Enhancing Ensemble Data Assimilation into One-Way-Coupled Models with One-Step-Ahead-Smoothing

Enhancing Ensemble Data Assimilation into One-Way-Coupled Models with One-Step-Ahead-Smoothing

Naila F. Raboudi[a] Boujemaa Ait-El-Fquih[a] Aneesh C. Subramanian[b] and Ibrahim Hoteit[a] *

[a] *King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia*

[b] *University of Colorado Boulder, Boulder, USA*

*Correspondence to: Ibrahim Hoteit, Division of Physical Science and Engineering, KAUST, Thuwal, Jeddah, 23955-6900, Kingdom of Saudi Arabia. E-mail: ibrahim.hoteit@kaust.edu.sa

This study investigates the filtering problem with one-way coupled (OWC) state-space systems, for which the joint ensemble Kalman filter (EnKF) is the standard solution. In this approach, the states of the two coupled sub-systems are jointly updated with all incoming observations. This enables transferring the information across the sub-systems, which should provide coupled-state estimates in better agreement with the observations. The state estimates of the joint EnKF highly depend on the relevance of the joint ensembles' cross-covariances between the sub-systems' variables. In this work, we propose a new joint EnKF scheme based on the One-Step-Ahead (OSA) smoothing formulation of the filtering problem for efficient assimilation into OWC systems. The scheme introduces an extra smoothing step for both states sub-systems with the future observations, followed by an analysis step for each sub-system state using only its own observation, all within a Bayesian consistent framework. The extra OSA-smoothing step enables to more efficiently exploit the observations, to enhance the representativeness of the EnKF covariances, and to mitigate for reported inconsistencies in the joint EnKF analysis step. We demonstrate the relevance of the proposed approach by presenting and analyzing results of various numerical experiments conducted with a OWC Lorenz-96 model.

Key Words: ensemble Kalman filter (EnKF), One-Step-Ahead (OSA) smoothing, one-way coupling (OWC).

Received . . .

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.3916

1. Introduction

Coupled earth models are designed to simulate the interactions between several dynamical processes (Zhang *et al.* 2007). These systems may couple two or more components, including the atmosphere, ocean, land surface, chemistry or hydrology. Accounting for the interactions between the different involved processes should help improve the predictability of such systems (Zhang *et al.* 2007; Luo and Hoteit 2014; Županski 2017). Here, we are interested in a specific class of coupled models, the one-way coupled (OWC) systems, consisting of two sub-systems where only one of them is forced with the state of the other (Benra *et al.* 2011). OWC systems are becoming increasingly used in many fields, as for instance for modeling of various transport phenomena, marine ecosystem modeling, etc (Travers *et al.* 2007).

Compared to classical single component models, coupled models are more complex as they involve multi-spatial and multi-temporal scales that interact nonlinearly. This makes them substantially more sensitive to the data assimilation (DA) strategies (Sakaguchi *et al.* 2012; Gharamti *et al.* 2014). In particular, the exchange of information across the system sub-components needs to be carefully treated during the assimilation process, referred to as coupled data assimilation (CDA) (e.g. Saha *et al.* 2010; Penny and Hamill 2017; Hoteit *et al.* 2018). Depending on the level of the exchange of information between the different components during assimilation, CDA can be broadly categorized into two types: weak CDA (WCDA) and strong CDA (SCDA) (Liu *et al.* 2013; Han *et al.* 2013). WCDA allows interactions between the sub-components only through the coupling dynamics, while each sub-component assimilating its own observations. This approach offers an important practical advantage of being promptly ready to apply when the sub-components are already equipped with their assimilation modules, at the cost of not exploiting the cross-information from the other components observations. This could also lead to scale mismatches and imbalances in the assimilated solutions (Han *et al.* 2013). In contrast, SCDA respects the full coupling between the components through both the coupling dynamics and during assimilation. The effect of any observation from one model component is indeed exploited in the estimation of the others through their joint distribution (Luo and Hoteit 2014). This should produce more physically balanced coupled state estimates and make the system in better physical agreement with all available observations (Liu *et al.* 2013; Hoteit *et al.* 2018). Algorithmically, SCDA can be implemented by concatenating the states of the sub-systems into one augmented state, on which conventional DA methods can be directly applied. While WCDA has been successfully implemented in several studies, mainly for weather prediction (e.g. Zhang *et al.* 2007; Saha *et al.* 2010; Lahoz and Schneider 2014; Laloyaux *et al.* 2016; Browne *et al.* 2019), SCDA is still not mature and remains an active area of research that requires more investigations to efficiently handle the exchange of information between the system components (Luo and Hoteit 2014; Penny and Hamill 2017; Hoteit *et al.* 2018).

The ensemble Kalman filter (EnKF) is a popular assimilation algorithm that was introduced by Evensen (1994) as a Monte Carlo implementation of the Kalman filter (KF) (Kalman 1960) to tackle large-scale, sequential, nonlinear DA problems. Using a set of realizations of model states, called ensemble, the filter estimates the first two moments, i.e., mean and covariance that can be computed exactly by the KF, in two steps (Evensen 2003; Hoteit *et al.* 2018): a forecast step that integrates the ensemble members forward with the model, and an analysis step to update the members with the incoming observations.

The implementation of SCDA within an EnKF framework still faces important challenges in real world applications, the most important of which being the involved multi-spatial/temporal scales and strongly nonlinear character of the coupled system components, and of course the increased computational cost (Han *et al.* 2013; Luo and Hoteit 2014). This makes it challenging to estimate reliable cross-correlation terms for the forecast error covariance matrix, which characterizes the exchange of information between the different components in the analysis step (Ballabrera-Poy *et al.* 2008). Strategies for dealing with these challenges have been investigated in several studies. For example, Lu *et al.* (2015) suggested using lagged covariances to tackle the difficulty of accounting for the different time scales, Luo and Hoteit (2014) proposed a divided state-space estimation strategy that accounts for the effect of coupling between the sub-systems and Frolov *et al.* (2016) followed a similar approach to build an

interface solver that retains independent DA sub-systems, while allowing for coupling during assimilation.

Considering the OWC filtering problem as a generalization of the state-parameter estimation problem, where the parameter vector evolves according to a given dynamical model that can be also observed, this problem can also be subject to issues related to stability and tractability when jointly updating the sub-systems states (Moradkhani *et al.* 2005; Hendricks Franssen and Kinzelbach 2008; Lü *et al.* 2011; Wen *et al.* 2005b). For instance, with highly nonlinear models, this produces non-Gaussian distributions and may result in some ensemble members being inconsistent with the model's dynamics (Wen *et al.* 2005b; Moradkhani *et al.* 2005; Lü *et al.* 2011; Subramanian *et al.* 2012). Other challenges were also raised in the literature among which we mainly cite: (i) the large degrees of freedom of the augmented state vector, potentially leading to intractable solutions, (ii) the possible time-updating inconsistencies, and (iii) potentially the difficulty of the joint schemes to handle different degrees of localization and inflation (Luo and Hoteit 2014; Wen *et al.* 2005b,a).

To enhance the EnKFs performances with OWC systems, we resort to the one-step-ahead (OSA) smoothing formulation of the filtering problem. Filtering with OSA-smoothing relies on the idea that the standard "forecast-then-update" path of the Bayesian filtering process is not unique, and can be reversed (Desbouvries and Ait-El-Fquih 2008; Desbouvries *et al.* 2011). This was shown to introduce an additional update step based on the future observation (OSA-smoothing), within a fully consistent Bayesian framework. In linear Gaussian systems, both standard and OSA-smoothing Kalman filtering approaches provide the same estimates (Desbouvries and Ait-El-Fquih 2008). In more general nonlinear and non-Gaussian systems, the OSA-smoothing scheme was shown to provide enhanced state and parameters estimates with the EnKFs (Gharamti *et al.* 2015; Ait-El-Fquih *et al.* 2016; Raboudi *et al.* 2018). The additional smoothing step indeed constrains the sampling of the forecast ensemble with the future observation, which provides improved background statistics for the analysis step (Raboudi *et al.* 2018).

EnKF with OSA-smoothing provided a Bayesian consistent framework for state-parameters estimation with the dual EnKF of Moradkhani *et al.* (2005) (Ait-El-Fquih *et al.* 2016) and for the running-in-place (RIP) scheme of Kalnay and Yang (2010) (Raboudi *et al.* 2018). El Gharamti *et al.* (2013) followed the dual state-parameter estimation method of Moradkhani *et al.* (2005) to derive a dual EnKF for DA into OWC models, demonstrating the relevance of using a time consistent sequential updating scheme to mitigate for the imbalances issues of the joint EnKF. Here, we generalize this work and derive a fully Bayesian consistent EnKF algorithm with OSA-smoothing for OWC assimilation. This introduces an extra smoothing step based on the future observations of both sub-systems to update their states, whereas the analysis steps are independently applied to each state component using only its own observations, resulting in the strong-type EnKF$_{OSA}$ scheme (SC-EnKF$_{OSA}$ hereafter). In its weak version (WC-EnKF$_{OSA}$), the smoothing step is performed on each state component using only its own observation. Extensive numerical experiments with a OWC multiscale Lorenz-96 model are conducted under different assimilation settings and scenarios to investigate the behavior of the proposed schemes. Our results suggest that SC-EnKF$_{OSA}$ is clearly more robust than SC-EnKF and WC-EnKF for estimating the coupled state components, mainly in the challenging scenarios of small ensembles, limited observations, assimilation of observations from one model component only, and biased forecast model. We further find that WC-EnKF$_{OSA}$ outperforms SC-EnKF$_{OSA}$ with small ensembles.

The remainder of the paper is organized as follows. Section 2 recalls the strong EnKF (SC-EnKF) and weak (WC-EnKF) algorithms. The proposed SC-EnKF$_{OSA}$ and WC-EnKF$_{OSA}$ schemes are then derived in Section 3. Results of numerical experiments with the OWC multi-scale Lorenz-96 model are analyzed and discussed in Section 4. Concluding remarks are finally given in Section 5.

2. The classical EnKF for OWC systems

2.1. Problem formulation

Consider the discrete-time OWC dynamical system:

$$\begin{cases} \mathbf{x}_n &= \mathcal{M}^x_{n-1}\left(\mathbf{x}_{n-1}\right) + \eta^x_{n-1} \\ \mathbf{z}_n &= \mathcal{M}^z_{n-1}\left(\mathbf{z}_{n-1}, \mathbf{x}_{n-1}\right) + \eta^z_{n-1} \\ \mathbf{y}^x_n &= \mathbf{H}^x_n \mathbf{x}_n + \boldsymbol{\varepsilon}^x_n \\ \mathbf{y}^z_n &= \mathbf{H}^z_n \mathbf{z}_n + \boldsymbol{\varepsilon}^z_n \end{cases}, \tag{1}$$

where $\mathbf{x}_n \in \mathbb{R}^{N_x}$ is the "free" system state of dimension N_x and $\mathbf{z}_n \in \mathbb{R}^{N_z}$ is the N_z dimensional "forced" state (by the state of the other sub-system) at time instant t_n. $\mathbf{y}^x_n \in \mathbb{R}^{N_{y_x}}$ and $\mathbf{y}^z_n \in \mathbb{R}^{N_{y_z}}$ are respectively the N_{y_x} and N_{y_z} dimensional observations of $\mathbf{x}_n$ and $\mathbf{z}_n$ at t_n. $\mathcal{M}^x_{n-1}(.)$ and $\mathcal{M}^z_{n-1}(.)$ are the (nonlinear) transition operators that respectively integrate the system states, $\mathbf{x}$ and $\mathbf{z}$, from time instant t_{n-1} to t_n. The observation operators, $\mathbf{H}^x_n$ and $\mathbf{H}^z_n$, respectively project $\mathbf{x}_n$ and $\mathbf{z}_n$ from the state space onto the observation space. We assume that $\mathbf{H}^x_n$ and $\mathbf{H}^z_n$ are linear for simplicity, but the proposed ensemble schemes can be easily extended to the case of nonlinear observation operators as discussed for example in Liu *et al.* (2016). The model noise terms, $\eta^x = \{\eta^x_n\}_{n\in\mathbb{N}}$ and $\eta^z = \{\eta^z_n\}_{n\in\mathbb{N}}$, and the observation noise terms, $\boldsymbol{\varepsilon}^x = \{\boldsymbol{\varepsilon}^x_n\}_{n\in\mathbb{N}}$ and $\boldsymbol{\varepsilon}^z = \{\boldsymbol{\varepsilon}^z_n\}_{n\in\mathbb{N}}$, are assumed to be independent, jointly independent and independent of the initial states $\mathbf{x}_0$ and $\mathbf{z}_0$. η^x_n and $\boldsymbol{\varepsilon}^x_n$ (respectively, η^z_n and $\boldsymbol{\varepsilon}^z_n$) are also assumed Gaussian with zero means and covariances, $\mathbf{Q}^x_n$ and $\mathbf{R}^x_n$ (respectively $\mathbf{Q}^z_n$ and $\mathbf{R}^z_n$). The observation operators for the two sub-systems are "separable" in the sense that the observation corresponding to each sub-system depends only on its corresponding state. The cases where the observation of one sub-system depends on the state variables of both sub-systems might be also treated as described in (Luo and Hoteit 2014).

We address the filtering problem, namely, the estimation at any t_n, of the joint state $\mathbf{X}_n = \left[\mathbf{x}^T_n, \mathbf{z}^T_n\right]^T$ from the observations up to t_n (i.e., from $\mathbf{y}^x_{0:n} = \{\mathbf{y}^x_0, \mathbf{y}^x_1, \cdots, \mathbf{y}^x_n\}$ and $\mathbf{y}^z_{0:n} = \{\mathbf{y}^z_0, \mathbf{y}^z_1, \cdots, \mathbf{y}^z_n\}$). This is achieved by transforming the coupled system (1) into a standard state-space system operating on the augmented state $\mathbf{X}_n$ and observation $\mathbf{Y}_n = \left[(\mathbf{y}^x_n)^T, (\mathbf{y}^z_n)^T\right]^T$:

$$\begin{cases} \mathbf{X}_n &= \mathcal{M}_{n-1}\left(\mathbf{X}_{n-1}\right) + \eta_{n-1} \\ \mathbf{Y}_n &= \mathbf{H}_n \mathbf{X}_n + \boldsymbol{\varepsilon}_n \end{cases}, \tag{2}$$

where $\mathcal{M}_{n-1}\left(\mathbf{X}_{n-1}\right) = \begin{bmatrix} \mathcal{M}^x_{n-1}(\mathbf{x}_{n-1}) \\ \mathcal{M}^z_{n-1}(\mathbf{z}_{n-1}, \mathbf{x}_{n-1}) \end{bmatrix}$, $\mathbf{H}_n = \begin{bmatrix} \mathbf{H}^x_n & 0 \\ 0 & \mathbf{H}^z_n \end{bmatrix}$, and $\boldsymbol{\varepsilon}_n = \left[(\boldsymbol{\varepsilon}^x_n)^T,\ (\boldsymbol{\varepsilon}^z_n)^T\right]^T$ and $\eta_{n-1} = \left[(\eta^x_{n-1})^T,\ (\eta^z_{n-1})^T\right]^T$ respectively denote the augmented model and observations' noise terms, which are actually Gaussian with zero means and covariances $\mathbf{R}_n = \begin{bmatrix} \mathbf{R}^x_n & 0 \\ 0 & \mathbf{R}^z_n \end{bmatrix}$ and $\mathbf{Q}_{n-1} = \begin{bmatrix} \mathbf{Q}^x_{n-1} & 0 \\ 0 & \mathbf{Q}^z_{n-1} \end{bmatrix}$, respectively. The Kalman Filter (KF) recursively computes the posterior mean estimate when the system is linear and the noise terms are Gaussian (Hoteit *et al.* 2018). EnKFs are efficient Gaussian-Monte Carlo approximations of the filtering problem solution designed for large dimensional nonlinear systems at reasonable computational cost (Evensen 2003). In the following, we first recall the strong EnKF (SC-EnKF) and its weak variant (WC-EnKF). Their OSA-smoothing counterparts, SC-EnKF$_{OSA}$ and WC-EnKF$_{OSA}$, are then presented in the next section. We derive the schemes following a stochastic EnKF formulation (Evensen 2003), but a deterministic formulation (Tippett *et al.* 2003) could be also considered as derived by Raboudi *et al.* (2018).

2.2. The SC-EnKF and WC-EnKF algorithms

The SC-EnKF algorithm is derived by directly applying the standard EnKF to the augmented system (2). Starting from an analysis ensemble at t_{n-1}, $\{\mathbf{X}^{a,i}_{n-1}\}_{i=1}^{N_e}$, with $\mathbf{X}^{a,i}_{n-1} = \left[\left(\mathbf{x}^{a,i}_{n-1}\right)^T, \left(\mathbf{z}^{a,i}_{n-1}\right)^T\right]^T$ being the i^{th} joint analysis member, the forecast members, $\mathbf{X}^{f,i}_n = \left[\left(\mathbf{x}^{f,i}_n\right)^T, \left(\mathbf{z}^{f,i}_n\right)^T\right]^T$, at the time of the next available observation t_n are derived from the coupled model as (Luo and Hoteit 2014),

$$\mathbf{x}^{f,i}_n = \mathcal{M}^x_{n-1}\left(\mathbf{x}^{a,i}_{n-1}\right) + \eta^{x,i}_{n-1}, \tag{3}$$

$$\mathbf{z}^{f,i}_n = \mathcal{M}^z_{n-1}\left(\mathbf{z}^{a,i}_{n-1}, \mathbf{x}^{a,i}_{n-1}\right) + \eta^{z,i}_{n-1}, \tag{4}$$

where $\eta^{x,i}_{n-1}$ and $\eta^{z,i}_{n-1}$ are samples from $\mathcal{N}(\mathbf{0}, \mathbf{Q}^x_{n-1})$ and $\mathcal{N}(\mathbf{0}, \mathbf{Q}^z_{n-1})$, respectively. Let for any ensemble $\{\mu^i\}_{i=1}^{N_e}$, $\overline{\mu}_n$ and $\mathbf{P}_\mu$ being respectively the sample ensemble mean and covariance,

and $\mathbf{S}_\mu$ the corresponding perturbation matrix whose i^{th} column is defined as $\frac{1}{\sqrt{N_e-1}}(\mu^i - \overline{\mu})$, i.e., $\mathbf{P}_\mu = \mathbf{S}_\mu \mathbf{S}_\mu^T$. Once a new observation, $\mathbf{Y}_n$, is available, the i^{th} forecast member $\mathbf{X}_n^{f,i}$ is updated with a Kalman correction-based step to obtain the i^{th} analysis member $\mathbf{X}_n^{a,i}$ as,

$$\mathbf{X}_n^{a,i} = \mathbf{X}_n^{f,i} + \mathbf{K}_{\mathbf{X}_n^f, \mathbf{Y}_n^f}(\mathbf{Y}_n - \mathbf{Y}_n^{f,i}), \tag{5}$$

where
$$\mathbf{Y}_n^{f,i} = \mathbf{H}_n \mathbf{X}_n^{f,i} + \boldsymbol{\varepsilon}_n^i \tag{6}$$

represents a perturbed observation such that $\mathbf{Y}_n^f = \left[(\mathbf{y}_n^{f_x})^T, (\mathbf{y}_n^{f_z})^T\right]^T$, with $\boldsymbol{\varepsilon}_n^i$ sampled from $\mathcal{N}(\mathbf{0}, \mathbf{R}_n)$, and $\mathbf{K}_{\mathbf{X}_n^f, \mathbf{Y}_n^f}$ the Kalman gain estimated from the ensemble members as $\mathbf{K}_{\mathbf{X}_n^f, \mathbf{Y}_n^f} = \mathbf{P}_{\mathbf{X}_n^f, \mathbf{Y}_n^f} \mathbf{P}^{-1}_{\mathbf{Y}_n^f}$ (Evensen 2003),

with,
$$\mathbf{P}_{\mathbf{X}_n^f, \mathbf{Y}_n^f} = \mathbf{P}_{\mathbf{X}_n^f} \mathbf{H}_n^T = \mathbf{S}_{\mathbf{X}_n^f} \widetilde{\mathbf{H}}_n^T, \tag{7}$$

$$\mathbf{P}_{\mathbf{Y}_n^f} = \mathbf{H}_n \mathbf{P}_{\mathbf{X}_n^f} \mathbf{H}_n^T + \mathbf{R}_n = \widetilde{\mathbf{H}}_n \widetilde{\mathbf{H}}_n^T + \mathbf{R}_n, \tag{8}$$

and $\widetilde{\mathbf{H}}_n = \mathbf{H}_n \mathbf{S}_{\mathbf{X}_n^f}$. Thus, in an SC-EnKF, both free and forced states are updated based on both observations, $\mathbf{y}_n^x$ and $\mathbf{y}_n^z$. WC-EnKF simply updates each state with its own observations only by omitting the cross-diagonal blocks of the forecast covariance $\mathbf{P}_{\mathbf{X}_n^f}$ (i.e., $\mathbf{P}_{\mathbf{x}_n^f, \mathbf{z}_n^f} = 0$) as

$$\mathbf{x}_n^{a,i} = \mathbf{x}_n^{f,i} + \mathbf{K}_{\mathbf{x}_n^f, \mathbf{y}_n^{f_x}}(\mathbf{y}_n^x - \mathbf{y}_n^{f_x,i}), \tag{9}$$

$$\mathbf{z}_n^{a,i} = \mathbf{z}_n^{f,i} + \mathbf{K}_{\mathbf{z}_n^f, \mathbf{y}_n^{f_z}}(\mathbf{y}_n^z - \mathbf{y}_n^{f_z,i}), \tag{10}$$

with $\mathbf{K}_{\mathbf{x}_n^f, \mathbf{y}_n^{f_x}} = \mathbf{P}_{\mathbf{x}_n^f, \mathbf{y}_n^{f_x}} \mathbf{P}^{-1}_{\mathbf{y}_n^{f_x}}$ and $\mathbf{K}_{\mathbf{z}_n^f, \mathbf{y}_n^{f_z}} = \mathbf{P}_{\mathbf{z}_n^f, \mathbf{y}_n^{f_z}} \mathbf{P}^{-1}_{\mathbf{y}_n^{f_z}}$. In our Gaussian EnKF framework, the WC-EnKF is obtained under the independence assumption:

$$p(\mathbf{x}_n, \mathbf{z}_n | \mathbf{Y}_{0:n-1}) = p(\mathbf{x}_n | \mathbf{y}^x_{0:n-1}) p(\mathbf{z}_n | \mathbf{Y}_{0:n-1}). \tag{11}$$

3. The EnKF with OSA-smoothing for OWC systems

Ensemble OSA-smoothing filters involve two update steps with the same data, in a fully Bayesian consistent way and under the common Gaussian assumptions (Ait-El-Fquih *et al.* 2016; Raboudi *et al.* 2018). Starting from an analysis ensemble, $\{\mathbf{r}_{n-1}^{a,i}\}_{i=1}^{N_e}$, EnKF$_{OSA}$ first performs a standard forecast step to obtain a forecast ensemble, $\{\mathbf{r}_n^{f,i}\}_{i=1}^{N_e}$. This is then used to smooth $\{\mathbf{r}_{n-1}^{a,i}\}_{i=1}^{N_e}$ with the incoming observation, $\mathbf{y}_n^r$, leading to the smoothed ensemble, $\{\mathbf{r}_{n-1}^{s,i}\}_{i=1}^{N_e}$. A second model integration is subsequently performed, now starting from $\{\mathbf{r}_{n-1}^{s,i}\}_{i=1}^{N_e}$, to compute a pseudo-forecast ensemble, which is, in turn, updated using the same observation to obtain the analysis ensemble $\{\mathbf{r}_n^{a,i}\}_{i=1}^{N_e}$. Ait-El-Fquih *et al.* (2016) showed that the resulting forecast and analysis members are respectively samples of the desired forecast and analysis pdfs, under the assumption of $p(\mathbf{r}_{n-1}, \mathbf{r}_n, \mathbf{y}_n^r | \mathbf{y}_{0:n-1}^r)$ being Gaussian. Unlike iterative algorithms, assimilating the data twice does not affect the Bayesian character of the EnKF$_{OSA}$ schemes as the data is used to update two different quantities, $\mathbf{r}_{n-1}$ and $\mathbf{r}_n$.

