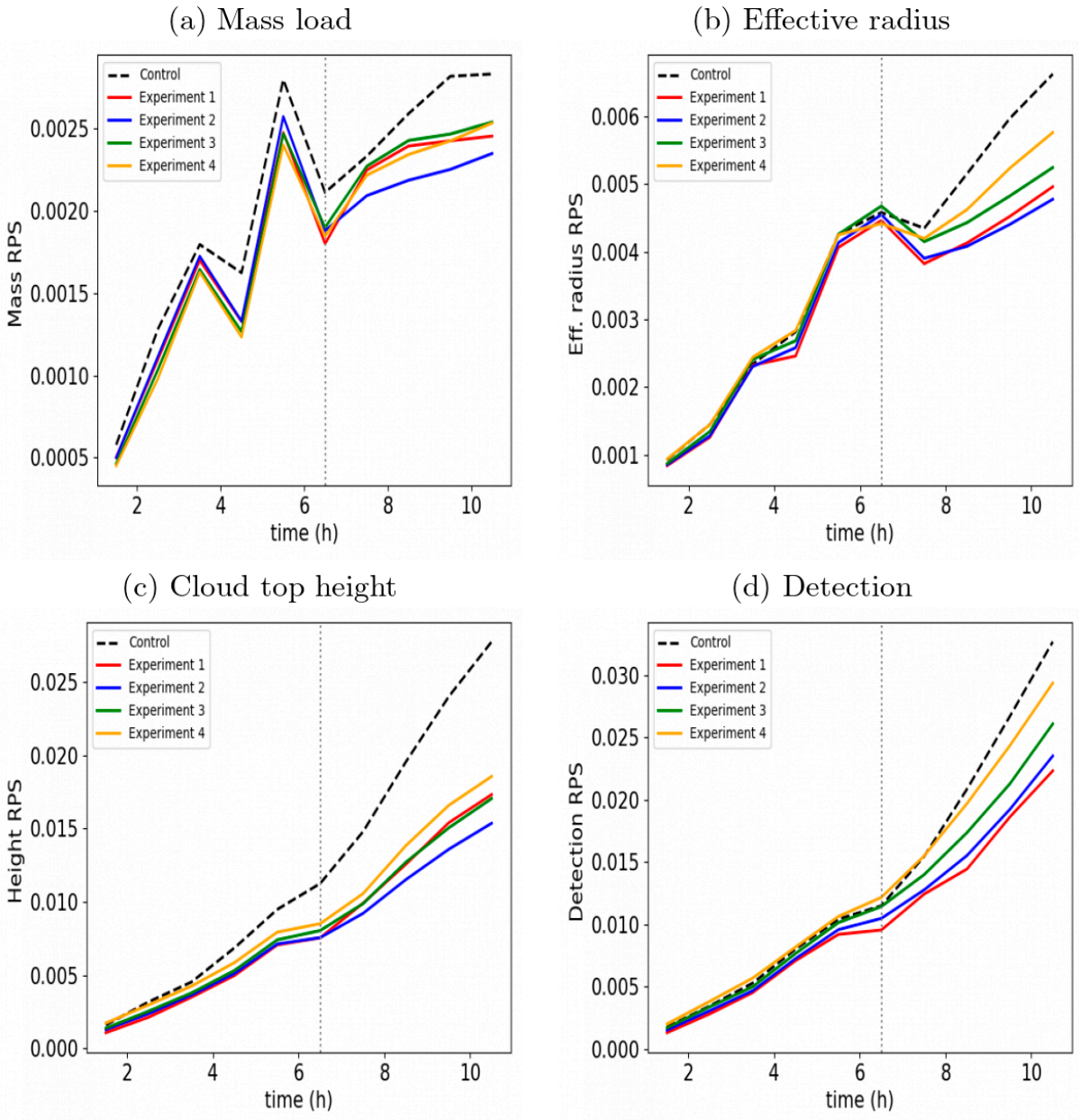

Figure 9. Ranked probability scores (RPS) for different fields as a function of time for the control run and experiments in this study. The experiments utilize different values of β (see Table 2). Lower RPS values indicate better probabilistic skill. The dotted vertical line demarcates the analysis and forecast periods.

Figure 10 shows the RMSESS and RPSS values relative to the control forecast. It should be noted that unlike the corresponding scores (in Figures 8 and 9), higher skill scores indicate better forecasting skill, and negative values indicate forecasting degradation relative to the control. The skills scores in Figure 10 are an average of the skill scores for the mass load, effective radius, cloud top height, and detection fields. The results indicate that the ensemble mean RMSE is improved by 10–25% and the RPS by about 15–30% on average during the forecasting phase. Hence, it can be concluded that the new time-varying formulation of Section 2.2 has been demonstrated to improve ensemble forecasts, as far as the case study considered in this paper is concerned. Experiments 1 and 2 (with lower β values) show the most improvement in skill scores. If the lower β values are considered to be the most appropriate, then RMSE improvements are just over 20%, and RPSS improvements are just over 25% for this case study.