Here, SC-EnKF$_{OSA}$ is derived by applying the EnKF$_{OSA}$ to the more general OWC system (2), by applying the equations in Raboudi *et al.* (2018, section 2.b) to the augmented state $\mathbf{X}_n$. The scheme therefore involves two joint update steps (smoothing and analysis) and two coupled model integrations (forecast and pseudo-forecast).

3.1. The smoothing step

Once the forecast ensemble $\{\mathbf{X}_n^{f,i} = \mathcal{M}_{n-1}\left(\mathbf{X}_{n-1}^{a,i}\right) + \eta_{n-1}^i\}_{i=1}^{N_e}$ generated according to (3) and (4), and the new observation $\mathbf{Y}_n$ is available, $\mathbf{X}_{n-1}^{a,i}$ is updated with the future observation using,

$$\mathbf{X}_{n-1}^{s,i} = \mathbf{X}_{n-1}^{a,i} + \mathbf{K}_{\mathbf{X}_{n-1}^a, \mathbf{Y}_n^f}(\mathbf{Y}_n - \mathbf{Y}_n^{f,i}), \tag{12}$$

with the Kalman gain $\mathbf{K}_{\mathbf{X}_{n-1}^a, \mathbf{Y}_n^f} = \mathbf{P}_{\mathbf{X}_{n-1}^a, \mathbf{Y}_n^f} \mathbf{P}^{-1}_{\mathbf{Y}_n^f}$, where $\mathbf{Y}_n^{f,i}$ and $\mathbf{P}_{\mathbf{Y}_n^f}$ are respectively given by (6) and (8), and the cross-covariance $\mathbf{P}_{\mathbf{X}_{n-1}^a, \mathbf{Y}_n^f}$ is computed from the corresponding ensembles as,

$$\mathbf{P}_{\mathbf{X}_{n-1}^a, \mathbf{Y}_n^f} = \mathbf{S}_{\mathbf{X}_{n-1}^a} \mathbf{S}^T_{\mathbf{Y}_n^f}, = \mathbf{S}_{\mathbf{X}_{n-1}^a} \widetilde{\mathbf{H}}_n^T. \tag{13}$$

Using $\mathbf{S}_{\mathbf{X}_{n-1}^a} = \left[\mathbf{S}^T_{\mathbf{x}_{n-1}^a}, \mathbf{S}^T_{\mathbf{z}_{n-1}^a}\right]^T$, (12) can be split into smoothing steps for the free and forced states as:

$$\mathbf{x}_{n-1}^{s,i} = \mathbf{x}_{n-1}^{a,i} + \mathbf{K}_{\mathbf{x}_{n-1}^a, \mathbf{Y}_n^f}(\mathbf{Y}_n - \mathbf{Y}_n^{f,i}), \tag{14}$$

$$\mathbf{z}_{n-1}^{s,i} = \mathbf{z}_{n-1}^{a,i} + \mathbf{K}_{\mathbf{z}_{n-1}^a, \mathbf{Y}_n^f}(\mathbf{Y}_n - \mathbf{Y}_n^{f,i}), \tag{15}$$

with $\mathbf{K}_{\mathbf{x}^a_{n-1},\mathbf{Y}^f_n} = \mathbf{P}_{\mathbf{x}^a_{n-1},\mathbf{Y}^f_n}\mathbf{P}^{-1}_{\mathbf{Y}^f_n}$ and $\mathbf{K}_{\mathbf{z}^a_{n-1},\mathbf{Y}^f_n} = \mathbf{P}_{\mathbf{z}^a_{n-1},\mathbf{Y}^f_n}\mathbf{P}^{-1}_{\mathbf{Y}^f_n}$. Consequently, both states are smoothed using the observations of both sub-systems ($\mathbf{y}^x_n$ and $\mathbf{y}^z_n$).

3.2. *The analysis step*

The analysis step integrates first the smoothed ensemble, $\{\mathbf{X}^{s,i}_{n-1}\}_{i=1}^{N_e}$, again with the coupled model to obtain the pseudo-forecast ensemble, $\widetilde{\mathbf{X}}^{f,i}_n = \left[(\widetilde{\mathbf{x}}^{f,i}_n)^T, (\widetilde{\mathbf{z}}^{f,i}_n)^T\right]^T$, as

$$\widetilde{\mathbf{x}}^{f,i}_n = \mathcal{M}^x_{n-1}\left(\mathbf{x}^{s,i}_{n-1}\right) + \widetilde{\eta}^{x,i}_{n-1}, \quad (16)$$

$$\widetilde{\mathbf{z}}^{f,i}_n = \mathcal{M}^z_{n-1}\left(\mathbf{z}^{s,i}_{n-1}, \mathbf{x}^{s,i}_{n-1}\right) + \widetilde{\eta}^{z,i}_{n-1}, \quad (17)$$

where $\widetilde{\eta}^{x,i}_{n-1}$ and $\widetilde{\eta}^{z,i}_{n-1}$ are samples from $\mathcal{N}(\mathbf{0}, \mathbf{Q}^x_{n-1})$ and $\mathcal{N}(\mathbf{0}, \mathbf{Q}^z_{n-1})$, respectively. These are then updated based on the same observation as,

$$\mathbf{X}^{a,i}_n = \widetilde{\mathbf{X}}^{f,i}_n + \mathbf{K}_{\widetilde{\mathbf{X}}^f_n,\widetilde{\mathbf{Y}}^f_n}(\mathbf{Y}_n - \widetilde{\mathbf{Y}}^{f,i}_n), \quad (18)$$

with $\widetilde{\mathbf{Y}}^{f,i}_n = \mathbf{H}_n\widetilde{\mathbf{X}}^{f,i}_n + \widetilde{\boldsymbol{\varepsilon}}^i_n$ and $\mathbf{K}_{\widetilde{\mathbf{X}}^f_n,\widetilde{\mathbf{Y}}^f_n} = \mathbf{P}_{\widetilde{\mathbf{X}}^f_n,\widetilde{\mathbf{Y}}^f_n}\mathbf{P}^{-1}_{\widetilde{\mathbf{Y}}^f_n}$. In Appendix A, we show that, in a Kalman setting (under the linear Gaussian assumption), the cross-terms in the analysis Kalman gain $\mathbf{K}_{\widetilde{\mathbf{X}}^f_n,\widetilde{\mathbf{Y}}^f_n}$ are zero. This implies that the analysis step of the EnKF with OSA-smoothing updates each of the two states using its own observation only, which is of great practical interest. Their analysis steps can thus be expressed as

$$\mathbf{x}^{a,i}_n = \widetilde{\mathbf{x}}^{f,i}_n + \mathbf{K}_{\widetilde{\mathbf{x}}^f_n,\widetilde{\mathbf{y}}^{f_x}_n}(\mathbf{y}^x_n - \widetilde{\mathbf{y}}^{f_x,i}_n), \quad (19)$$

$$\mathbf{z}^{a,i}_n = \widetilde{\mathbf{z}}^{f,i}_n + \mathbf{K}_{\widetilde{\mathbf{z}}^f_n,\widetilde{\mathbf{y}}^{f_z}_n}(\mathbf{y}^z_n - \widetilde{\mathbf{y}}^{f_z,i}_n), \quad (20)$$

with,

$$\mathbf{K}_{\widetilde{\mathbf{x}}^f_n,\widetilde{\mathbf{y}}^{f_x}_n} = \mathbf{P}_{\widetilde{\mathbf{x}}^f_n,\widetilde{\mathbf{y}}^{f_x}_n}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}^{f_x}_n}, \quad (21)$$

$$\mathbf{K}_{\widetilde{\mathbf{z}}^f_n,\widetilde{\mathbf{y}}^{f_z}_n} = \mathbf{P}_{\widetilde{\mathbf{z}}^f_n,\widetilde{\mathbf{y}}^{f_z}_n}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}^{f_z}_n} \quad (22)$$

where $\widetilde{\mathbf{y}}^{f_x,i}_n = \mathbf{H}^x_n\mathbf{x}^{f,i}_n + \boldsymbol{\varepsilon}^{x,i}_n$ and $\widetilde{\mathbf{y}}^{f_z,i}_n = \mathbf{H}^z_n\mathbf{z}^{f,i}_n + \boldsymbol{\varepsilon}^{z,i}_n$, with $\boldsymbol{\varepsilon}^{x,i}_n$ and $\boldsymbol{\varepsilon}^{z,i}_n$ sampled from $\mathcal{N}(\mathbf{0}, \mathbf{R}^x_n)$ and $\mathcal{N}(\mathbf{0}, \mathbf{R}^z_n)$, respectively. Note that, although $\mathbf{x}^{a,i}_n$ and $\mathbf{z}^{a,i}_n$ are updated separately, their associated (analysis) pdfs are still dependent, i.e., the covariance of the joint state analysis $\mathbf{P}_{\mathbf{X}^a_n}$ is not block-diagonal. Furthermore, in a linear-Gaussian setting, the $\mathbf{P}_{\mathbf{X}^a_n}$ of KF-OSA exactly matches that of KF (Desbouvries *et al.* 2011).

Summary of SC-EnKF$_{OSA}$ algorithm: Starting from an analysis ensemble, $\{\mathbf{X}^{a,i}_{n-1}\}_{i=1}^{N_e}$, with $\mathbf{X}^{a,i}_{n-1} = \left[(\mathbf{x}^{a,i}_{n-1})^T, (\mathbf{z}^{a,i}_{n-1})^T\right]^T$,

- **Smoothing Step**
  - $\mathbf{s}_1-$ The forecast ensemble, $\{\mathbf{X}^{f,i}_n\}_{i=1}^{N_e}$, is computed by integrating $\{\mathbf{X}^{a,i}_{n-1}\}_{i=1}^{N_e}$ with the coupled model to the time of the next available observation following (3) and (4).
  - $\mathbf{s}_2-$ $\{\mathbf{X}^{f,i}_n\}_{i=1}^{N_e}$ is then used with the incoming observation,$\mathbf{Y}_n$, to jointly smooth $\{\mathbf{X}^{a,i}_{n-1}\}_{i=1}^{N_e}$ and obtain the smoothed ensemble, $\{\mathbf{X}^{s,i}_{n-1}\}_{i=1}^{N_e}$, as in (12).
- **Analysis Step**
  - $\mathbf{a}_1-$ $\{\mathbf{X}^{s,i}_{n-1}\}_{i=1}^{N_e}$ is integrated with the model to compute the pseudo-forecast ensemble $\widetilde{\mathbf{X}}^{f,i}_n = \left[(\widetilde{\mathbf{x}}^{f,i}_n)^T, (\widetilde{\mathbf{z}}^{f,i}_n)^T\right]^T$ using (16) and (17).
  - $\mathbf{a}_2-$ $\{\widetilde{\mathbf{x}}^{f,i}_n\}_{i=1}^{N_e}$ are updated using $\mathbf{y}^x_n$ to compute $\{\mathbf{x}^{a,i}_n\}_{i=1}^{N_e}$ following (19).
  - $\mathbf{a}_3-$ $\{\widetilde{\mathbf{z}}^{f,i}_n\}_{i=1}^{N_e}$ are updated using $\mathbf{y}^z_n$ to compute $\{\mathbf{z}^{a,i}_n\}_{i=1}^{N_e}$ following (20).

SC-EnKF$_{OSA}$ applies two update steps (smoothing and analysis). Its weak version (WC-EnKF$_{OSA}$) is derived by separating the smoothing step of the free and the forced states (i.e., by updating each state using its own observations only), as the analysis update is readily in this form. This can stem from neglecting the off-diagonal blocks of $\mathbf{P}_{\mathbf{X}^a_{n-1},\mathbf{Y}^f_n}$ and $\mathbf{P}_{\mathbf{Y}^f_n}$, which could be derived in our Gaussian framework under the independence assumptions,

$$p(\mathbf{x}_{n-1}, \mathbf{z}_n|\mathbf{Y}_{0:n-1}) = p(\mathbf{x}_{n-1}|\mathbf{Y}_{0:n-1})p(\mathbf{z}_n|\mathbf{Y}_{0:n-1}), \quad (23)$$

$$p(\mathbf{z}_{n-1}, \mathbf{x}_n|\mathbf{Y}_{0:n-1}) = p(\mathbf{z}_{n-1}|\mathbf{Y}_{0:n-1})p(\mathbf{x}_n|\mathbf{Y}_{0:n-1}), \quad (24)$$

for $\mathbf{P}_{\mathbf{X}^a_{n-1},\mathbf{Y}^f_n}$, and the assumption (11) for $\mathbf{P}_{\mathbf{Y}^f_n}$. The WC-EnKF$_{OSA}$ has therefore the same algorithm except that the joint smoothing step ($\mathbf{s}_2$) is performed separately for each state sub-system using only its own observations.

4. Numerical experiments

4.1. *Experimental setup*

We use the multiscale Lorenz 96 (MS-L96) model (Lorenz 1996), in which N_x slow variables, $\{\mathbf{x}_i\}_{i=1}^{N_x}$, are coupled to $N_z = K \times$

N_x fast variables, $\{\mathbf{z}_{j,i}\}_{(j,i)=(1,1)}^{(K,N_x)}$, as,

$$\frac{d\mathbf{x}_i}{dt} = (\mathbf{x}_{i+1} - \mathbf{x}_{i-2})\,\mathbf{x}_{i-1} - \mathbf{x}_i + F - \frac{hc}{b}\sum_{J=1}^{K}\mathbf{z}_{j,i}, \qquad (25)$$

$$\frac{d\mathbf{z}_{j,i}}{dt} = \left(\mathbf{z}_{j-1,i} - \mathbf{z}_{j+2,i}\right) cb\mathbf{z}_{j+1,i} - c\mathbf{z}_{j,i} + \frac{hc}{b}\mathbf{x}_i, \qquad (26)$$

for $i = 1, \cdots, N_x$ and $j = 1, \cdots, K$. To avoid any confusion with the previous notations, we emphasize that the variable indices here refer to the number of state elements. The state variables $\mathbf{x}_i$ and $\mathbf{z}_{j,i}$ are periodic, i.e., $\mathbf{x}_{i+N_x} = \mathbf{x}_i$, $\mathbf{z}_{j,i+N_x} = \mathbf{z}_{j,i}$ and $\mathbf{z}_{j+K,i} = \mathbf{z}_{j,i+1}$. The MS-L96 simulates the weather variations at the mid-latitude and is often used to study the influence of different spatio-temporal scales on the predictability of atmospheric flows (e.g., Lorenz 1996; Palmer *et al.* 2005; Abramov 2016). $\mathbf{x}_i$ and $\mathbf{z}_{j,i}$ are respectively the slow and fast variables representing some atmospheric quantities discretized respectively into N_x and $N_x \times K$ sectors along the latitude circle (Lorenz 1996). Both variables are driven by quadratic (nonlinear) interaction terms modeling advection, constant forcing, linear damping, and coupling between a slow variable in one sector and K fast variables in the corresponding subsectors (Lorenz 1996; Frakkullin and Vanden-Eijnden 2004). Each $\mathbf{x}_i$ forces the $\mathbf{z}_{j,i}$ that are linked to it; the $\mathbf{z}_{j,i}$ then interact with each other with wave-like dynamics and feedback to the slow variables with a combined forcing. The parameters F, c, b, and h are constant and stand respectively for the forcing term, the spatial and temporal scale ratios, and the coupling coefficient.

We formulate a OWC MS-L96 model (consistent with system (1)) by neglecting the coupling term $(-\frac{hc}{b}\sum_{J=1}^{K}\mathbf{z}_{j,i})$ in Eq. (25). This yields the classical L-96 evolution equations for $\mathbf{x}_i$'s variables;

$$\frac{d\mathbf{x}_i}{dt} = (\mathbf{x}_{i+1} - \mathbf{x}_{i-2})\,\mathbf{x}_{i-1} - \mathbf{x}_i + F. \qquad (27)$$

A OWC system is so obtained where the variable $\mathbf{z}$ does not act on $\mathbf{x}$, whereas $\mathbf{z}$ is forced by $\mathbf{x}$ at each time step. Physically, this suggests that the chaotic fast variable system is also forced by a chaotic system, having an inherent slower timescale and leading to interesting chaotic dynamics in the $\mathbf{z}$ variable system. More precisely, the wave dynamics in $\mathbf{z}$ do not couple with the $\mathbf{x}$ dynamics whereas the $\mathbf{x}$ dynamics can still force the $\mathbf{z}$ variables with no feedback to $\mathbf{x}$. Examples of OWC models include one-way coupled hydro-meteorology models, where a meteorology model (e.g. WRF) can be one-way coupled to a hydrology model (e.g. WRF-Hydro) (Sampson and Gochis 2018). Similar OWC models are also found in applications related to marine ecosystem modeling (Triantafyllou *et al.* 2006) and more generally in marine and atmosphere transport models (e.g. pollution) and downscaled wind energy modeling. Resonance behavior in such coupled nonlinear systems as a consequence of the intrinsic dynamics has been studied previously (Gang *et al.* 1993; Lee *et al.* 1998; Jiang and Xin 2000). Let $N_x = 8$, $K = 16$ (i.e., $N_z = N_x \times K = 128$), $c = b = 10$ and $h = 1$, which corresponds to a coupling strength of $hc/b = 1$. However, more experiments (not shown here), confirmed that the conclusions from this work remain valid with stronger couplings. We also set $F = 10$ so that both scales are chaotic (Lorenz 1996). MS-L96 is discretized using the Runge-Kutta 4th-order scheme with a constant time step $\Delta t = 0.005$, which corresponds to 36 minutes in real time. We run the model for 10 years for spin-up, starting from an initial state sampled from a Gaussian density of zero-mean and identity covariance. The data assimilation experiments start from the end of the 10-year spin-up. The model is again integrated for 3 years and the resulting trajectory is considered as the truth. These are then used to generate synthetic observations by adding zero-mean Gaussian noises with covariances $\sigma_x^2\mathbf{I}_{N_x}$ and $\sigma_z^2\mathbf{I}_{N_z}$. The observational error standard deviations for the slow ($\mathbf{x}$) and fast ($\mathbf{z}$) variables are respectively set to $\sigma_x = 1$ and $\sigma_z = 0.1$, which correspond to approximately 20% and 35% of the long term average standard deviation of the two states, respectively.

An initial ensemble is generated by perturbing the aforementioned initial state with Gaussian noise of zero mean and identity covariance. We implement all filters using the covariance inflation (Anderson and Anderson 1999) and localization techniques. Both states' ensembles ($\mathbf{x}$ and $\mathbf{z}$) are inflated each time an observation (of at least one of them) becomes available. Inflation in the OSA-smoothing schemes is applied for both forecast and pseudo-forecast ensembles. For localization, the most common techniques, local analysis and covariance localization (Sakov and Bertino 2011), are both distance-based strategies. This requires

the model and observation variables to be associated with physical locations defining the distances between them. This is however not the case in our application in which the distances between the slow and fast variables are not physical. Luo and Hoteit (2014) neglected the different spatial scales between each $\mathbf{x}_i$ and the corresponding $\mathbf{z}_{j,i}$ by assuming that they have the same physical location. By doing so, the same localization scale was imposed for both fast and slow components. In this work, we instead follow the correlation-based (adaptive) localization method of Luo *et al.* (2017), which estimates different cut-off radii for each state variable, based on the spatial distributions of the correlations between the model variables and the corresponding simulated observations (called hereafter correlation field). The model variables are then updated using only the observations with which they significantly correlate. This is practically done via a Schur product between a tapering matrix (constructed based on the correlation field) and the Kalman gain, such that the tapering coefficients are 0 or 1 depending on the significance of the correlation (defined by a threshold) between the model variables and the observations. We follow Luo *et al.* (2017) and use a large ensemble (of 10000 members) to compute a robust approximation of the correlation field, ρ^{∞}. An "oracle threshold" for each available observation, is then computed based on the standard derivation of the error between ρ_k^{∞} and the correlation field computed using the EnKF ensemble. This is separately performed at every filtering step and for each model state. In cases where the computation of ρ^{∞} is computationally demanding, one can still use a wavelet-based denoising algorithm to estimate the noise level in the wavelet domain, from which the threshold value can be deduced. The reader is referred to Luo *et al.* (2017) for more details about the method.

The filters' performances are evaluated using the root-mean-square error (RMSE) between the reference states, $\mathbf{U}_{(i,n)}$, and the filters' estimates, $\hat{\mathbf{U}}_{(i,n)}$ (the index i stands for the number of state elements and n for the time), averaged over all its components and over the assimilation period of N_n cycles:

$$\text{RMSE} = \frac{1}{N_n}\sum_{n=1}^{N_n}\sqrt{\frac{1}{N_x}\sum_{i=1}^{N_x}(\mathbf{u}_{(i,n)} - \hat{\mathbf{u}}_{(i,n)})^2}\,. \qquad (28)$$

Here, $\mathbf{u}$ can be either $\mathbf{x}$ or $\mathbf{z}$ and therefore two different RMSEs are computed: the marginal RMSE for $\mathbf{x}$ (x-RMSE) and the marginal RMSE for $\mathbf{z}$ (z-RMSE). We further analyze the ensemble spread and evaluate the cross-correlation terms. To reduce statistical fluctuations, our results are averaged over 10 independent random realizations (with different initial ensembles and observational errors).

4.2. *Results and discussions*

We evaluate the filters' performances under various scenarios. We particularly study the sensitivities to the ensemble size, the length of the assimilation window and the number of assimilated observations. Three observational scenarios are investigated for both model components: full (i.e., all model variables are observed), half (i.e., every second variable is observed) and quarter (i.e., every fourth variable is observed). To give an idea about the different temporal scales between the state variables of the fast and slow components, we present in Fig. 1 the time series for some components of the free state ($\mathbf{x}_1$, $\mathbf{x}_4$ and $\mathbf{x}_8$) and the forced state ($\mathbf{z}_{1,1}$, $\mathbf{z}_{4,4}$ and $\mathbf{z}_{8,8}$) during the first 5000 model steps, confirming that the $\mathbf{z}$ variables change much faster in time compared to the $\mathbf{x}$ variables.

4.2.1. *Sensitivity to the ensemble size*

We assimilate the data of both states ($\mathbf{y}^x$ and $\mathbf{y}^z$) every day (i.e, every 40 model steps (ms)), and consider the scenario in which half of the observations of both states ($\mathbf{x}$ and $\mathbf{z}$) are assimilated. We report in Figs. 2, 3 and 4 the RMSEs, as a function of the inflation factor, as they result from SC-EnKF, SC-EnKF$_{OSA}$, WC-EnKF and WC-EnKF$_{OSA}$ with 20, 40 and 80 members, respectively. Comparing the strong and weak filtering schemes for both standard and OSA-smoothing formulations, one can see that the advantage of SCDA over WCDA becomes noticeable only when "enough" members are used (starting from $N_e = 40$ for SC-EnKF$_{OSA}$ and 80 for SC-EnKF). With the smallest ensemble size, $N_e = 20$, WC-EnKF and WC-EnKF$_{OSA}$ significantly outperform SC-EnKF and SC-EnKF$_{OSA}$, respectively. In this case, SC-EnKF reaches a minimum x-RMSE and z-RMSE of 1.81 and 0.232, respectively, compared to 0.96 and 0.218 for WC-EnKF. This supports the idea that the cross-correlations should be exploited

during the update step when enough ensemble members are considered to obtain statistically meaningful estimates.

As for the comparison of SC-EnKF with SC-EnKF$_{OSA}$, and WC-EnKF with WC-EnKF$_{OSA}$, it is clear that the OSA-smoothing-based filters outperform their standard counterparts. This is particularly pronounced when a small ensemble is used. For example, with 20 members, the minimum x-RMSE and z-RMSE achieved by SC-EnKF are respectively 1.81 and 0.232 compared to 0.94 and 0.201 with SC-EnKF$_{OSA}$, which corresponds to relative improvements of about 48% and 13%, respectively. Similar results are obtained for WC-EnKF$_{OSA}$ compared with WC-EnKF. As the ensemble size increases, the effect of the extra OSA-smoothing step becomes less pronounced for both weak and strong schemes, as can be seen in Fig. 4. The results further suggest that with an under-sampled ensemble of 20 members, SC-EnKF$_{OSA}$ yields substantially improved estimates, which are even slightly better than those of WC-EnKF. In coupled systems, the cross-covariances between the two coupled systems seem to generally be more prone to sampling errors than the covariances, and these have been damped by the OSA formulation. This confirms the effect of the extra smoothing step, which improves the background representation by constraining the forecast ensemble sampling with the "future" data, notably with small ensembles. In addition, the separate updates of the analysis step of SC-EnKF$_{OSA}$ (without abuse of language, but keeping in mind that SC-EnKF$_{OSA}$ does not neglect the cross-covariance terms in the analysis Kalman gain, as these are indeed null) further helps mitigating the impact of spurious cross-correlations, making the SC-EnKF$_{OSA}$ update more robust to sampling errors. For instance, the cross-covariances between $\mathbf{z}_{i,j}$ and all $\mathbf{z}_{i,j}$ estimated by the SC-EnKF, SC-EnKF$_{OSA}$, WC-EnKF and WC-EnKF$_{OSA}$ at the 1000^{th} assimilation step using 20 members in Fig. 5 show that the sampling errors are indeed reduced by SC-EnKF$_{OSA}$ compared to SC-EnKF, making its performance closer to that of WC-EnKFs. This is further discussed in the next section.

The results also emphasize the importance of adequate tuning of the inflation values to achieve good performances. As expected, increasing the ensemble size reduces the need for inflation. Moreover, SC-EnKFs (with or without OSA-smoothing) require larger inflation values compared to their weak counterparts, which is more pronounced with smaller ensembles. This suggests a need for larger inflation values to better exploit the cross-observations, as a way to mitigate the sampling errors in the (ensemble-approximated) cross-correlations. With large enough ensembles ($N_e = 160$ and 320), strong and weak filtering schemes require similar inflation values. Although WC-EnKF and WC-EnKF$_{OSA}$ require comparable inflation values, SC-EnKF$_{OSA}$ could work with less inflation than SC-EnKF, which may translate a better estimation of the cross-correlations in the smoothing step, or could be also related to the separate updates in the analysis step (second update) of SC-EnKF$_{OSA}$.

To understand to which extent observations from one state affects the other with the strong coupled assimilation schemes, we conduct four experiments that only differ in the cross-observations assimilation strategies. The first experiment (Exp1) corresponds to the most general setting where all observations are used to update both states (i.e., strongly coupled assimilation). In the second experiment (Exp2), $\mathbf{x}$-observations ($\mathbf{y}^x$) are not used to update $\mathbf{z}$. In Exp3, $\mathbf{z}$-observations ($\mathbf{y}^z$) are not used to update $\mathbf{x}$. Finally, in Exp4, each observation updates only its state (i.e., weakly coupled assimilation). Fig. 6 displays barplots of the minimum RMSEs using a small ensemble (20 members) and a large ensemble (160 members) to also assess the effect of the ensemble size on the results. With 160 members and for the estimation of $\mathbf{x}$, the results of Exp3 are of course the same as those of the weak formulation (Exp4), which is actually expected since $\mathbf{x}$ is not dependent on $\mathbf{z}$ and uses only its observations in both scenarios. Results of Exp2 are comparable to those of Exp1, which suggests that the improvement introduced by the strong formulation (over the weak formulation) for $\mathbf{x}$ is mainly due to the assimilation of $\mathbf{y}^z$. As for the forced (fast) state $\mathbf{z}$, comparing Exp3 and Exp4 suggests that assimilating $\mathbf{y}^x$ only marginally improves the results compared to the weak formulation, suggesting that $\mathbf{y}^x$ are not very informative for $\mathbf{z}$. This is because $\mathbf{y}^x$ does not provide high-frequency information relevant for the fast variable $\mathbf{z}$, as has been also suggested by Liu *et al.* (2013). This is further supported by the fact that the results of Exp2 are close to the those of Exp1. Hence, the improvements introduced by the strong formulation over the weak one for the estimation of $\mathbf{z}$ are not

related to the update of $\mathbf{z}$ by $\mathbf{y}^x$, but to the cross-assimilation of $\mathbf{y}^z$ by $\mathbf{x}$, suggesting the important effect of the coupling with better $\mathbf{x}$-estimates. One may therefore choose to update the fast variables using only the corresponding observations, whereas the slow state variables are still updated with observations from both components. Fig. 6 further suggests that with a relatively small ensemble of 20 members, assimilating the cross-observations generally introduces noise to both system components, which is alleviated by the OSA-smoothing formulation.

Fig. 7 plots the evolution of the minimum x-RMSE and z-RMSE using six different ensemble sizes varying between 20 to 320. The results suggest three important findings. First, applying the OSA-smoothing step improves the assimilation results, particularly when the filters are implemented with small ensembles. Second, in these small ensemble cases, the (extra update) OSA-smoothing step significantly reduces the gap (in terms of minimum RMSE) between the weak and strong coupled assimilation strategies. For instance, with 10 members, SC-EnKF and WC-EnKF respectively achieve 2.69 and 1.33 minimum x-RMSEs, about 50% difference, while SC-EnKF$_{OSA}$ provides a minimum x-RMSE of 1.41 compared to 0.97 with WC-EnKF$_{OSA}$, a relative improvement of 31%. Third, SC-EnKF$_{OSA}$ generally outperforms both SC-EnKF and WC-EnKF. With very small ensembles, WC-EnKF$_{OSA}$ provides the best results. Overall, these results suggest the relevance of the OSA-smoothing step for enhancing the performances of the EnKF with OWC by providing more reliable background covariances (and cross-covariances). With small ensembles, this may still not be enough, but WC-EnKF$_{OSA}$ could still be quite beneficial. Hereafter, we will focus on analyzing the behavior of the filter when both state sub-systems are updated with both states observations, i.e. the strongly coupled DA framework.

We report in Fig. 8 the percentages of relative improvements resulting from SC-EnKF$_{OSA}$ with respect to SC-EnKF for different ensemble sizes (only positive percentages are reported). We are in particular interested in evaluating the SC-EnKF$_{OSA}$ results when implemented with half of the ensemble size used for SC-EnKF, so that both schemes require roughly the same computational load. With the same ensemble size, SC-EnKF$_{OSA}$ outperforms SC-EnKF in practically all cases (except for the estimation of $\mathbf{x}$ with 160 and 320 members where both schemes yield comparable results). This tends to be more pronounced with relatively small ensembles and for the state of the slow components $\mathbf{x}$ (up to 50%), while for the fast state $\mathbf{z}$, these improvements do not exceed 13%. When implemented with half the number of members, SC-EnKF$_{OSA}$ still generally outperforms SC-EnKF, mainly in the small ensemble size cases ($N_e = 10$ or 20). In the rest of our experiments, we will use two ensemble sizes, 40 and 80, to assess the filters' performances under the same computational cost.

4.2.2. *Sensitivity to the assimilation windows*

We study in this section the effect of varying the length of the assimilation window of both sub-systems. We assimilate half of the observations from both sub-systems and conduct two sets of experiments to illustrate the sensitivity to $\mathbf{y}^x$ and $\mathbf{y}^z$. Fig. 9 displays the evolution of the minimum RMSEs for different assimilation windows of $\mathbf{y}^z$ (ranging between every 5 ms and every 1000 ms). In these runs, $\mathbf{y}^x$ are assimilated every 40 ms in Figs. 9.a-b and every 10 ms in Figs. 9.c-d. As expected, the RMSE values for both $\mathbf{x}$ and $\mathbf{z}$ decrease as the ensemble size and the frequency of assimilating $\mathbf{y}^z$ increase. One can also notice that SC-EnKF$_{OSA}$ generally outperforms SC-EnKF in the tested scenarios, and this becomes more pronounced when only 40 members are used. Comparing the results of Fig. 9.a-b, when $\mathbf{y}^x$ are assimilated every 40 ms, with those of 9.c-d, when they are observed every 10 ms, one can see that a more frequent assimilation of $\mathbf{y}^x$ has a very small impact on improving the estimation of $\mathbf{z}$ (plots (b) and (d)). The effect on the estimation of $\mathbf{x}$ is much more noticeable (plots (a) and (c)). This is consistent with the results of Section 4.2.1 where $\mathbf{y}^x$ was not very informative for $\mathbf{z}$. The results further suggest that with the assimilation of $\mathbf{y}^z$ every 5 to 40 ms, SC-EnKF$_{OSA}$'s estimates with 40 members are close to those of SC-EnKF with 80 members. When $\mathbf{y}^z$ is assimilated less frequently (beyond 40 ms), SC-EnKF$_{OSA}$ tends to provide similar or even better estimates than SC-EnKF, especially when $\mathbf{z}$ is observed every 40 ms. Fig. 9.a indeed suggests that SC-EnKF$_{OSA}$ can outperform SC-EnKF with only half the ensemble size. The effect of the OSA-smoothing formulation seems to

become much more pronounced as less observations from the fast component are assimilated in time.

Fig. 10 illustrates the sensitivity of the filtering schemes to the assimilation frequency of $\mathbf{y}^x$. As above, we explore two values for the assimilation interval of $\mathbf{y}^z$ (40 and 10 ms) and analyze the RMSEs values for different assimilation periods of $\mathbf{y}^x$. Comparing SC-EnKF and SC-EnKF$_{OSA}$, we notice that when the fast variables are frequently observed (every 10 ms), SC-EnKF$_{OSA}$ outperforms SC-EnKF in all tested scenarios, and again, this is more pronounced with small ensembles. When assimilating $\mathbf{y}^z$ frequently enough (every 10 ms), the increase in the estimation error with the assimilation window of $\mathbf{y}^x$ for both components (especially $\mathbf{x}$) is expected to be alleviated as the system is well-constrained by $\mathbf{y}^z$. By increasing the ensemble size from 40 to 80, both SC-EnKF and SC-EnKF$_{OSA}$ estimates become less sensitive to the assimilation frequency of $\mathbf{y}^x$ as $\mathbf{y}^z$ is more efficiently assimilated. Similarly, the two-stage update of SC-EnKF$_{OSA}$ is less sensitive than that of SC-EnKF to the assimilation window of $\mathbf{y}^x$, in particular when 80 members are used. Increasing the assimilation window of $\mathbf{y}^z$ to 40 ms results in different (opposite) behaviors of SC-EnKF and SC-EnKF$_{OSA}$, depending on the assimilation frequency of $\mathbf{y}^x$. Indeed, when the latter is shorter than 40 ms (i.e., the system assimilation step is also shorter than 40 ms), SC-EnKF$_{OSA}$ clearly outperforms SC-EnKF, mainly when the filters are implemented with 40 members. With longer assimilation windows of $\mathbf{y}^x$ (i.e., the system assimilation step is set to 40 ms), however, the behavior of SC-EnKF$_{OSA}$ starts degrading, especially for the estimation of the slow variable $\mathbf{x}$. Indeed, when $\mathbf{y}^x$ are assimilated every more often than 40 ms, SC-EnKF$_{OSA}$ performs successive OSA-smoothing updates based on $\mathbf{y}^z$ only, before $\mathbf{y}^x$ becomes available. It seems that assimilating $\mathbf{y}^z$ based on a standard update is more efficient since a smoothing window of 40 ms is relatively large. The smoothing step is indeed beneficial only up to a certain window within which the linear smoothing step is viable, beyond which the OSA smoothing filters performances start to degrade (Raboudi *et al.* 2018). Obviously, this problem does not occur when $\mathbf{y}^x$ are assimilated less frequently than 40 ms, and is further alleviated when $\mathbf{y}^x$ are assimilated every 40 or 80 ms. As an illustration, we conducted an experiment assimilating $\mathbf{y}^x$ every 200 ms and $\mathbf{y}^z$ every 40 ms using 40 members such that, whenever $\mathbf{y}^x$ is available, both observations are assimilated based on a OSA-smoothing update, whereas all the assimilation steps of $\mathbf{y}^z$ only (i.e., between two successive occurrences of $\mathbf{y}^x$) are standard. This results in more accurate estimates than those of SC-EnKF and SC-EnKF$_{OSA}$.

To further assess the effect of the smoothing step on the estimation of $\mathbf{x}$ and $\mathbf{z}$, we plot in Fig. 11 the time evolution of the forecast and analysis RMSEs as estimated by SC-EnKF and SC-EnKF$_{OSA}$, along with the forecast ensemble standard deviation. The assimilation windows for $\mathbf{y}^x$ and $\mathbf{y}^z$ are respectively 40 and 10 ms and the ensemble size is 40. The results correspond to the inflation factor that provides the lowest (joint) RMSEs. We plot the results between the 100^{th} and the 300^{th} assimilation steps, but report the averaged RMSEs over the three-year simulation period. For both $\mathbf{x}$ and $\mathbf{z}$, the forecast and analysis RMSEs resulting from SC-EnKF$_{OSA}$ are lower than those of SC-EnKF, with relative improvements of respectively 40% and 27%. As for the ensemble spread, it is also reduced by SC-EnKF$_{OSA}$, suggesting less uncertainty on the estimated state. Similarly, we plot in Figs. 12 the time series of the cross-covariances (between $\mathbf{x}_1$ and $\mathbf{z}_{1,1}$, $\mathbf{x}_4$ and $\mathbf{z}_{4,4}$ and $\mathbf{x}_8$ and $\mathbf{z}_{8,8}$) between the 100^{th} and 300^{th} assimilation steps, as estimated by SC-EnKF and SC-EnKF$_{OSA}$, using three different ensemble sizes: 20 (Figs. 12.a-d-g), 40 (Figs. 12.b-e-h) and 320 (Figs. 12.c-f-i). With a relatively large ensemble, of 320 members, both SC-EnKF and SC-EnKF$_{OSA}$ provide similar (and reliable) estimates of the cross-covariances, which by comparison to those estimated by only 40 and 20 members, suggest better SC-EnKF$_{OSA}$ estimates of cross-covariances. The results further suggest that the OSA-smoothing formulation reduces the noise in the cross-covariance terms, and this becomes more pronounced with smaller ensembles. For example, by comparing the results with 20 and 40 members, one can notice that the ensemble cross-covariances as computed by both SC-EnKF and SC-EnKF$_{OSA}$ become more dominated by noise, but to a clearly lesser extent for SC-EnKF$_{OSA}$, confirming the improved robustness of this filter to sampling errors.

We further examine in Fig. 13 the posterior ensemble and uncertainties as they result from both schemes. The first row plots

the analysis ensemble as estimated by the SC-EnKF and SC-EnKF$_{OSA}$ for $\mathbf{x}_1$, using 20, 40 and 320 members (between the 100^{th} and the 300^{th} assimilation steps) as well as the observations (black stars). The figure shows that with 320 members, the filters ensembles spreads are similar, whereas with 40 and particularly 20 members, SC-EnKF$_{OSA}$ exhibits less spread, suggesting less uncertainty in the estimated state. Moreover, in both schemes, the observations are generally within these uncertainties from the estimates, suggesting their consistency. The second row plots the analysis RMSE and ensemble standard deviation for $\mathbf{x}_1$ during the same period, and suggests overall less errors and variance for SC-EnKF$_{OSA}$. The RMSEs are further consistent with the standard deviations in the case of SC-EnKF, and to a larger extent for SC-EnKF$_{OSA}$. The third row displays the rank histograms of the ensembles of schemes using 20, 40 and 320 members. Rank histograms are common indicators of the reliability of the filter ensemble (Harrison *et al.* 1995; Talagrand 1999). For a properly samples ensemble, small analysis error should be accompanied by uniformly distributed rank histograms (Hamill and Colucci 1997). Ideally, the rank of the true state relative to a sorted N_e-ensemble should be equally likely to occur in any of the $N_e - 1$ possible ranks. Hence, over many samples, a histogram of the ranks of observations relative to the sorted ensemble should be approximately uniform. Non-uniformity could indicate potential problems in the ensemble. The results suggest that SC-EnKF$_{OSA}$ exhibits relatively flatter rank histograms, particularly will small ensembles. For instance, with 20 members, the SC-EnKF rank histogram is concave downward, suggesting an over-dispersive ensemble, while SC-EnKF$_{OSA}$ shows a more desired histogram.

4.2.3. *Sensitivity to the number of observations*

This section investigates the effect of the number of assimilated observations on the filters' performances. Based on the results of the previous section, we assimilate $\mathbf{y}^x$ every 40 ms and $\mathbf{y}^z$ every 10 ms. Fig. 14 plots the RMSEs, as function of the inflation factor, as they result from SC-EnKF and SC-EnKF$_{OSA}$ when respectively all, half and quarter of both $\mathbf{y}^x$ and $\mathbf{y}^z$ are assimilated. As expected, the filters' behavior improves with the assimilation of more observations and SC-EnKF$_{OSA}$ is shown to significantly outperform SC-EnKF in all tested scenarios. In Fig. 15, we report the results of the experiment assimilating observations from half of the variables of one state ($\mathbf{x}$ or $\mathbf{z}$) and vary the number of observed variables from the other state. As expected, the results mainly suggest that the number of assimilated $\mathbf{y}^x$ does not have an important impact on the estimation of $\mathbf{z}$, while the number of assimilated $\mathbf{y}^z$ have a clear impact on the estimation of $\mathbf{x}$, suggesting as before that $\mathbf{y}^x$ does not provide much information about $\mathbf{z}$.