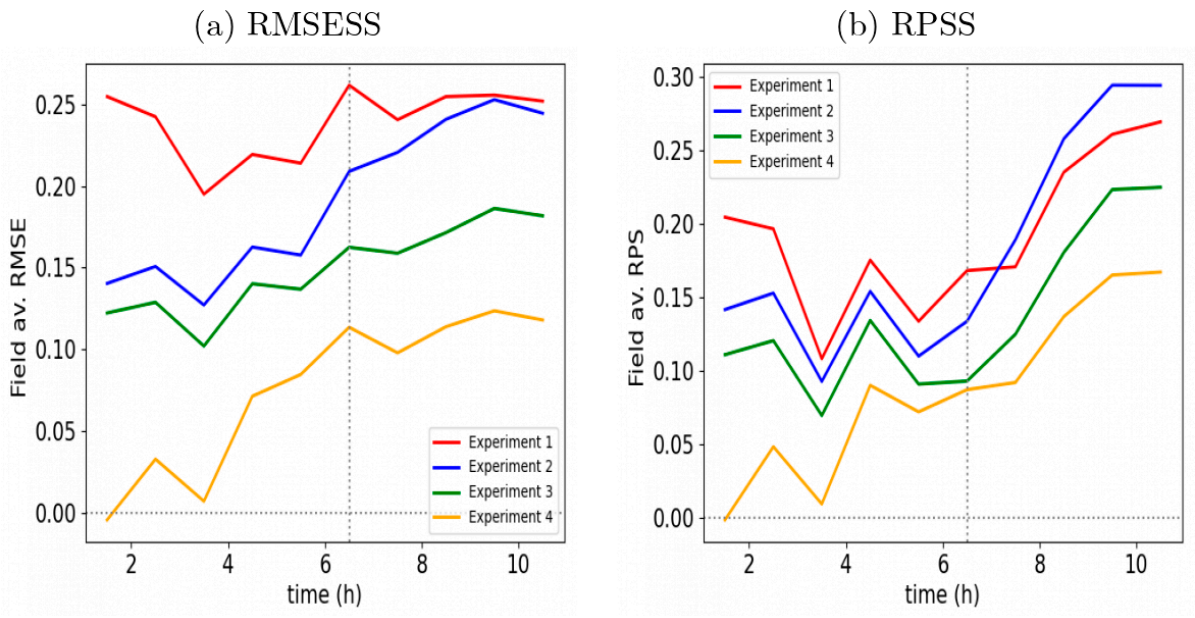

Figure 10. RMSESS and RPSS values (see Section 3.3) as a function of time for the experiments in this study. The experiments utilize different values of β (see Table 2). Positive skill scores indicate improved skill relative to the control experiment. The dotted vertical line demarcates the analysis and forecast periods.

5. Discussion

In this paper, a new scheme to account for time variation in the volcanic ash source term was introduced. In this scheme, temporal variations in source emission rates occur due to variations in the source height. The temporal source height function is obtained by random sampling of an a priori specified uncertainty range at different times within an analysis time window. The random sampling is performed for all ensemble members in the forecasting system. An ensemble filtering scheme, whose development was described in previous papers (see Section 1), is then used to create a new 'posterior' ensemble in which members that do not match well with observations (represented by VOLCAT satellite retrievals) are removed from the forecasting phase of the algorithm. This process was shown to lead to a forecasting system that better matches the observations during the analysis time window. In contrast, the original methodology considered in previous studies simply represents the temporal variation of emissions as a top hat function; emissions are constant over the duration of the eruption. This was shown to result in an unrealistic distribution of mass over the analysis time window in an eruption with strong temporal fluctuations.

The algorithm was applied to the 30 May 2014 Sangeang Api eruption case study. This eruption was of interest because it involved at least three distinct explosive episodes spaced about three hours apart. This kind of activity is difficult to model with the simple source formulation used in previous studies, in which the temporal variation is a simple top hat function, but it is in principle possible to model with the new formulation presented in this paper. Hence, this case study was an ideal testbed for the new algorithm. The results obtained with the new formulation were very encouraging. The posterior source heights displayed the expected temporal pattern with peak heights spaced 3 h apart, consistent with the observed eruption characteristics. The ensemble mean mass load patterns were also in very good agreement with satellite retrievals during the analysis time window. This contrasts with the control run, in which the original formulation, with a top hat function in the temporal dimension, was used. The mass load pattern in the control run did not display the multimodal structure seen in the retrievals due to the lack of appropriate temporal variability in the source term.