We further investigate the cases where observations of only one state variable are available. Fig. 16 illustrates the filters' performances when only $\mathbf{y}^x$ are assimilated (thus, the estimation of $\mathbf{x}$ is equivalent to the weak formulation; see Fig. 3), whereas the forced variable $\mathbf{z}$ is updated through the coupling with $\mathbf{x}$. When $\mathbf{y}^x$ are frequently available in time, standard and OSA-smoothing schemes yield comparable results. The benefit of SC-EnKF$_{OSA}$ becomes clear when less $\mathbf{y}^x$ are assimilated. This is obviously up to a certain frequency beyond which the smoothing step becomes less robust, as discussed in Section 4.2.2 ($\sim$ 60 ms). Fig. 17 presents the results of the second experiment when only $\mathbf{y}^z$ are assimilated. With an assimilation window of 40 ms, SC-EnKF outperforms SC-EnKF$_{OSA}$, again because of the large assimilation window. Assimilating over a shorter window of 10 ms, SC-EnKF$_{OSA}$ clearly outperforms SC-EnKF in the estimation of both $\mathbf{x}$ and $\mathbf{z}$, with relative improvements of about 36% for $\mathbf{x}$ and 25% for $\mathbf{z}$.

4.2.4. *Robustness to bias in the forecast model*

We finally study the sensitivity of the filtering schemes to a biased forecast model. We choose to introduce a bias in the MS-L96 model by reducing the values of the parameters F, c, b, and h in (27) and (26) by 20% ($c = b = 8$, $h = 0.8$ and $F = 8$) and 40% ($c = b = 6$, $h = 0.6$ and $F = 6$) compared to their reference values during the forecast steps. We further consider the challenging scenario where we assimilate only one-quarter of the observations from each model component every 40 ms. Fig. 18 shows that, with 20% of bias, SC-EnKF$_{OSA}$ results are about 20% and 9% better in terms of x-RMSE and z-RMSE respectively. With 40% of bias, SC-EnKF$_{OSA}$ can still outperform SC-EnKF, suggesting its robustness to bias in the forecast model.

5. Discussions

This study considered the ensemble data assimilation (DA) problem into one-way coupled (OWC) models, a class of coupled models in which one of the two sub-systems is forced by the other, using the Ensemble Kalman Filter (EnKF). Although the implementation of the EnKF with such systems can be commonly performed by concatenating the states and observations of the coupled system components, so that the coupled system is treated as a standard state-space-model, several studies reported several challenges associated with this implementation in realistic applications, namely the differences in time and temporal scales between the different components, the strongly nonlinear character of the coupling between the system components, and the increase in the computational cost. In these situations, it is challenging to estimate reliable cross-correlations, characterizing the exchange of information between the components at the analysis step.

Here, we considered the one-step-ahead (OSA) smoothing formulation of the filtering problem to derive new EnKF schemes better suited to deal with OWC DA problems. The OSA-smoothing formulation reverses the order of the standard "forecast-then-update" path of the filtering problem, which introduces an extra update step based on the future observations. Retaining the sampling of the forecast ensemble with the future observations provides enhanced background statistics that were shown to be quite beneficial for efficient cross-transferring of the information between the system components. An EnKF algorithm with strong coupling was then derived based on the OSA-smoothing formulation (SC-EnKF$_{OSA}$), with the particularity of involving a joint smoothing step for both state components using the future observations of both variables and a separate analysis step for both states (involving two separate updates of each state component, each assimilating its own observations only). We further derived its weak version (WC-EnKF$_{OSA}$) by neglecting the cross-updates in the joint smoothing step.

The performances of the EnKF$_{OSA}$ schemes were tested and compared with their standard EnKF counterparts (SC-EnKF and WC-EnKF) by conducting extensive numerical experiments with the OWC multiscale Lorenz 96 (MS-L96) model, which couples slow and fast variables. Our numerical results demonstrated that EnKF with OSA-smoothing is more efficient and robust for estimating both state components. This was particularly more pronounced in the challenging situations when the filters were implemented with small ensembles, limited number and frequency of observations, assimilated observations from one model component only, and biased forecast model. In our experiments, SC-EnKF$_{OSA}$ outperformed all other filtering schemes, except when very small ensembles were used, cases in which WC-EnKF$_{OSA}$ becomes more reliable. Acknowledging that the computational costs of the OSA-smoothing schemes is almost double compared to those of the standard schemes, we also demonstrated that the SC-EnKF$_{OSA}$ is able to outperform the SC-EnKF using only half the ensemble members.

Our results further suggested that SC-EnKF$_{OSA}$ has a particular benefit on the slow component, which could be related to two factors: the characteristics of the considered OWC model and the slow-fast nature of its coupled dynamics. In the considered OWC model, the fast variable is function of the slow variable and therefore an observation of the fast component is also an observation of the slow one. One should therefore expect a larger impact on the slow component, particularly with the SC-EnKF$_{OSA}$ through its two-stage update. Second, the observed information about the slow dynamics does not provide high-frequency information relevant for the fast variable and therefore the improvements should be more pronounced for the slow component, mainly with SC-EnKF$_{OSA}$ which better handles the transfer of cross-information between the coupled subsystems.

The proposed OSA-smoothing EnKF schemes were found quite efficient in the context of twin experiments with the OWC Lorenz-96 model. Our next step will be to test these schemes for data assimilation into realistic ocean and atmospheric transport models.

Acknowledgement

This work was supported by the Office of Sponsored Research (OSR) at King Abdullah University of Science and Technology (KAUST) under the Virtual Red Sea Initiative (Grant #REP/1/3268-01-01). The research made use of the KAUST supercomputing facilities.

6. Appendix [A]

Seperability of the analysis step of SC-EnKF$_{OSA}$

Under the assumption of $p(\mathbf{X}_{n-1}, \mathbf{X}_n, \mathbf{Y}_n|\mathbf{Y}_{0:n-1})$ Gaussian (which actually holds for linear Gaussian systems), the analysis Kalman gain $\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f}$ can be written as

$$\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f} = \underbrace{cov(\mathbf{X}_n, \mathbf{Y}_n|\mathbf{X}_{n-1}, \mathbf{Y}_{0:n-1})}_{\mathbf{C}_1} \big[\underbrace{cov(\mathbf{Y}_n|\mathbf{X}_{n-1}, \mathbf{Y}_{0:n-1})}_{\mathbf{C}_2}\big]^{-1} \tag{29}$$

where cov stands for the (cross)-covariance between the underlying variables (Desbouvries *et al.* 2011; Ait-El-Fquih *et al.* 2016). We focus here on the off-diagonal blocks of the (cross)-covariance blocks in $\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f}$:

$$\mathbf{C}_1 = \begin{bmatrix} \bullet & cov(\mathbf{x}_n, \mathbf{z}_n|\mathbf{X}_{n-1}, \mathbf{Y}_{0:n-1})(\mathbf{H}_n^z)^T \\ cov(\mathbf{z}_n, \mathbf{x}_n|\mathbf{X}_{n-1}, \mathbf{Y}_{0:n-1})(\mathbf{H}_n^x)^T & \bullet \end{bmatrix}, \tag{30}$$

and,

$$\mathbf{C}_2 = \begin{bmatrix} \bullet & \mathbf{H}_n^x cov(\mathbf{x}_n, \mathbf{z}_n|\mathbf{X}_{n-1}, \mathbf{Y}_{0:n-1})(\mathbf{H}_n^z)^T + cov(\boldsymbol{\varepsilon}_n^x, \boldsymbol{\varepsilon}_n^z) \\ \mathbf{H}_n^z cov(\mathbf{z}_n, \mathbf{x}_n|\mathbf{X}_{n-1}, \mathbf{Y}_{0:n-1})(\mathbf{H}_n^x)^T + cov(\boldsymbol{\varepsilon}_n^z, \boldsymbol{\varepsilon}_n^x) & \bullet \end{bmatrix}. \tag{31}$$

Assuming that $cov(\mathbf{x}_n, \mathbf{z}_n|\mathbf{X}_{n-1}, \mathbf{Y}_{0:n-1}) = 0$ as the model noise terms η_{n-1}^x and η_{n-1}^z are independent and independent of $\mathbf{x}_0$ and $\mathbf{z}_0$, and $cov(\boldsymbol{\varepsilon}_n^x, \boldsymbol{\varepsilon}_n^z) = 0$ as the observation noise terms $\boldsymbol{\varepsilon}^x$ and $\boldsymbol{\varepsilon}^z$ are independent, $\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f}$ becomes a block diagonal matrix of the form

$$\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f} = \begin{bmatrix} \widetilde{\mathbf{K}}_x & \widetilde{\mathbf{K}}_{xz} = 0_{N_x \times N_{y_z}} \\ \widetilde{\mathbf{K}}_{zx} = 0_{N_z \times N_{y_x}} & \widetilde{\mathbf{K}}_z \end{bmatrix}. \tag{32}$$

This could be also obtained by using the alternative expression of the analysis gain $\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f} = \mathbf{Q}_{n-1}\mathbf{H}_n^T(\mathbf{H}_n\mathbf{Q}_{n-1}\mathbf{H}_n^T + \mathbf{R}_n)^{-1}$ (Desbouvries *et al.* (2011); Raboudi *et al.* (2018)), with the augmented matrices $\mathbf{Q}_{n-1} = \begin{bmatrix} \mathbf{Q}_{n-1}^x & 0 \\ 0 & \mathbf{Q}_{n-1}^z \end{bmatrix}$, $\mathbf{R}_n = \begin{bmatrix} \mathbf{R}_n^x & 0 \\ 0 & \mathbf{R}_n^z \end{bmatrix}$ and $\mathbf{H}_n = \begin{bmatrix} \mathbf{H}_n^x & 0 \\ 0 & \mathbf{H}_n^z \end{bmatrix}$. $\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f}$ can thus be expressed as:

$$\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f} = \begin{bmatrix} \widetilde{\mathbf{K}}_x = \mathbf{Q}_{n-1}^x(\mathbf{H}_n^x)^T(\mathbf{H}_n^x\mathbf{Q}_{n-1}^x(\mathbf{H}_n^x)^T + \mathbf{R}_n^x)^{-1} & \widetilde{\mathbf{K}}_{xz} = 0_{N_x \times N_{y_z}} \\ \widetilde{\mathbf{K}}_{zx} = 0_{N_z \times N_{y_x}} & \widetilde{\mathbf{K}}_z = \mathbf{Q}_{n-1}^z(\mathbf{H}_n^z)^T(\mathbf{H}_n^z\mathbf{Q}_{n-1}^z(\mathbf{H}_n^z)^T + \mathbf{R}_n^z)^{-1} \end{bmatrix}. \tag{33}$$

With $\widetilde{\mathbf{K}}_{xz} = \widetilde{\mathbf{K}}_{zx} = 0$, one can show that $\widetilde{\mathbf{K}}_x = \mathbf{P}_{\widetilde{\mathbf{x}}_n^f, \widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx}}^{-1}$ and $\widetilde{\mathbf{K}}_z = \mathbf{P}_{\widetilde{\mathbf{z}}_n^f, \widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz}}^{-1}$ as given respectively in Eqs. (21) and (22). For brevity, we present only the derivations for $\widetilde{\mathbf{K}}_x$, as the expression of $\widetilde{\mathbf{K}}_z$ could be similarly derived. One has,

$$\mathbf{K}_{\widetilde{\mathbf{X}}_n^f, \widetilde{\mathbf{Y}}_n^f} = \begin{bmatrix} \mathbf{P}_{\widetilde{\mathbf{x}}_n^f, \widetilde{\mathbf{y}}_n^{fx}} & \mathbf{P}_{\widetilde{\mathbf{x}}_n^f, \widetilde{\mathbf{y}}_n^{fz}} \\ \mathbf{P}_{\widetilde{\mathbf{z}}_n^f, \widetilde{\mathbf{y}}_n^{fx}} & \mathbf{P}_{\widetilde{\mathbf{z}}_n^f, \widetilde{\mathbf{y}}_n^{fz}} \end{bmatrix} \times \begin{bmatrix} \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx}} & \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx}, \widetilde{\mathbf{y}}_n^{fz}} \\ \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz}, \widetilde{\mathbf{y}}_n^{fx}} & \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz}} \end{bmatrix}^{-1}. \tag{34}$$

Using the Schur complement theorem and some matrix inversion lemmas, one can write

$$
\begin{aligned}
\widetilde{\mathbf{K}}_x &= \left[\mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}} - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\right] \times \left[\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx}} - \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\right]^{-1}, \\
&= \left[\mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}} - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\right] \\
&\times \left[\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}} - \underbrace{\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}}\left(\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}} - \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz}}\right)^{-1}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}}_{\Sigma}\right], \\
&= \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}} + \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\left(\Sigma - \mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\right) - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\Sigma\,, \\
&= \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}} - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\left(\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx}} - \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fz}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\right)^{-1} - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\Sigma\,, \\
&= \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}} + \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fz}}\left(\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}} - \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz}}\right)^{-1}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}} - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\Sigma\,, \\
&= \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\,, \\
&+ \left(\mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fz}} - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}}\right)\left(\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz}} - \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}}\right)^{-1}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\,. \qquad (35)
\end{aligned}
$$

Using again the Schur complement theorem, one obtains

$$
\widetilde{\mathbf{K}}_{xz} = \left(\mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fz}} - \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}}\right)\left(\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz}} - \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}_{\widetilde{\mathbf{y}}_n^{fx},\widetilde{\mathbf{y}}_n^{fz}}\right)^{-1}, \qquad (36)
$$

so that (35) can be written as

$$
\widetilde{\mathbf{K}}_x = \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}} + \widetilde{\mathbf{K}}_{xz} \times \mathbf{P}_{\widetilde{\mathbf{y}}_n^{fz},\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}. \qquad (37)
$$

Since $\widetilde{\mathbf{K}}_{xz} = 0$, the expression of $\widetilde{\mathbf{K}}_x$ is reduced to $\widetilde{\mathbf{K}}_x = \mathbf{P}_{\widetilde{\mathbf{x}}_n^f,\widetilde{\mathbf{y}}_n^{fx}}\mathbf{P}^{-1}_{\widetilde{\mathbf{y}}_n^{fx}}$, hence (21).

References

Abramov R. 2016. A simple stochastic parameterization for reduced models of multiscale dynamics. *Fluids* **1**(1): 2.

Ait-El-Fquih B, El Gharamti M, Hoteit I. 2016. A Bayesian consistent dual ensemble Kalman filter for state-parameter estimation in subsurface hydrology. *Hydrology and Earth System Sciences* **20**(8): 3289–3307, doi: 10.5194/hess-20-3289-2016.

Anderson JL, Anderson SL. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review* **127**: 2741–2758.

Ballabrera-Poy J, Kalnay E, Yang SC. 2008. Data assimilation in a system with two scalescombining two initialization techniques. *Tellus A: Dynamic Meteorology and Oceanography* **61**(4): 539–549.

Benra FK, Dohmen HJ, Pei J, Schuster S, Wan B. 2011. A comparison of one-way and two-way coupling methods for numerical analysis of fluid-structure interactions. *Journal of applied mathematics* **2011**.

Browne PA, de Rosnay P, Zuo H, Bennett A, Dawson A. 2019. Weakly coupled ocean–atmosphere data assimilation in the ecmwf nwp system. *Remote Sensing* **11**(3): 234.

Desbouvries F, Ait-El-Fquih B. 2008. Direct, prediction-based and smoothing-based particle filter algorithms. In: *Proc. 4th world conference of the Int. Ass. for Statistical Computing (IASC 2008) Yokohama, Japan*.

Desbouvries F, Petetin Y, Ait-El-Fquih B. 2011. Direct, prediction- and smoothing-based Kalman and particle filter algorithms. *Signal Process.* **91**(8): 2064–2077, doi:10.1016/j.sigpro.2011.03.013.

El Gharamti M, Hoteit I, Valstar J. 2013. Dual states estimation of a subsurface flow-transport coupled model using ensemble kalman filtering. *Advances in water resources* **60**: 75–88.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans* **99**(C5): 10 143–10 162.

Evensen G. 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics* **53**(4): 343–367.

Fatkullin I, Vanden-Eijnden E. 2004. A computational strategy for multiscale systems with applications to lorenz 96 model. *Journal of Computational Physics* **200**(2): 605–638.

Frolov S, Bishop CH, Holt T, Cummings J, Kuhl D. 2016. Facilitating strongly coupled ocean–atmosphere data assimilation with an interface solver. *Monthly Weather Review* **144**(1): 3–20.

Gang H, Ditzinger T, Ning CZ, Haken H. 1993. Stochastic resonance without external periodic force. *Physical Review Letters* **71**(6): 807.

Gharamti ME, Ait-El-Fquih B, Hoteit I. 2015. An iterative ensemble Kalman filter with one-step-ahead smoothing for state-parameters estimation of contaminant transport models. *Journal of Hydrology* **527**: 442–457.

Gharamti ME, Kadoura A, Valstar J, Sun S, Hoteit I. 2014. Constraining a compositional flow model with flow-chemical data using an ensemble-based kalman filter. *Water Resources Research* **50**(3): 2444–2467.

Hamill TM, Colucci SJ. 1997. Verification of eta–rsm short-range ensemble forecasts. *Monthly Weather Review* **125**(6): 1312–1327.

n G, Wu X, Zhang S, Liu Z, Li W. 2013. Error covariance estimation for coupled data assimilation using a lorenz atmosphere and a simple pycnocline ocean model. *Journal of Climate* **26**(24): 10 218–10 231.

rrison M, Richardson D, Robertson K, Woodcock A. 1995. Medium-range ensembles using both the ecmwf t63 and unified modelsan initial report. *UK Meteorological Office Tech. Rep* **153**: 25.

Hendricks Franssen HJ, Kinzelbach W. 2008. Real-time groundwater flow modeling with the ensemble kalman filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water Resources Research* **44**(9).

Hoteit I, Luo X, Bocquet M, Kohl A, Ait-El-Fquih B. 2018. *New frontiers in operational oceanography*, ch. 17: Data assimilation in oceanography: current status and new directions. GODAE Ocean View, pp. 465–512.

ang Y, Xin H. 2000. Coherent resonance in a one-way coupled system. *Physical Review E* **62**(2): 1846.

Kalman RE. 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82**(1): 35–45.

Kalnay E, Yang SC. 2010. Accelerating the spin-up of ensemble kalman filtering. *Quarterly Journal of the Royal Meteorological Society* **136**(651): 1644–1651.

Lahoz WA, Schneider P. 2014. Data assimilation: making sense of earth observation. *Frontiers in Environmental Science* **2**: 16.

Laloyaux P, Balmaseda M, Dee D, Mogensen K, Janssen P. 2016. A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society* **142**(694): 65–78.

Lee SG, Neiman A, Kim S. 1998. Coherence resonance in a hodgkin-huxley neuron. *Physical Review E* **57**(3): 3292.

Liu B, Ait-El-Fquih B, Hoteit I. 2016. Efficient kernel-based ensemble Gaussian mixture filtering. *Monthly Weather Review* **144**(2): 781–800.

Liu Z, Wu S, Zhang S, Liu Y, Rong X. 2013. Ensemble data assimilation in a simple coupled climate model: The role of ocean-atmosphere interaction. *Advances in Atmospheric Sciences* **30**(5): 1235–1248.

Lorenz EN. 1996. Predictability: A problem partly solved. In: *Proc. Seminar on predictability*, vol. 1.

Lu F, Liu Z, Zhang S, Liu Y. 2015. Strongly coupled data assimilation using leading averaged coupled covariance (lacc). Part I: Simple model study. *Monthly Weather Review* **143**(9): 3823–3837.

Lü H, Yu Z, Zhu Y, Drake S, Hao Z, Sudicky EA. 2011. Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended kalman filter data assimilation. *Advances in water resources* **34**(3): 395–406.

Luo X, Bhakta T, Nævdal G, *et al.* 2017. Data driven adaptive localization with applications to ensemble-based 4d seismic history matching. In: *SPE Bergen One Day Seminar*. Society of Petroleum Engineers.

Luo X, Hoteit I. 2014. Ensemble kalman filtering with a divided state-space strategy for coupled data assimilation problems. *Monthly Weather Review* **142**(12): 4542–4558.

Moradkhani H, Sorooshian S, Gupta HV, Houser PR. 2005. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources* **28**(2): 135–147.

Palmer TN, Shutts GJ, Hagedorn R, Doblas-Reyes FJ, Jung T, Leutbecher M. 2005. Representing model uncertainty in weather and climate prediction. *Annu. Rev. Earth Planet. Sci.* **33**: 163–193.

Penny SG, Hamill TM. 2017. Coupled data assimilation for integrated earth system analysis and prediction. *Bulletin of the American Meteorological Society* **97**(7): ES169–ES172.

Raboudi NF, Ait-El-Fquih B, Hoteit I. 2018. Ensemble kalman filtering with one-step-ahead smoothing. *Monthly Weather Review* **146**(2): 561–581.

Saha S, Moorthi S, Pan HL, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, *et al.* 2010. The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society* **91**(8): 1015–1058.

Sakaguchi K, Zeng X, Brunke MA. 2012. The hindcast skill of the CMIP ensembles for the surface air temperature trend. *Journal of Geophysical Research: Atmospheres* **117**(D16).

Sakov P, Bertino L. 2011. Relation between two common localisation methods for the EnKF. *Computational Geosciences* **15**(2): 225–237.

Sampson K, Gochis D. 2018. Wrf hydro gis pre-processing tools, version 5.0 documentation. *Boulder, CO: National Center for Atmospheric Research, Research Applications Laboratory* .

Subramanian AC, Hoteit I, Cornuelle B, Miller AJ, Song H. 2012. Linear versus nonlinear filtering with scale-selective corrections for balanced dynamics in a simple atmospheric model. *Journal of the Atmospheric Sciences* **69**(11): 3405–3419.

Talagrand O. 1999. Evaluation of probabilistic prediction systems. In: *Workshop proceedings" Workshop on predictability", 20-22 October 1997, ECMWF, Reading, UK*.

Tippett MK, Anderson JL, Bishop CH, Hamill TM, Whitaker JS. 2003. Ensemble square root filters. *Monthly Weather Review* **131**: 1485 – 1490.

Travers M, Shin YJ, Jennings S, Cury P. 2007. Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. *Progress in oceanography* **75**(4): 751–770.

Triantafyllou G, Korres G, Hoteit I, Petihakis G, Banks A. 2006. Assimilation of ocean colour data into a biochemical flux model of the eastern

mediterranean sea .

Wen XH, Chen WH, *et al.* 2005a. Real-time reservoir model updating using ensemble kalman filter. In: *SPE reservoir simulation symposium*. Society of Petroleum Engineers.

Wen XH, Chen WH, *et al.* 2005b. Some practical issues on real-time reservoir model updating using ensemble kalman filter. In: *International Petroleum Technology Conference*. International Petroleum Technology Conference.

Zhang S, Harrison MJ, Rosati A, Wittenberg A. 2007. System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Monthly Weather Review* **135**(10): 3541–3564.

panski M. 2017. Data assimilation for coupled modeling systems. In: *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III)*, Springer, pp. 55–70.

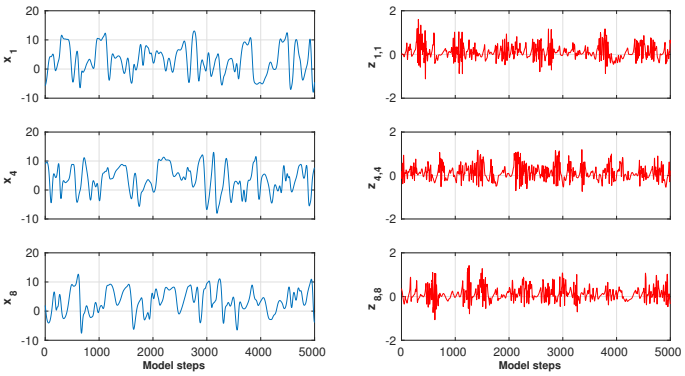


Figure 1. Time series for three components ($\mathbf{x}_1$, $\mathbf{x}_4$ and $\mathbf{x}_8$) of the slow state and three components ($\mathbf{z}_{1,1}$, $\mathbf{z}_{4,4}$ and $\mathbf{z}_{8,8}$) of the fast state for the first 5000 model steps.

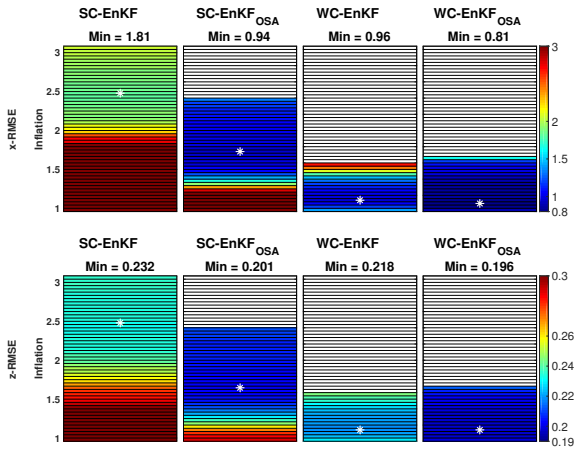


Figure 2. Time-averaged RMSEs as a function of inflation factor (y axis), as they result from SC-EnKF, SC-EnKF$_{OSA}$, WC-EnKF and WC-EnKF$_{OSA}$. All the filters were implemented with 20 members, and half of the observations of both model components were assimilated every 40 ms. The minimum-averaged RMSEs are indicated by asterisks and their associated values are given in the title. White lines indicate divergence of the filter.

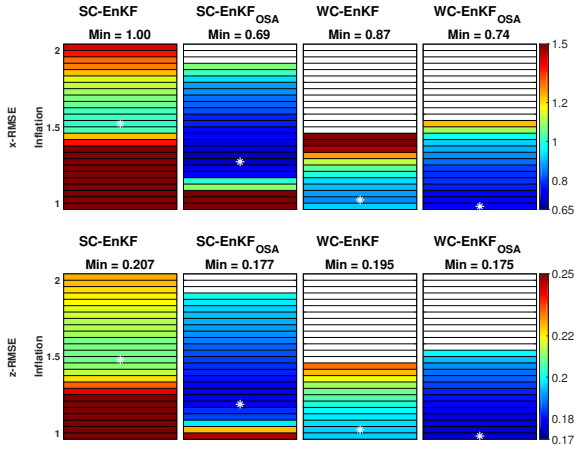


Figure 3. Same as Fig. 2, but for 40 ensemble members.

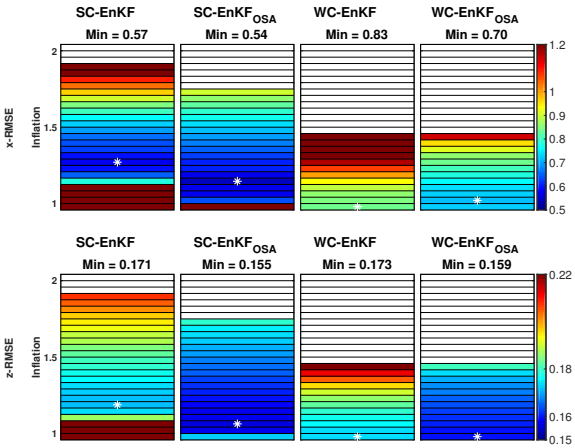


Figure 4. Same as Fig. 2, but for 80 ensemble members.

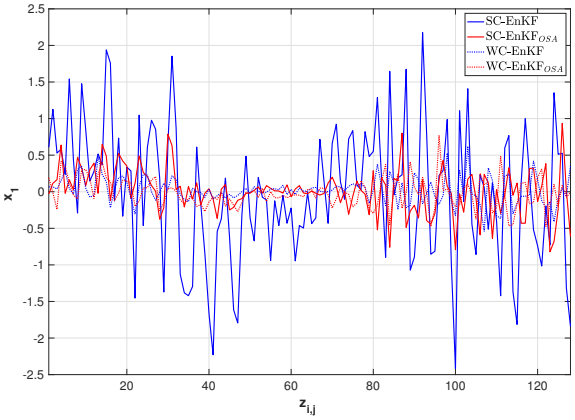


Figure 5. Plots of the ensemble-based cross-covariances between $\mathbf{x}_1$ and all $\mathbf{z}_{i,j}$ as estimated by the SC-EnKF, SC-EnKF$_{OSA}$, WC-EnKF and WC-EnKF$_{OSA}$ at the 1000^{th} assimilation step using 40 members. Half of the observations from both model components were assimilated every 40 ms. Results are presented using the inflation factor that provides the lowest joint RMSE for each filter.

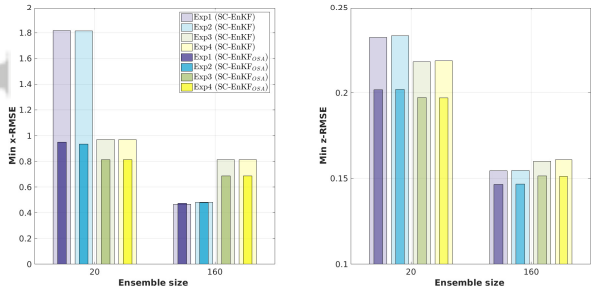


Figure 6. Bar plots of the minimum x-RMSE and z-RMSE as resulting from SC-EnKF and SC-EnKF$_{OSA}$ in four different scenarios that differ in the cross-observations assimilation strategies. Half of the observations from both model components were assimilated every 40 ms.

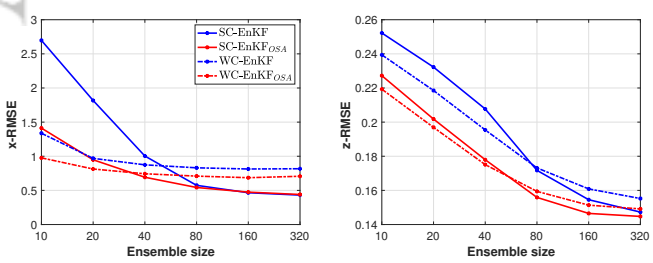


Figure 7. Minimum averaged x-RMSE and z-RMSE as a function of ensemble size and as estimated by SC-EnKF, SC-EnKF$_{OSA}$, WC-EnKF and WC-EnKF$_{OSA}$. Half of the observations from both model components were assimilated every 40 ms.

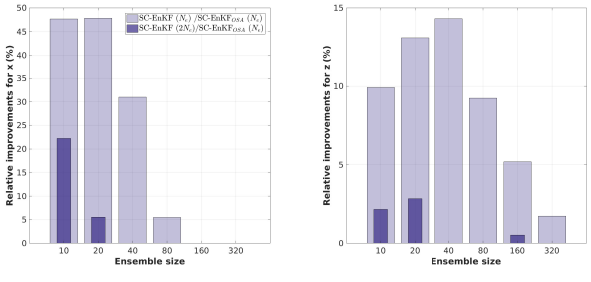


Figure 8. Bar plots of relative improvements introduced by SC-EnKF$_{OSA}$ compared to SC-EnKF as function of the ensemble size (only positive percentages are reported). Half of the observations from both model components were assimilated every 40 ms.

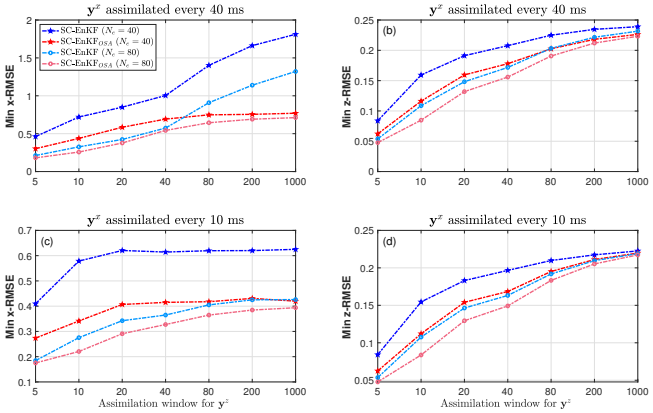


Figure 9. Minimum averaged x-RMSE and z-RMSE as a function of the assimilation window of $\mathbf{y}^z$.

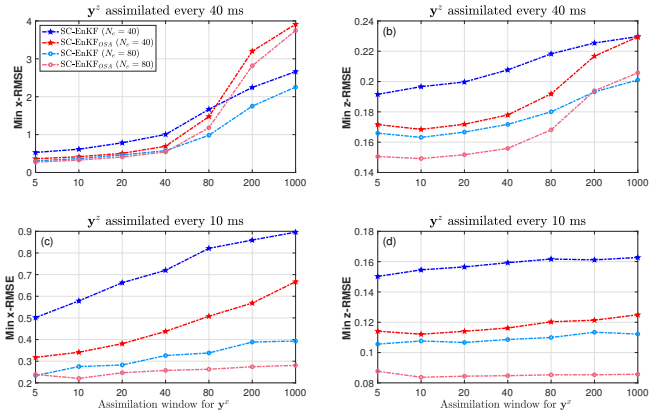


Figure 10. Minimum averaged x-RMSE and z-RMSE as a function of the assimilation window of $\mathbf{y}^x$.

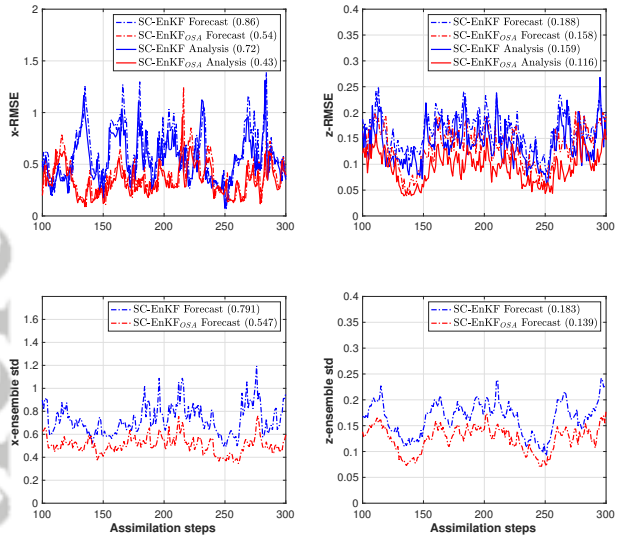


Figure 11. Time evolution of the forecast and analysis RMSEs (upper panels) and the ensemble spread (lower panels) as they result from SC-EnKF and SC-EnKF$_{OSA}$ between the 100^{th} and the 300^{th} assimilation steps. The filters were implemented with 40 members and half of the observations from both model components were assimilated every 40 ms for $\mathbf{y}^x$ and every 10 ms for $\mathbf{y}^z$. Results are presented using the inflation factor that provides the lowest joint RMSE for each filter.

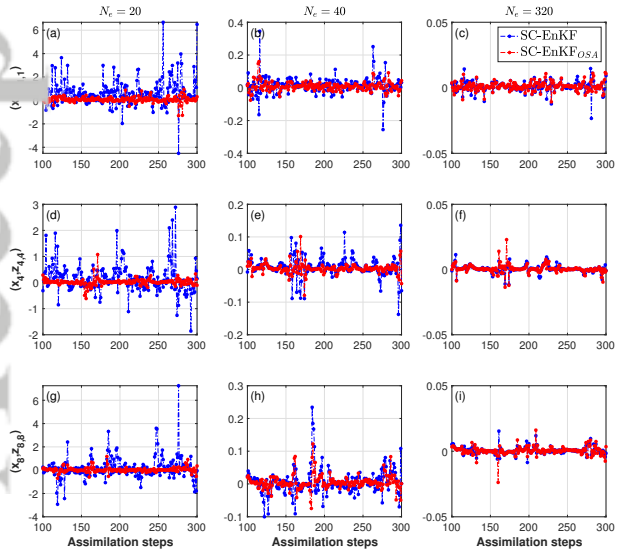


Figure 12. Plots of the ensemble-based cross-covariances between $\mathbf{x}_1$ (upper panels), $\mathbf{x}_4$ (middle panels), $\mathbf{x}_8$ (lower panels) and all $\mathbf{z}_{i,j}$ as estimated by the SC-EnKF and SC-EnKF$_{OSA}$ using 20 and 40 members (between the 100^{th} and the 300^{th} assimilation steps). Half of the observations from both model components were assimilated every 40 ms for $\mathbf{y}^x$ and every 10 ms for $\mathbf{y}^z$. Results are presented using the inflation factor that provides the lowest joint RMSE for each filter.

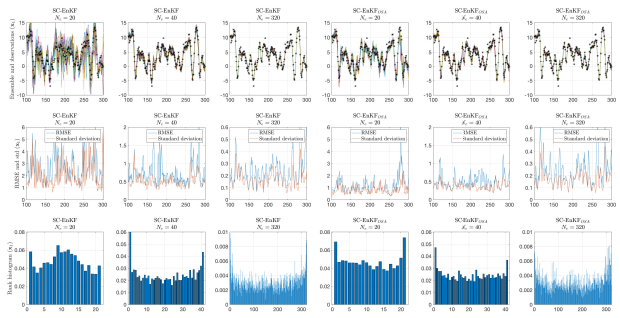


Figure 13. Plots of the analysis ensemble and observations (first row), analysis RMSE and ensemble standard deviation (second row) and rank histograms (third row), as estimated by the SC-EnKF and SC-EnKF$_{OSA}$ for $\mathbf{x}_1$ using 20, 40 and 320 members (between the 100^{th} and the 300^{th} assimilation steps). Half of the variables from both model components were assimilated every 40 ms for $\mathbf{y}^x$ and every 10 ms for $\mathbf{y}^z$. Results are presented using the inflation factor that provides the lowest joint RMSE for each filter.

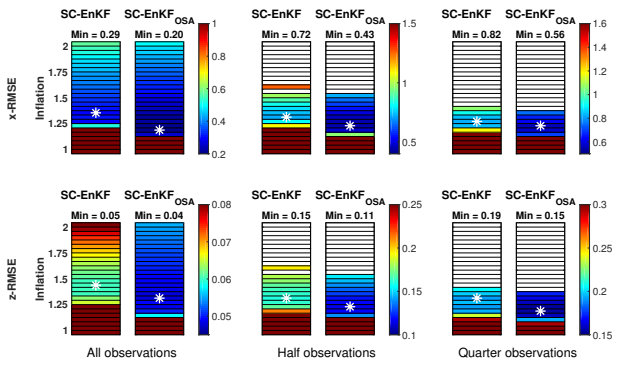


Figure 14. Time-averaged RMSEs, as a function of inflation factor (y axis), as they result from SC-EnKF and SC-EnKF$_{OSA}$, when all, half and one-quarter of the observations from each model component are assimilated. The filters were implemented with 40 members. $\mathbf{y}^x$ are assimilated every 40 ms and $\mathbf{y}^z$ every 10 ms.

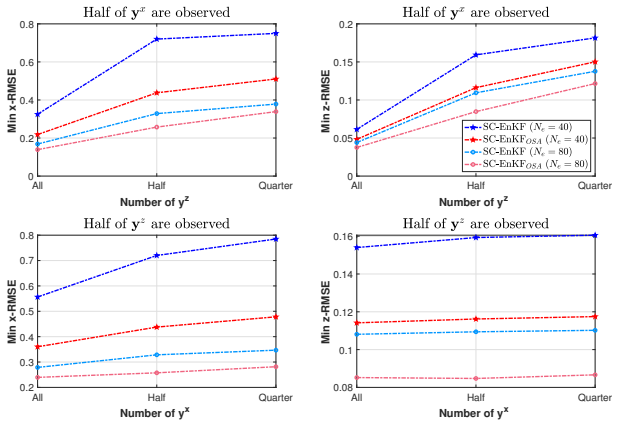


Figure 15. Minimum averaged x-RMSE and z-RMSE as a function of the number of assimilated $\mathbf{y}^z$ and $\mathbf{y}^x$.

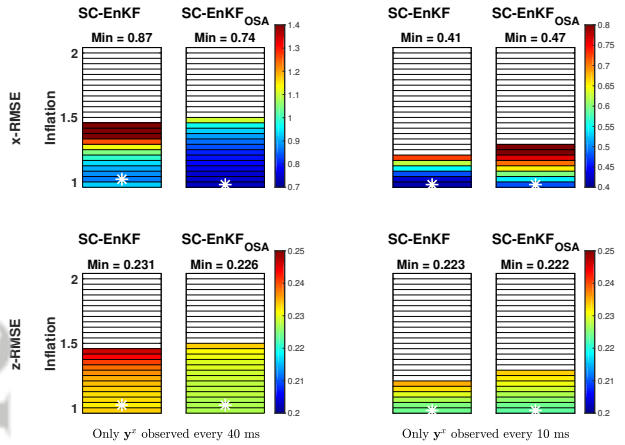


Figure 16. Time-averaged RMSEs, as a function of inflation factor (y axis), as they result from SC-EnKF and SC-EnKF$_{OSA}$. The filters were implemented with 40 members and only half of $\mathbf{y}^x$ are assimilated every 40 and 10 ms.

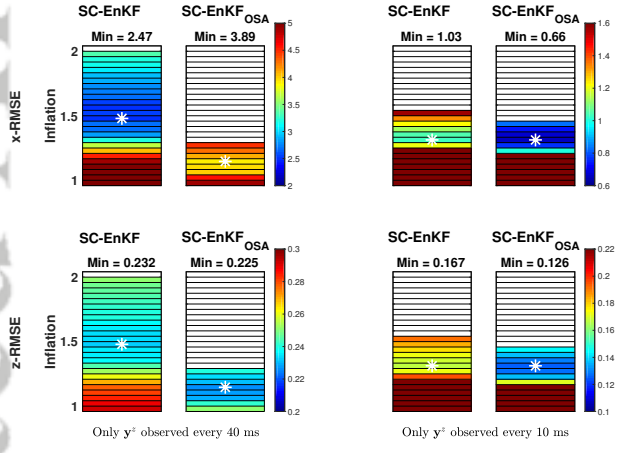


Figure 17. Time-averaged RMSEs, as a function of inflation factor (y axis), as they result from SC-EnKF and SC-EnKF$_{OSA}$. The filters were implemented with 40 members and only half of $\mathbf{y}^z$ are assimilated every 40 and 10 ms.