In this case study, a quantitative verification analysis also demonstrated the superiority of this scheme over the original scheme. It was shown that the RMSE scores of the ensemble

mean fields relative to observations, which measure the skill of the ensemble, were significantly improved compared to the control forecasts. A similar degree of improvement was also demonstrated for the RPS metric, which measures the reliability of the probabilistic forecasts. Hence, it was demonstrated for the 30 May 2014 Sangeang Api case study that the new source variation estimation scheme improves the ensemble forecasts. Experiments employing different degrees of correlation between random temporal height perturbations and the prior height function suggested that it might be advantageous not to constrain the temporal pattern of the random perturbation too tightly. With low correlation between random perturbations and the prior function, improvements to RMSE and RPS of over 20% were demonstrated. However, a difficulty with the verification in this study is that the detection capability of the VOLCAT system typically deteriorates within a few hours of tracking a distal cloud, making it difficult to verify forecasts beyond a few hours. That is the principal reason for the relatively short duration of 4 h that was used for verification. Hence, more case studies with retrievals available over longer periods are needed to provide a better understand of the utility of this method over longer timescales.

6. Conclusions

It was shown in this study that a new scheme for representing source time variation in a volcanic ash ensemble forecasting system that utilizes an ensemble filtering scheme can be successfully applied to a case study involving strong temporal variability in eruptive activity, namely the 30 May Sangeang Api eruption event. The resulting posterior source height function was shown to have temporal characteristics which were in very good agreement with observed eruption variability. Mass loads and cloud top height patterns and verification scores obtained with the system were shown to be superior to corresponding fields and scores obtained with a system used in previous studies, in which the time variability was represented by a simple top hat function.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: VOLCAT satellite retrieval data used in this study are available at http://doi.org/10.5281/zenodo.3579613. ACCESS-GE NWP fields in HYSPLIT ARL format are available at http://dx.doi.org/10.25914/6142ad42d977e.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Casadevall, T.J. *Volcanic Ash and Aviation Safety: Proceedings of the First International Symposium on Volcanic Ash and Aviation Safety*; DIANE Publishing: Collingdale, PA, USA, 1995; Volume 2047.
2. O'Kane, T.J.; Naughton, M.J.; Xiao, Y. The Australian community climate and earth system simulator global and regional ensemble prediction scheme. *ANZIAM J.* **2008**, *50*, 385–398. [CrossRef]
3. Bessho, K.; Date, K.; Hayashi, M.; Ikeda, A.; Imai, T.; Inoue, H.; Yoshida, R. An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites. *J. Meteorol. Soc. Jpn.* **2016**, *94*, 151–183. [CrossRef]
4. Corradini, S.; Merucci, L.; Prata, A.J.; Piscini, A. Volcanic ash and SO2 in the 2008 Kasatochi eruption: Retrievals comparison from different IR satellite sensors. *J. Geophys. Res.* **2010**, *115*. [CrossRef]
5. Francis, P.N.; Cooke, M.C.; Saunders, R.W. Retrieval of physical properties of volcanic ash using Meteosat: A case study from the 2010 Eyjafjallajökull eruption. *J. Geophys. Res.* **2012**, *117*. [CrossRef]
6. Pavolonis, M.J.; Heidinger, A.K.; Sieglaff, J. Automated retrievals of volcanic ash and dust cloud properties from upwelling infrared measurements. *J. Geophys. Res. Atmos.* **2013**, *118*, 1436–1458. [CrossRef]
7. Pavolonis, M.J.; Sieglaff, J.; Cintineo, J. Spectrally Enhanced Cloud Objects—A generalized framework for automated detection of volcanic ash and dust clouds using passive satellite measurements: 1. Multispectral analysis. *J. Geophys. Res. Atmos.* **2015**, *120*, 7813–7841. [CrossRef]
8. Pavolonis, M.J.; Sieglaff, J.; Cintineo, J. Spectrally Enhanced Cloud Objects—A generalized framework for automated detection of volcanic ash and dust clouds using passive satellite measurements: 2. Cloud object analysis and global application. *J. Geophys. Res. Atmos.* **2015**, *120*, 7842–7870. [CrossRef]