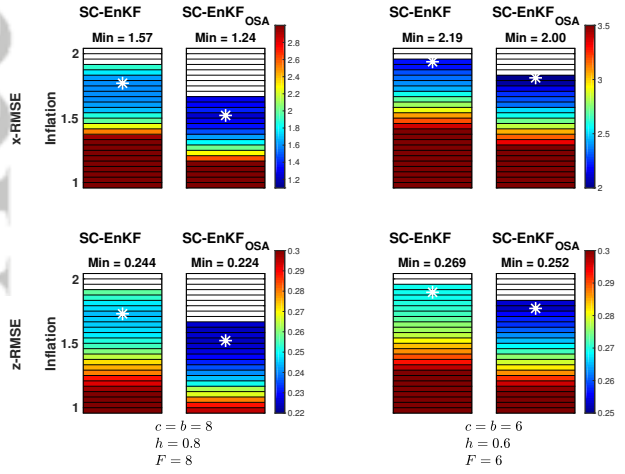
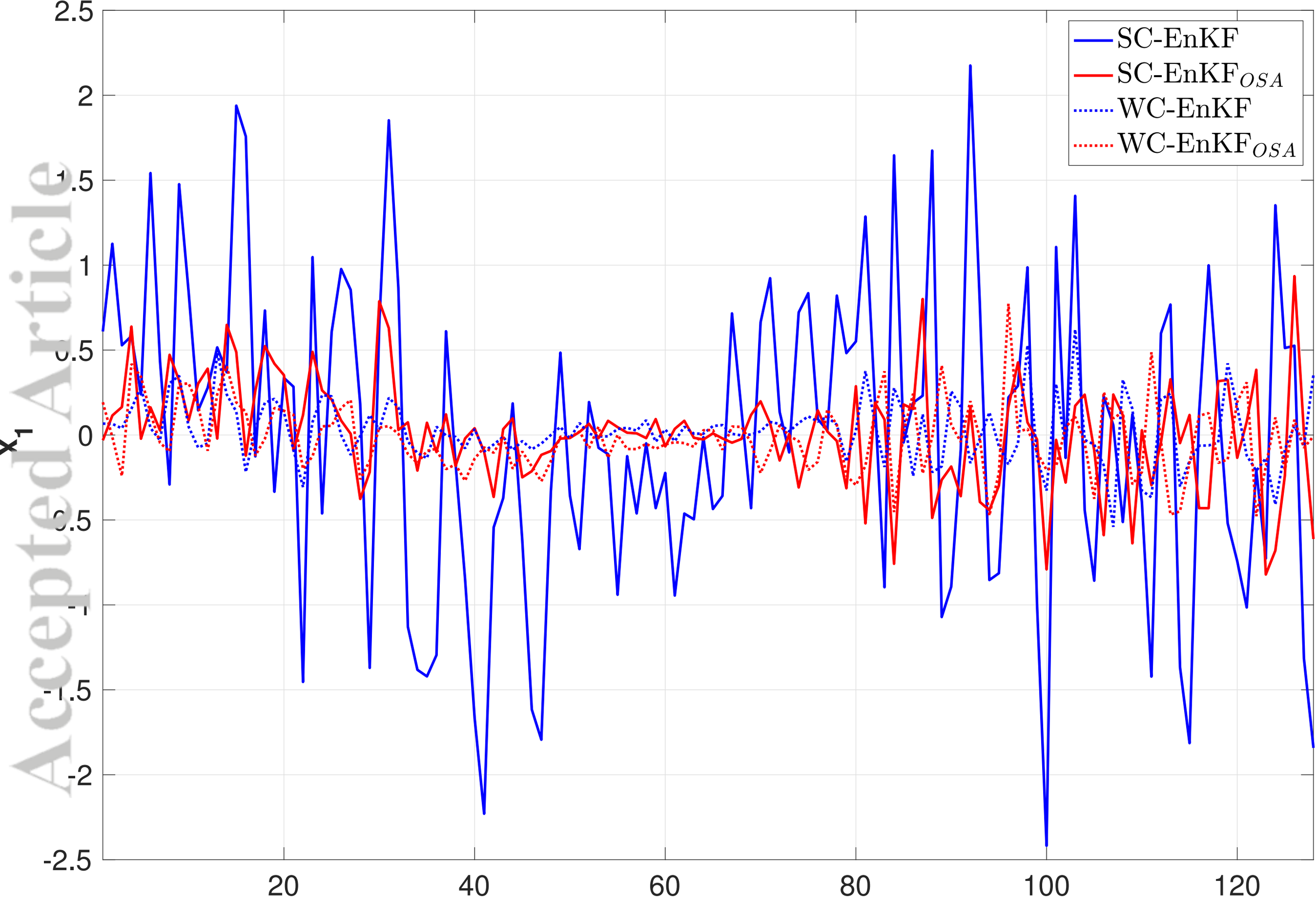
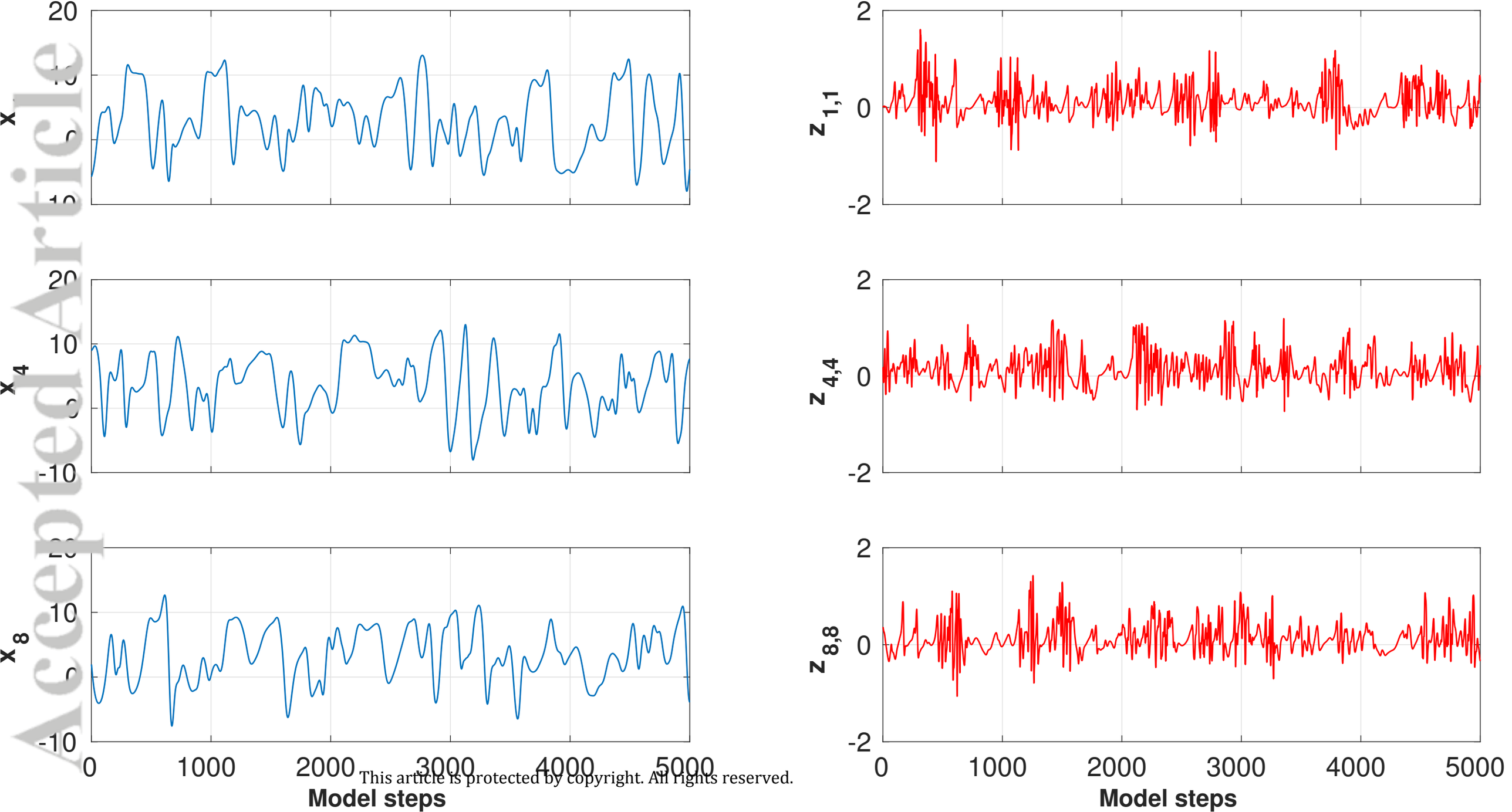
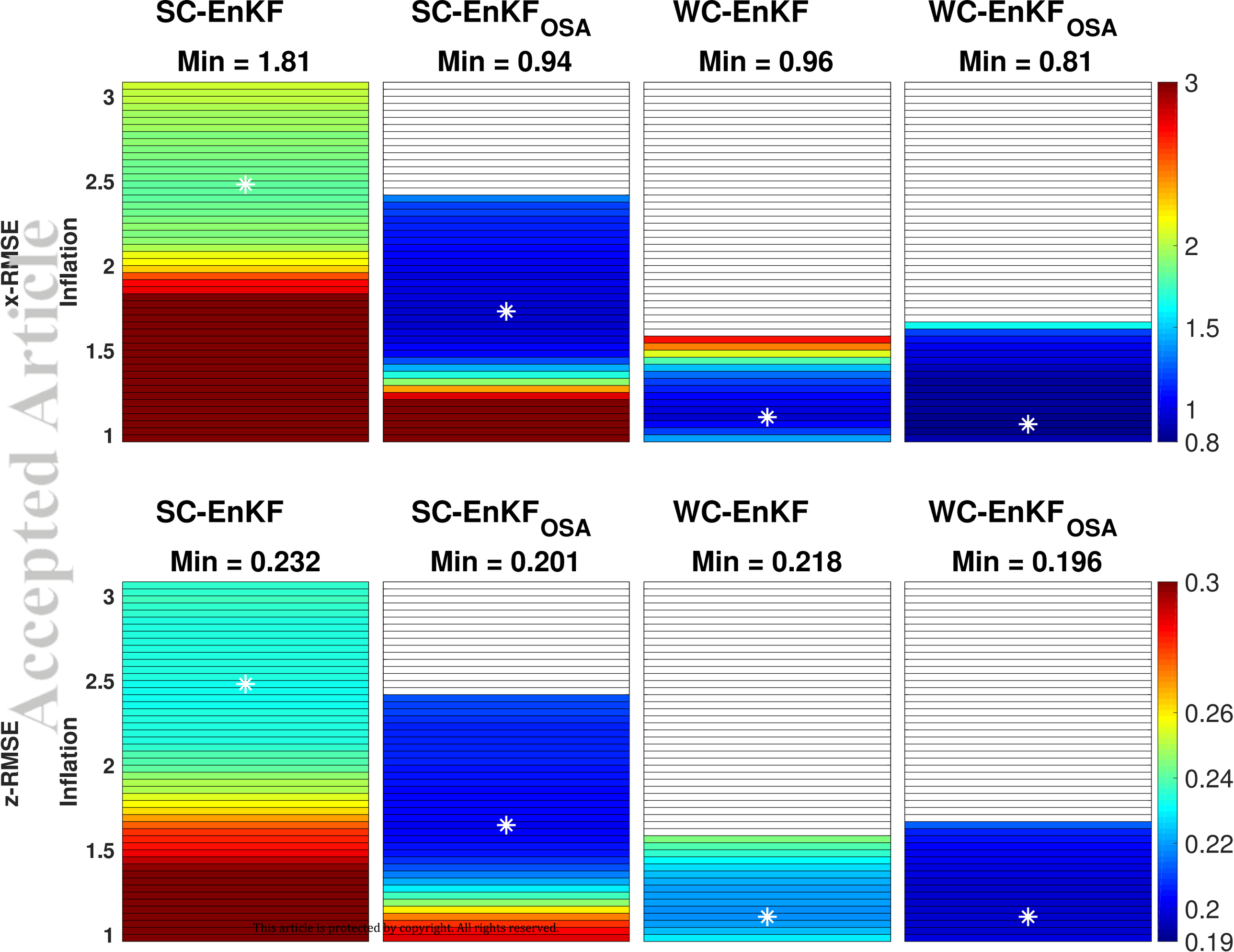
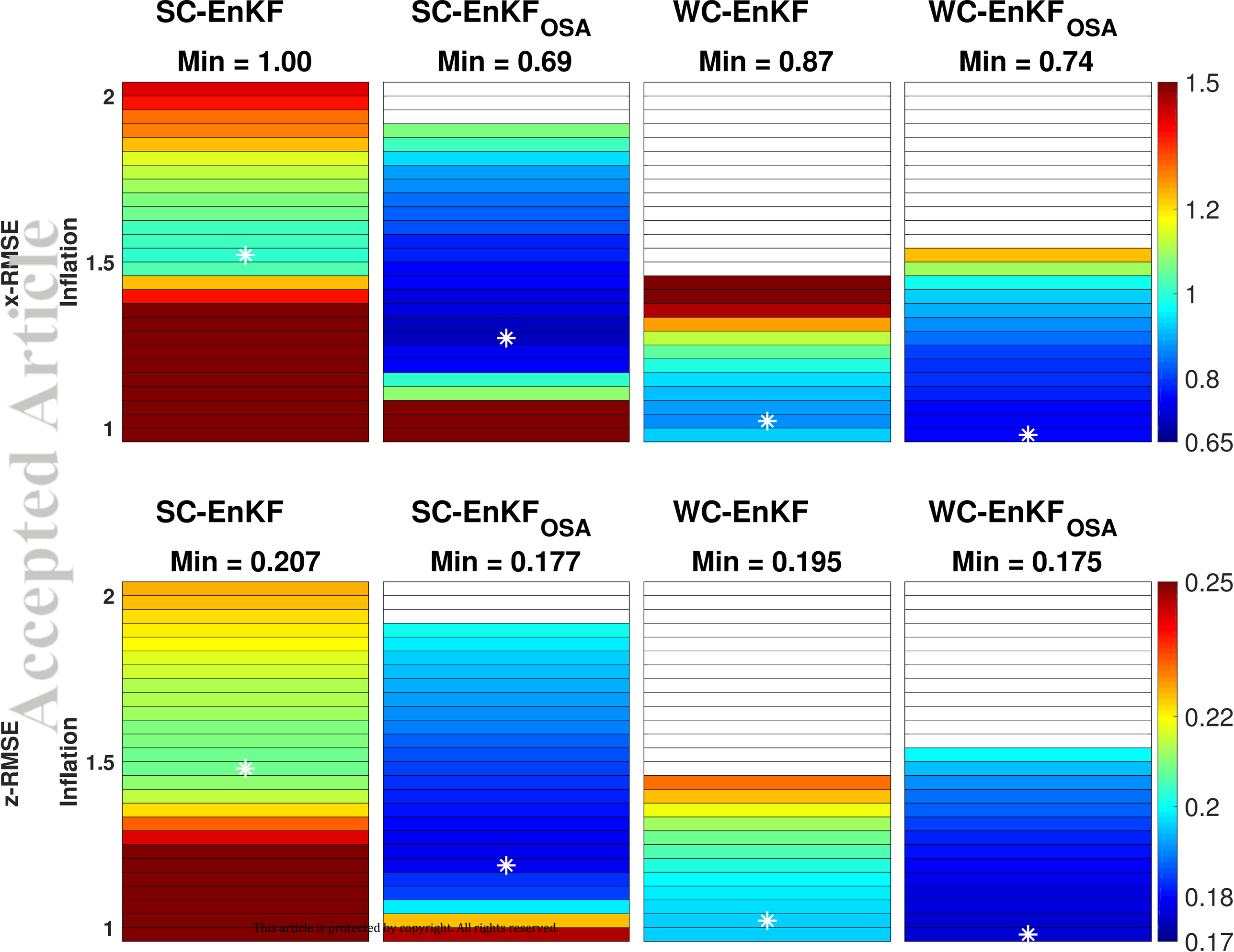


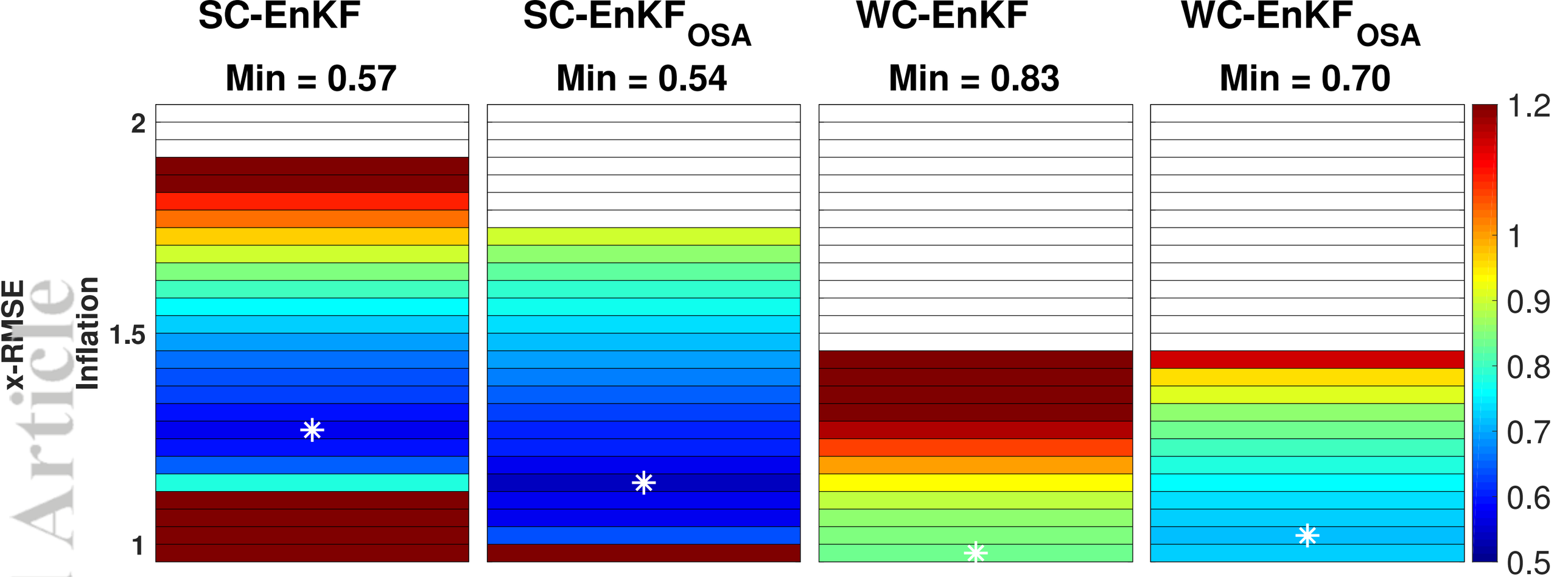
Figure 18. Time-averaged RMSEs, as a function of inflation factor (y axis), as they result from SC-EnKF and SC-EnKF$_{OSA}$ when assimilation is carried out with a biased forecast model. The filters were implemented with 40 and 80 members and quarter of the observations from both model components are assimilated every 40 ms.

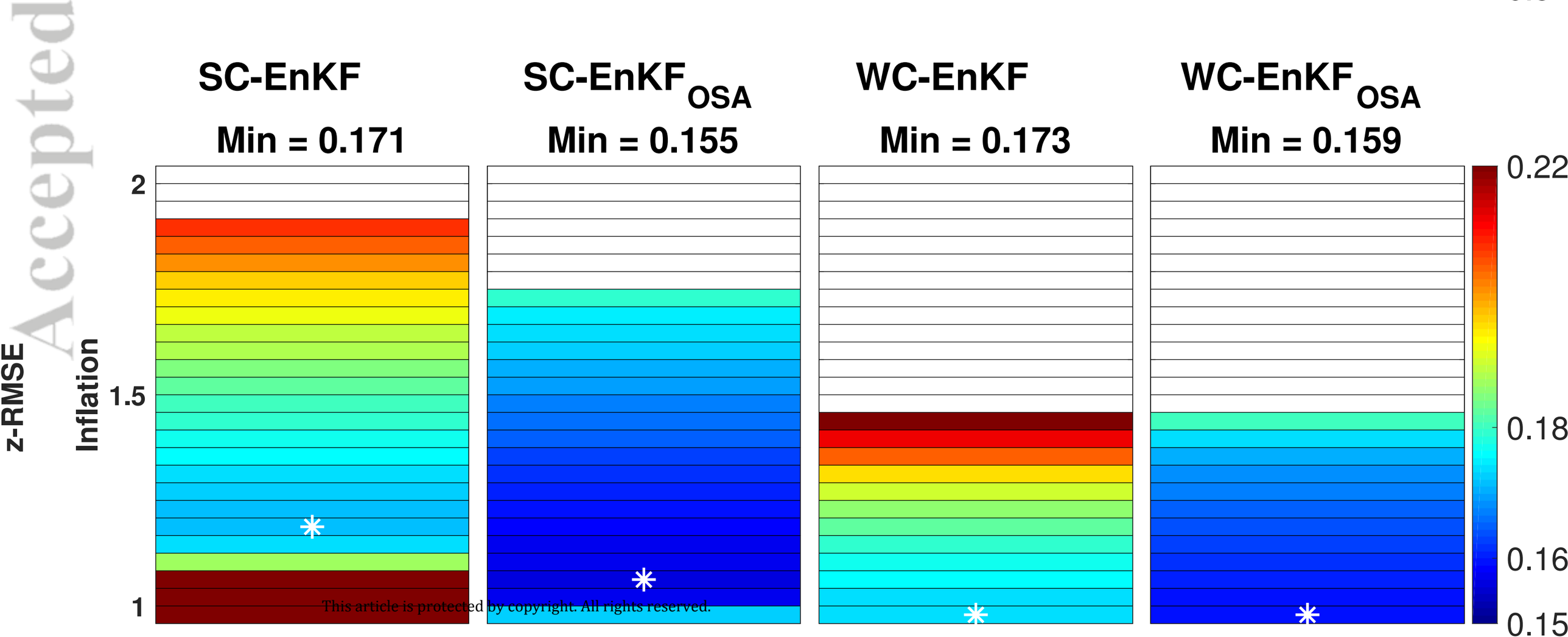


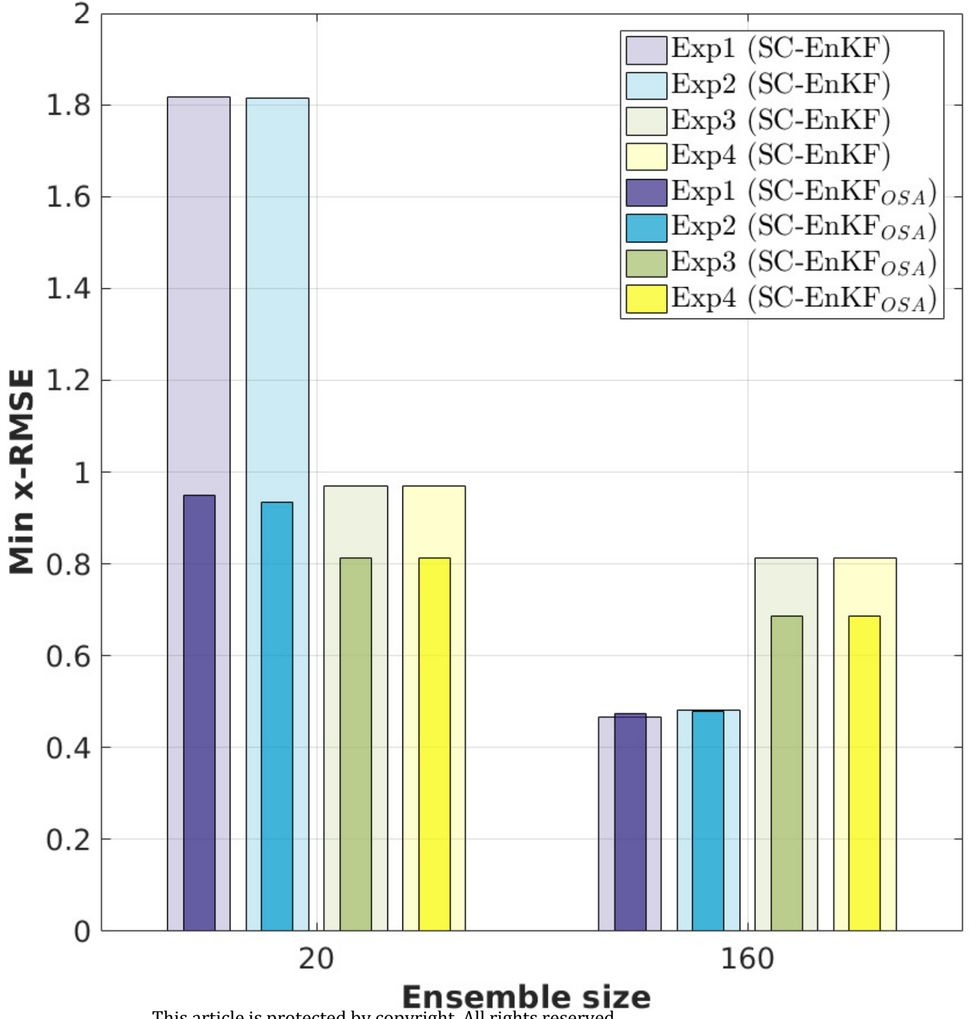


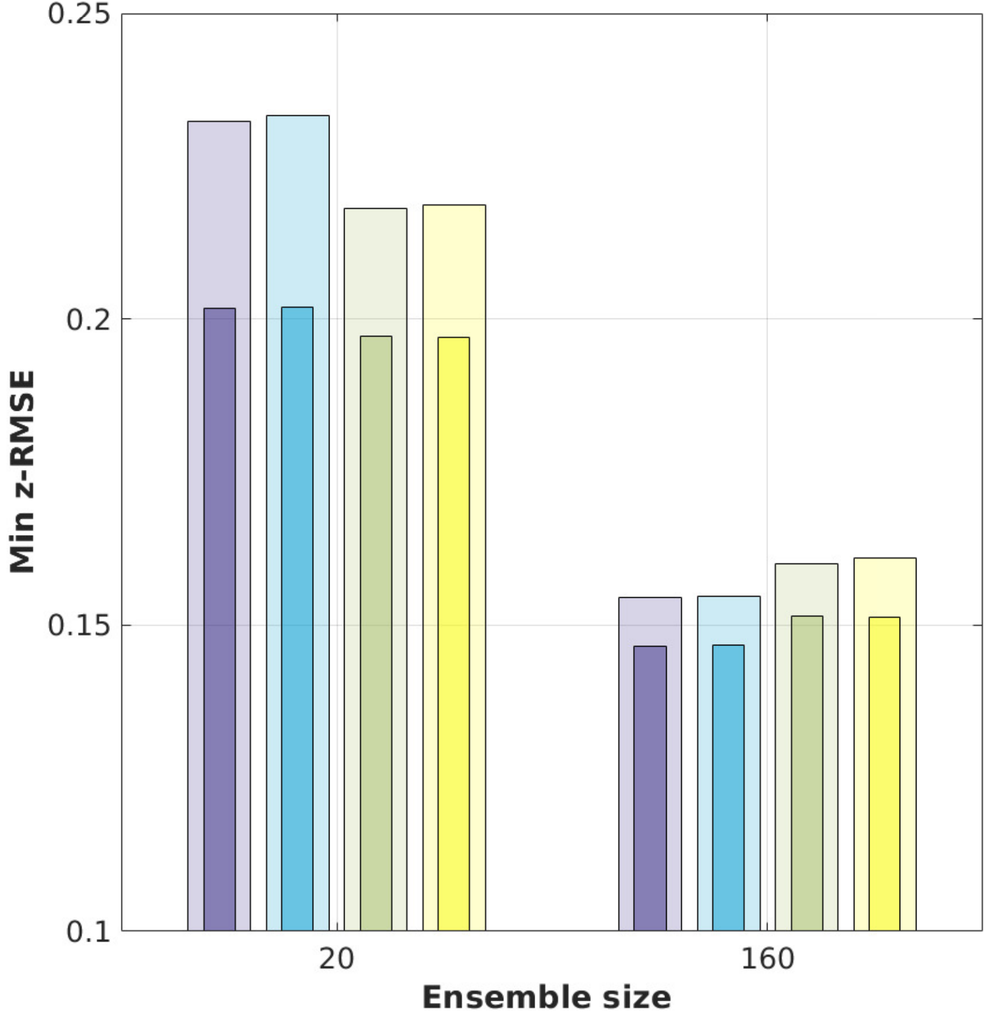


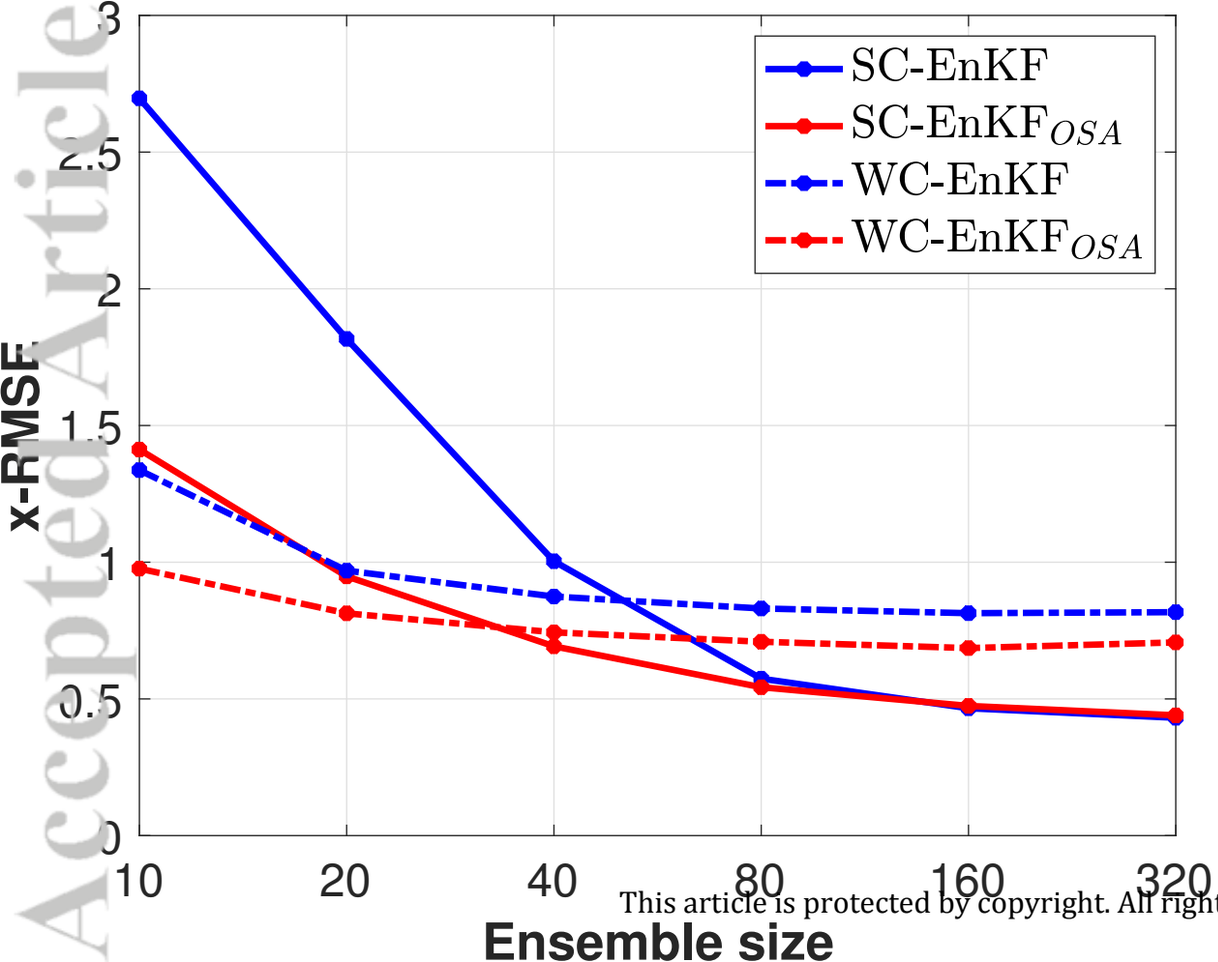


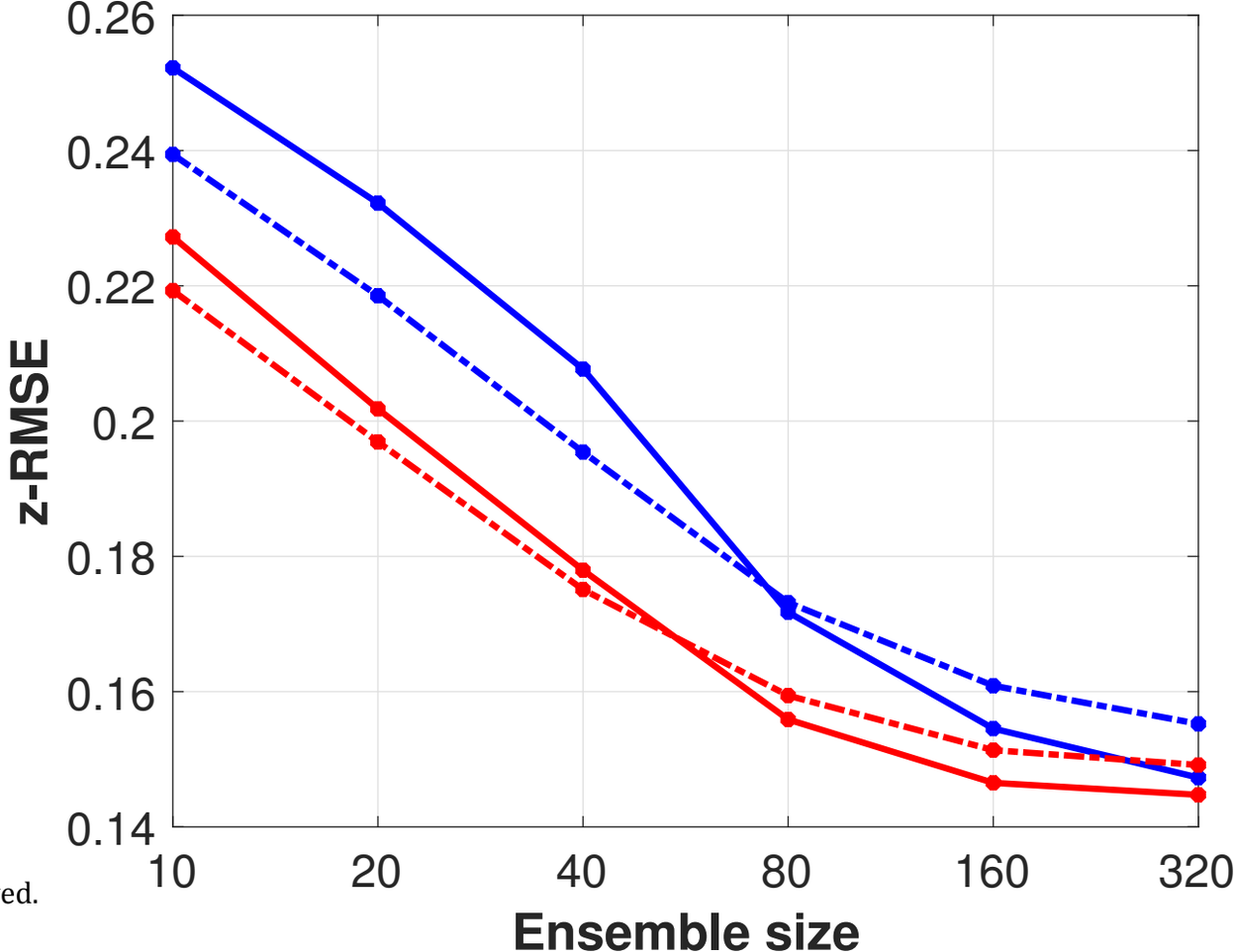


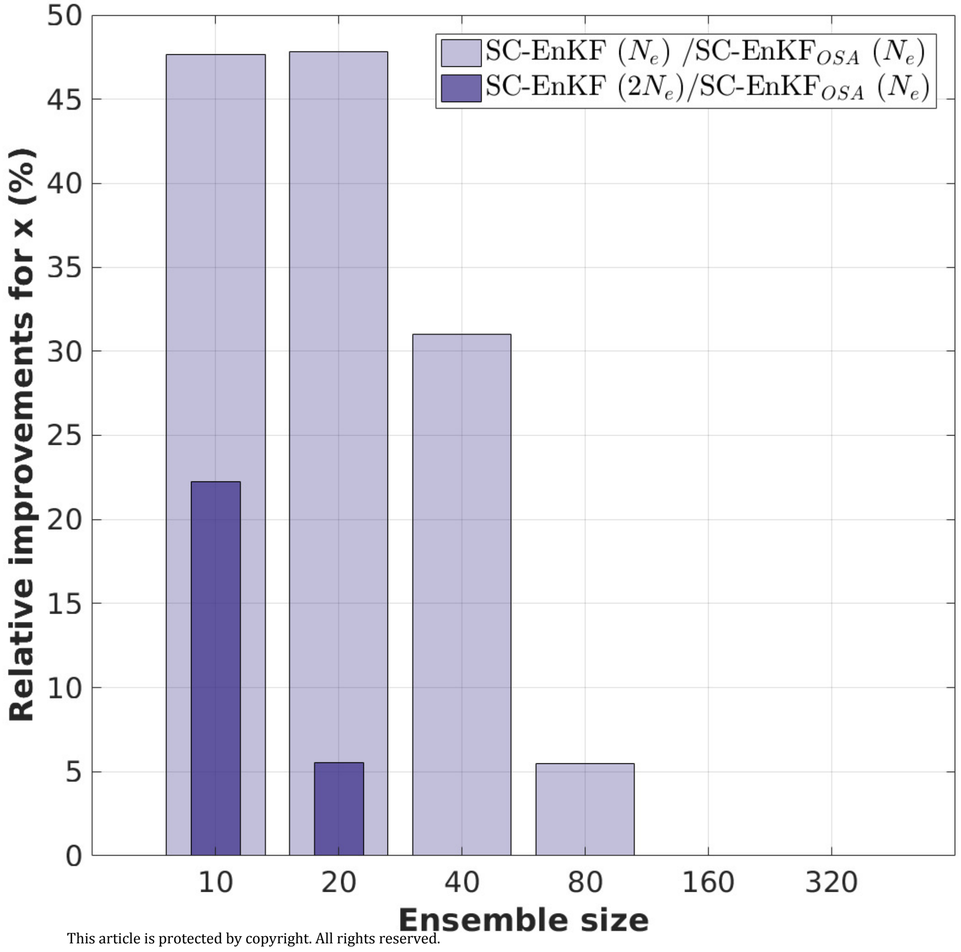


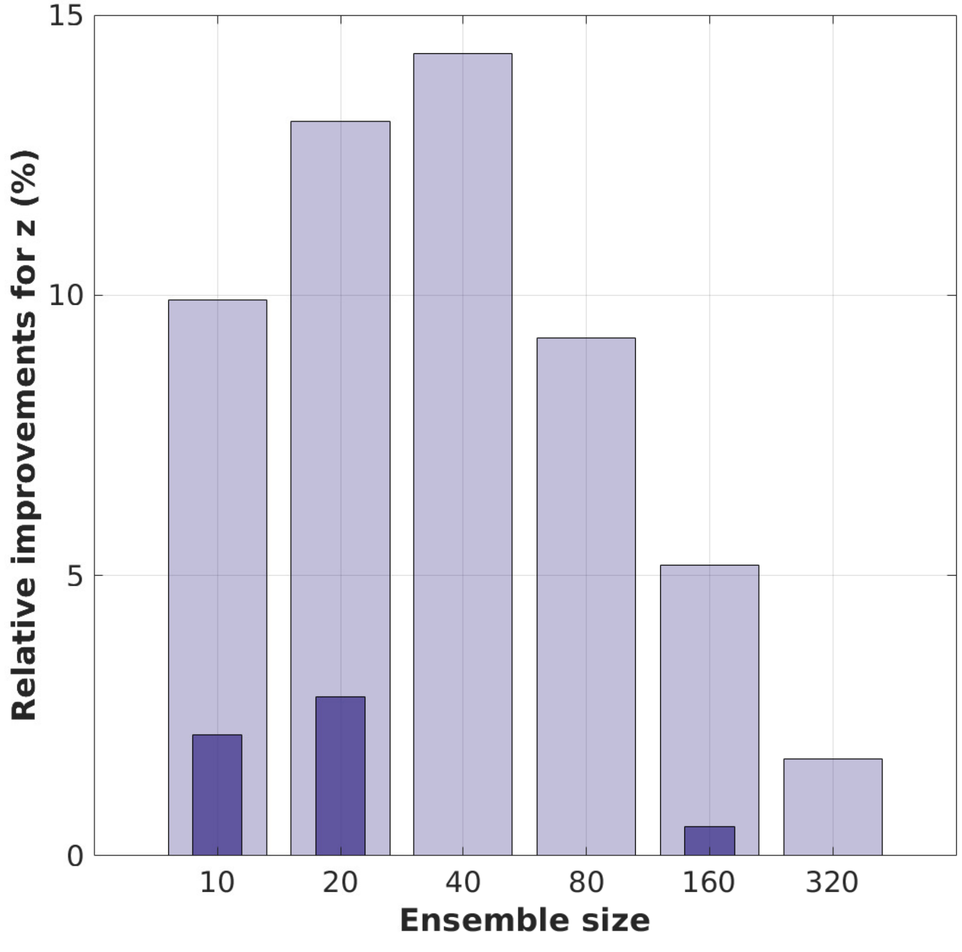