9. Eckhardt, S.; Prata, A.J.; Seibert, P.; Stebel, K.; Stohl, A. Estimation of the vertical profile of sulfur dioxide injection into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modeling. *Atmos. Chem. Phys.* **2008**, *8*, 3881–3897. [CrossRef]
10. Stohl, A.; Prata, A.J.; Eckhardt, S.; Clarisse, L.; Durant, A.; Henne, S. Determination of time-and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: The 2010 Eyjafjallajökull eruption. *Atmos. Chem. Phys.* **2011**, *11*, 4333–4351. [CrossRef]
11. Kristiansen, N.I.; Prata, A.J.; Stohl, A.; Carn, S.A. Stratospheric volcanic ash emissions from the 13 February 2014 Kelut eruption. *Geophys. Res. Lett.* **2015**, *42*, 588–596. [CrossRef]
12. Chai, T.; Crawford, A.; Stunder, B.; Pavolonis, M.J.; Draxler, R.; Stein, A. Improving volcanic ash predictions with the HYSPLIT dispersion model by assimilating MODIS satellite retrievals. *Atmos. Chem. Phys.* **2017**, *17*, 2865–2879. [CrossRef]
13. Harvey, N.J.; Dacre, H.F.; Webster, H.N.; Taylor, I.A.; Khanal, S.; Grainger, R.G.; Cooke, M.C. He impact of ensemble meteorology on inverse modeling estimates of volcano emissions and ash dispersion forecasts: Grímsvötn 2011. *Atmosphere* **2020**, *11*, 1022. [CrossRef]
14. Pelley, R.E.; Thomson, D.J.; Webster, H.N.; Cooke, M.C.; Manning, A.J.; Witham, C.S.; Hort, M.C. A Near-Real-Time Method for Estimating Volcanic Ash Emissions Using Satellite Retrievals. *Atmosphere* **2021**, *12*, 1573. [CrossRef]
15. Mastin, L.G.; Guffanti, M.; Servranckx, R.; Webley, P.; Barsotti, S.; Dean, K. A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions. *J. Volcanol. Geotherm. Res.* **2009**, *186*, 10–21. [CrossRef]
16. Zidikheri, M.J.; Lucas, C. Using satellite data to determine empirical relationships between volcanic ash source parameters. *Atmosphere* **2020**, *11*, 342. [CrossRef]
17. Zidikheri, M.J.; Potts, R.J. A simple inversion method for determining optimal dispersion model parameters from satellite detections of volcanic sulfur dioxide. *J. Geophys. Res. Atmos.* **2015**, *120*, 9702–9717. [CrossRef]
18. Zidikheri, M.J.; Potts, R.; Lucas, C. A probabilistic inverse method for volcanic ash dispersion modelling. *ANZIAM J.* **2014**, *56*, 194–209. [CrossRef]
19. Zidikheri, M.J.; Lucas, C.; Potts, R. Estimation of optimal dispersion model source parameters using satellite detections of volcanic ash. *J. Geophys. Res. Atmos.* **2017**, *122*, 8207–8232. [CrossRef]
20. Zidikheri, M.J.; Lucas, C.; Potts, R.J. Toward quantitative forecasts of volcanic ash dispersal: Using satellite retrievals for optimal estimation of source terms. *J. Geophys. Res. Atmos.* **2017**, *122*, 8187–8206. [CrossRef]
21. Zidikheri, M.J.; Lucas, C.; Potts, R.J. Quantitative verification and calibration of volcanic ash ensemble forecasts using satellite data. *J. Geophys. Res. Atmos.* **2018**, *123*, 4135–4156. [CrossRef]
22. Zidikheri, M.J.; Lucas, C. A computationally efficient ensemble filtering scheme for quantitative volcanic ash forecasts. *J. Geophys. Res. Atmos.* **2021**, *126*. [CrossRef]
23. Zidikheri, M.J.; Lucas, C. Improving Ensemble Volcanic Ash Forecasts by Direct Insertion of Satellite Data and Ensemble Filtering. *Atmosphere* **2021**, *12*, 1215. [CrossRef]
24. Draxler, R.R.; Hess, G.D. An overview of the HYSPLIT_4 modelling system for trajectories. *Aust. Meteorol. Mag.* **1998**, *47*, 295–308.
25. Pavolonis, M.J.; Calvert, C.C.; Cintineo, J.; Sieglaff, J. Transforming satellite data to products in the era of big data. In *Adding Value: Applications of Weather and Climate Services-Abstracts of the Eleventh CAWCR Workshop 27 November-1 December 2017, Melbourne, Australia*; 2017. Available online: http://www.bom.gov.au/research/publications/cawcrreports/CTR_081.pdf (accessed on 3 August 2022).
26. Ohkawara, N. *Multifunctional Transport Satellite (MTSAT)*; Meteorological Satellite Center, Japan Meteorological Agency: Tokyo, Japan, 2003; pp. 1–8.