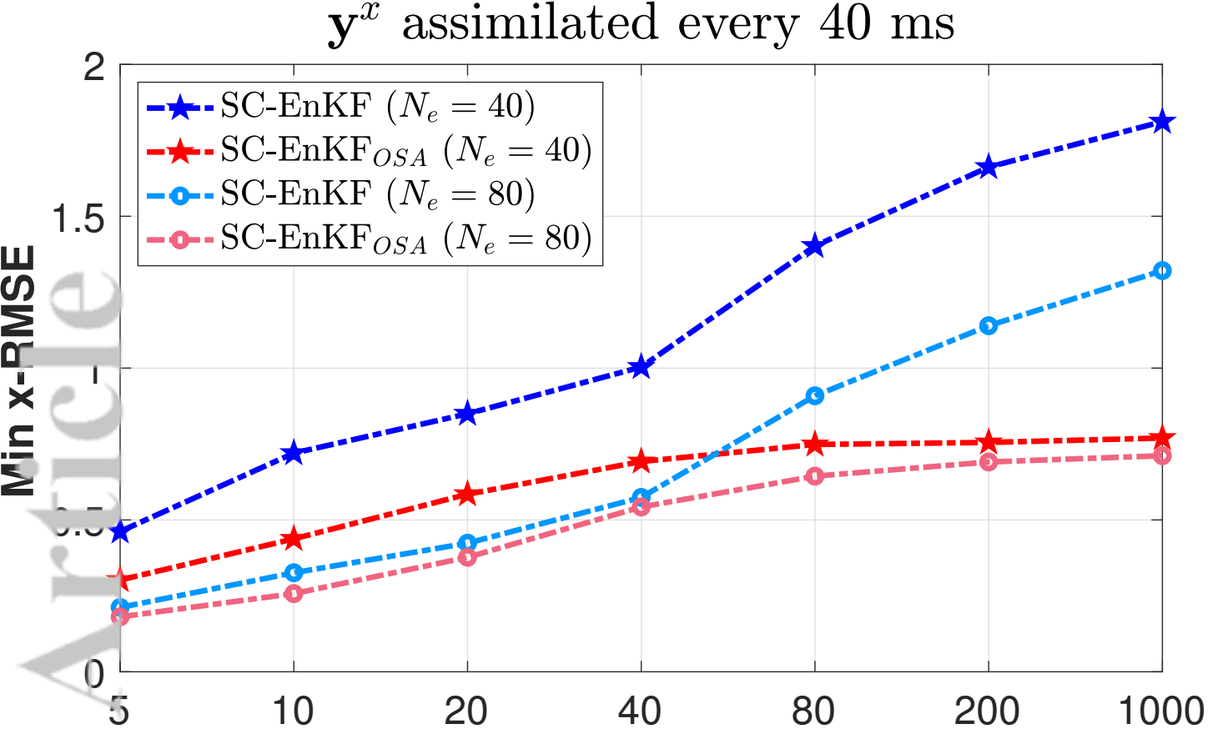


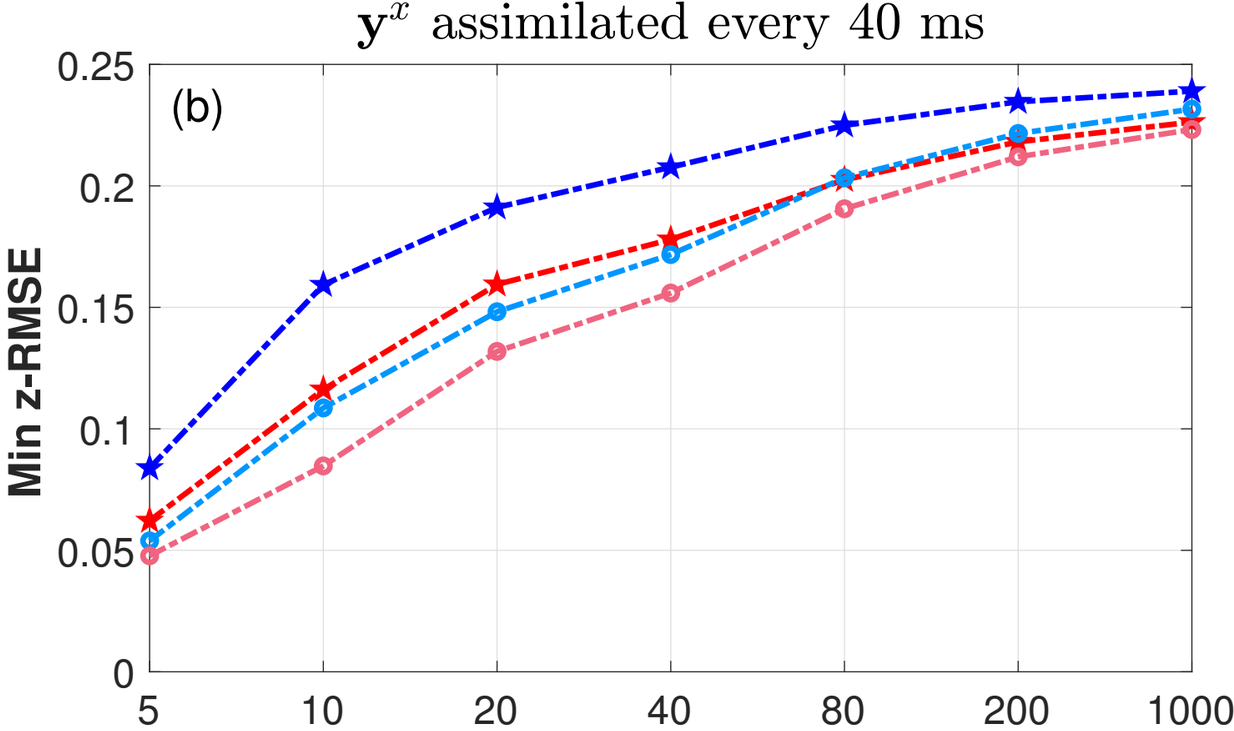


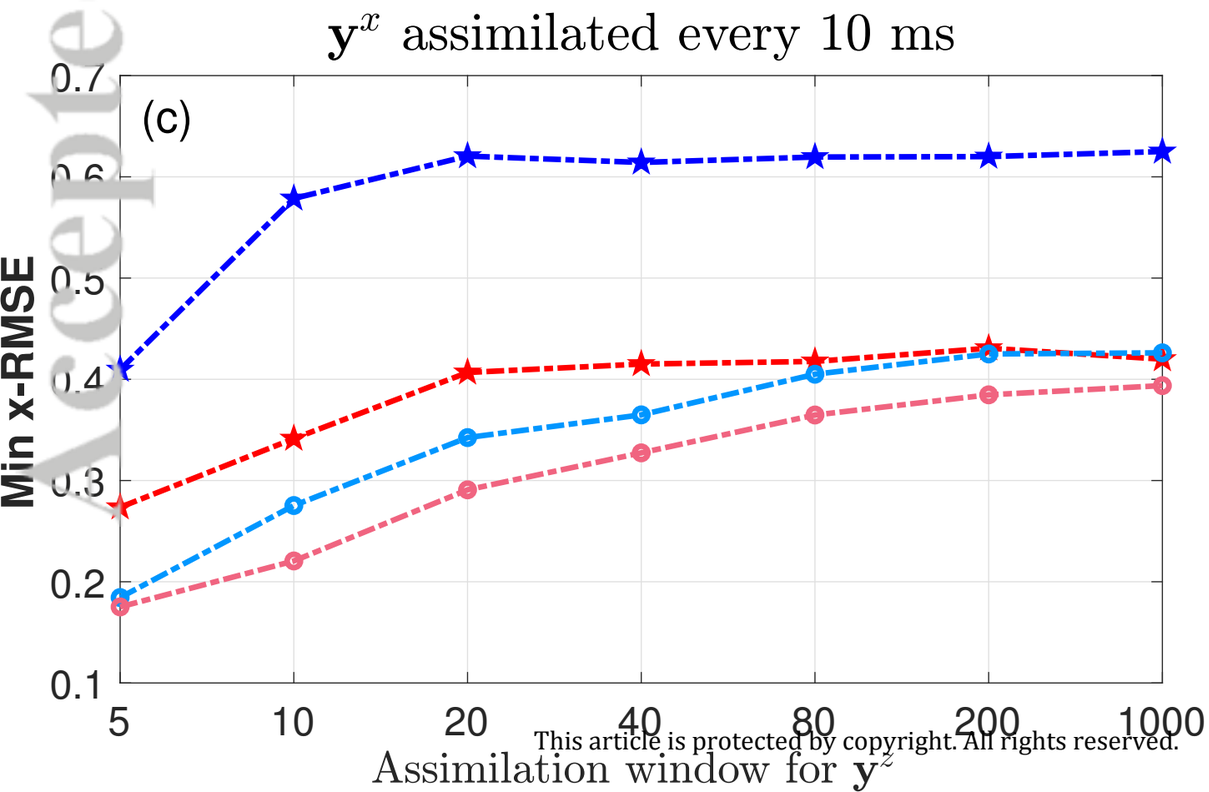


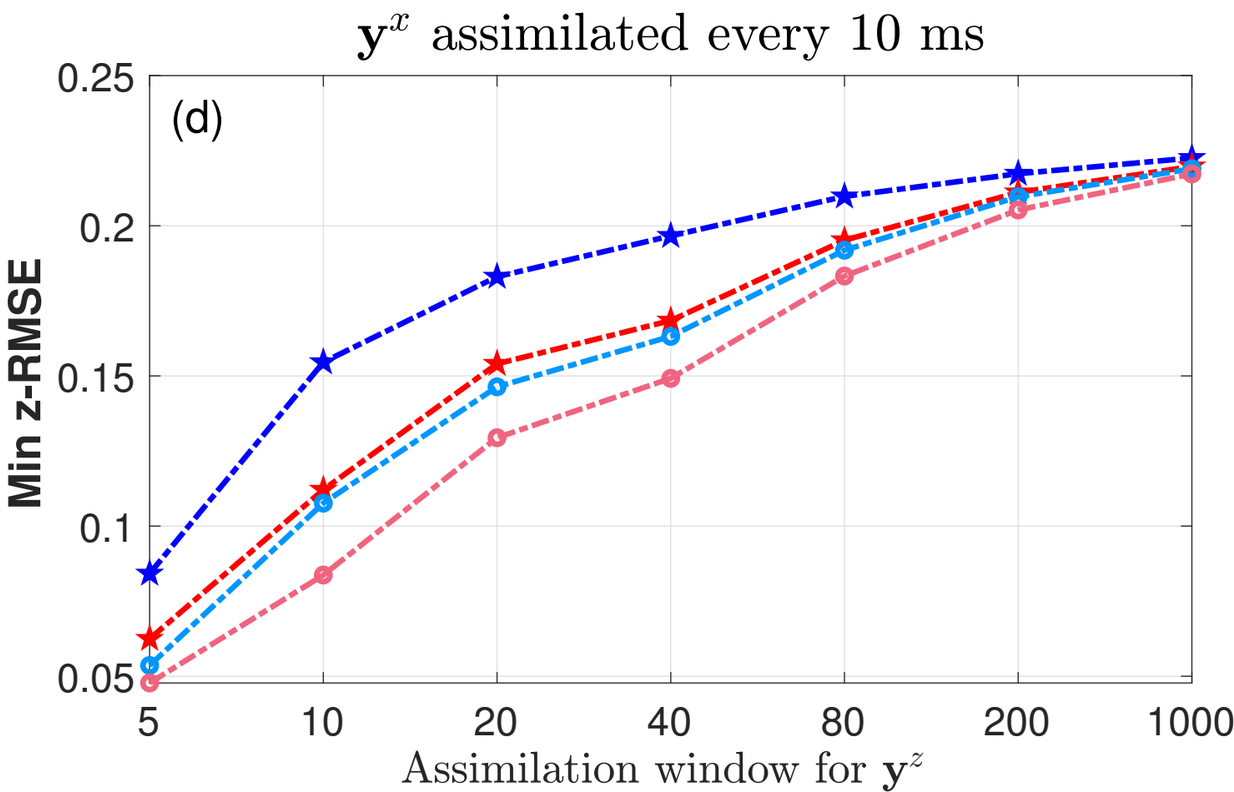


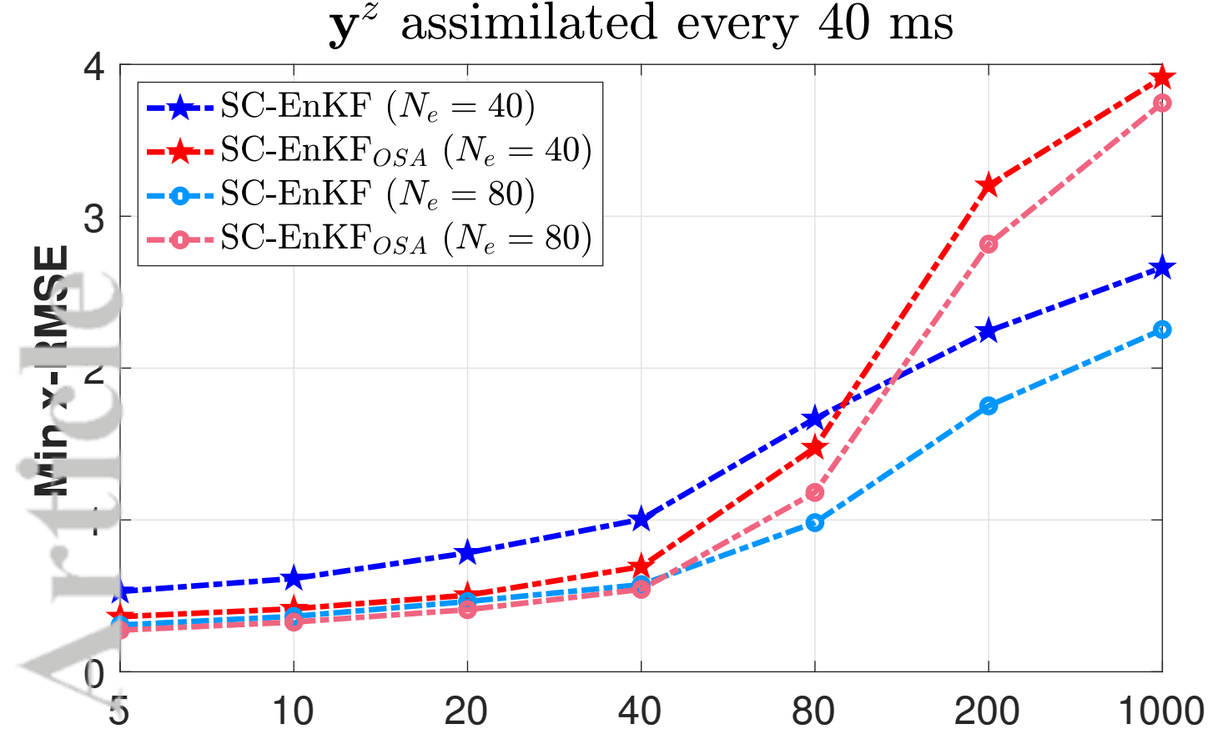


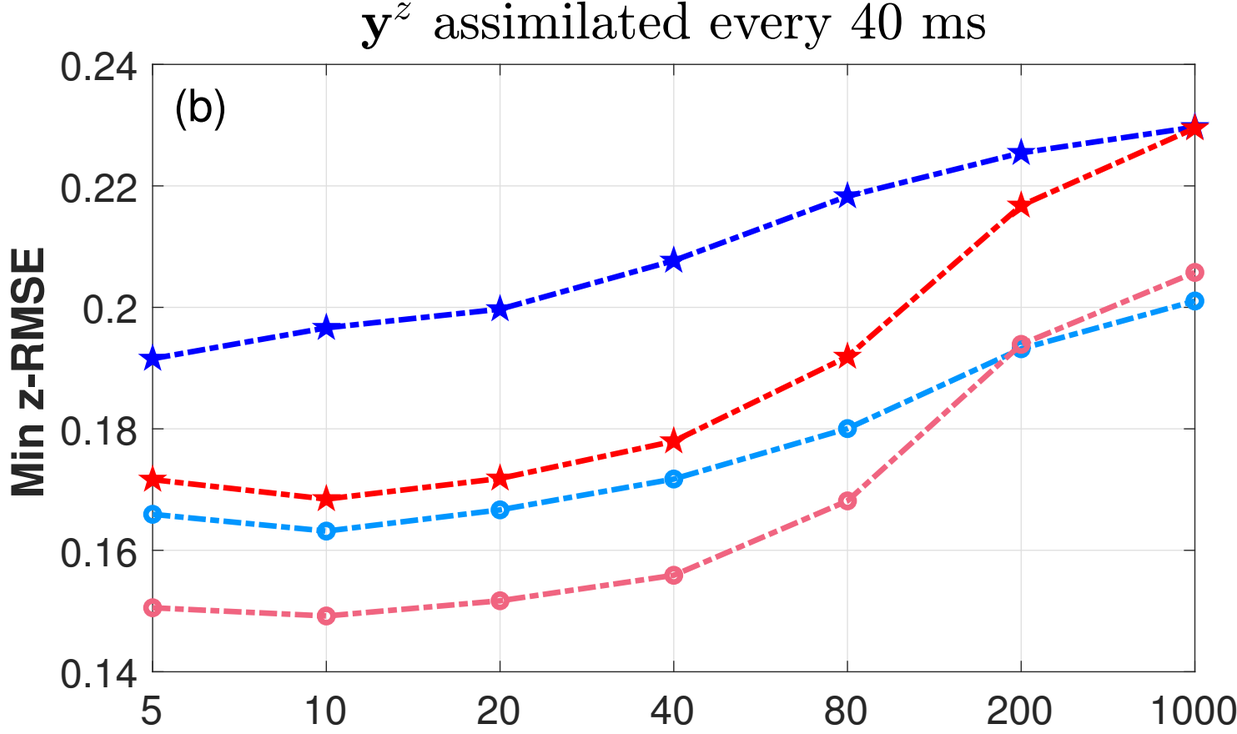


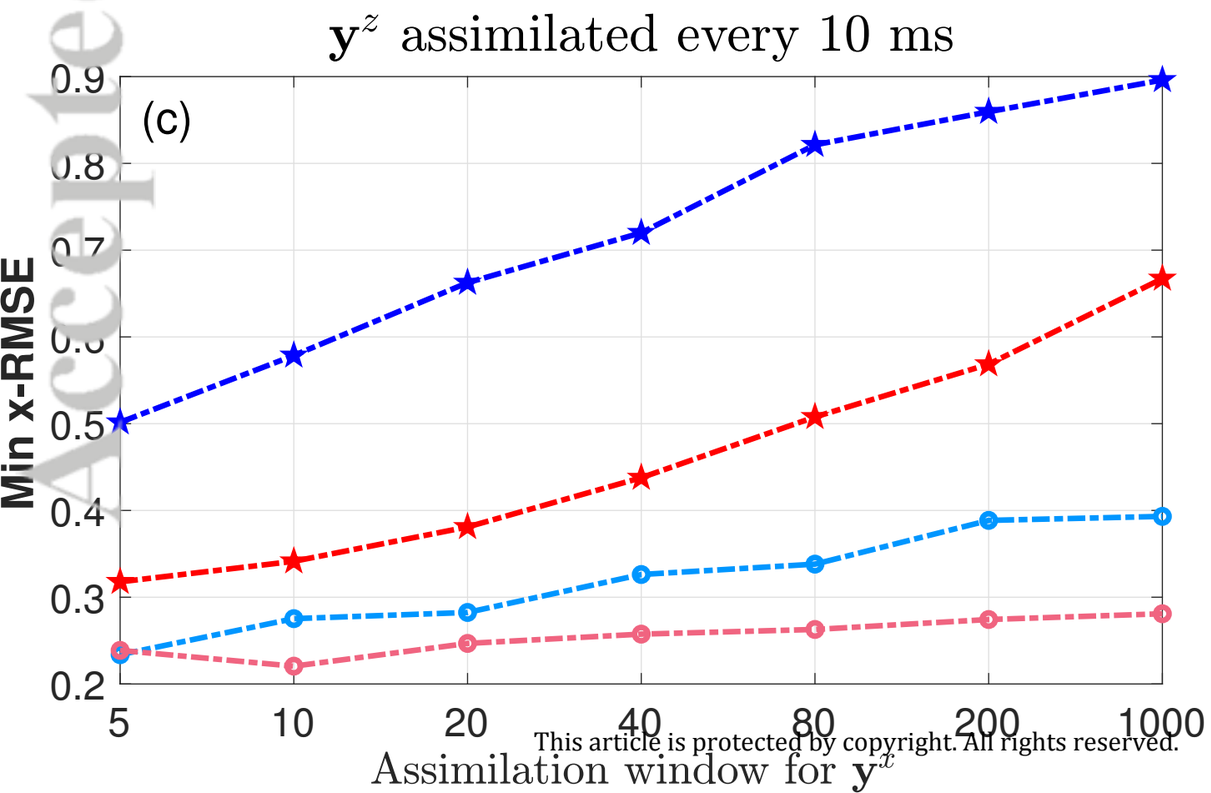


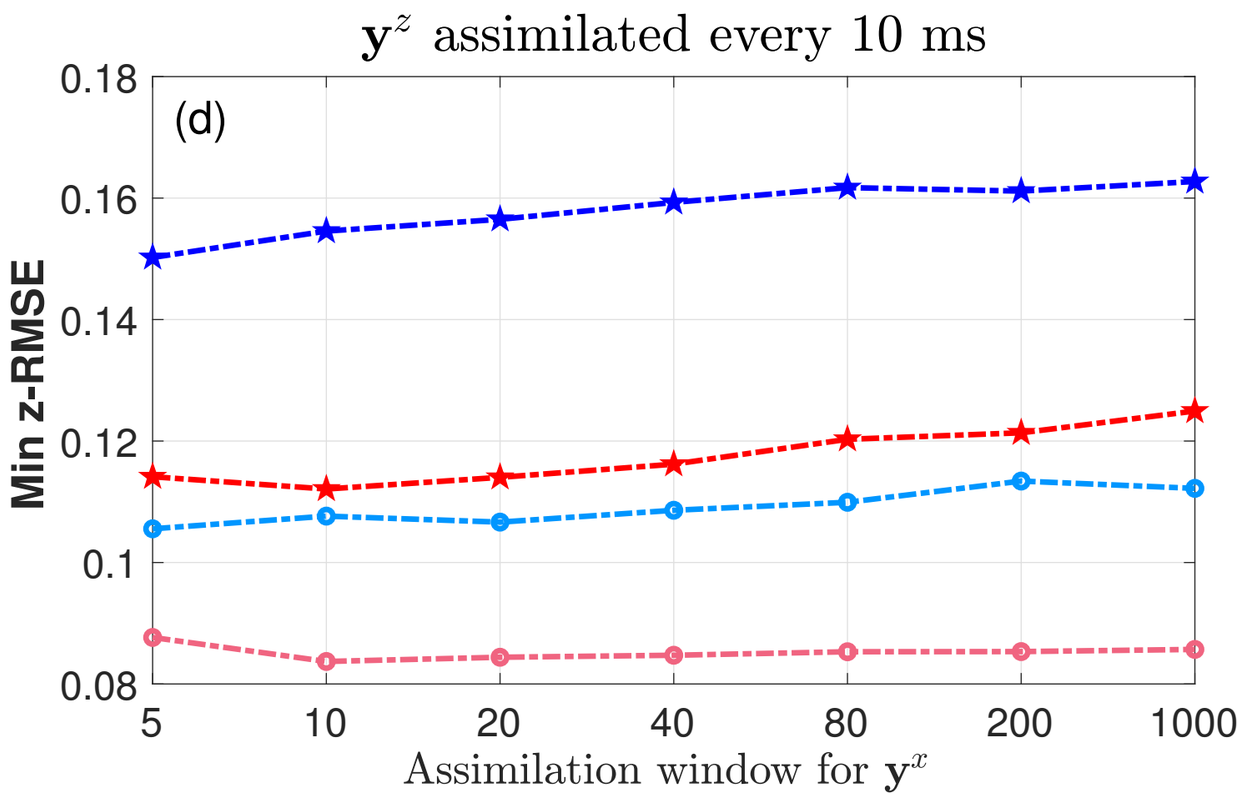


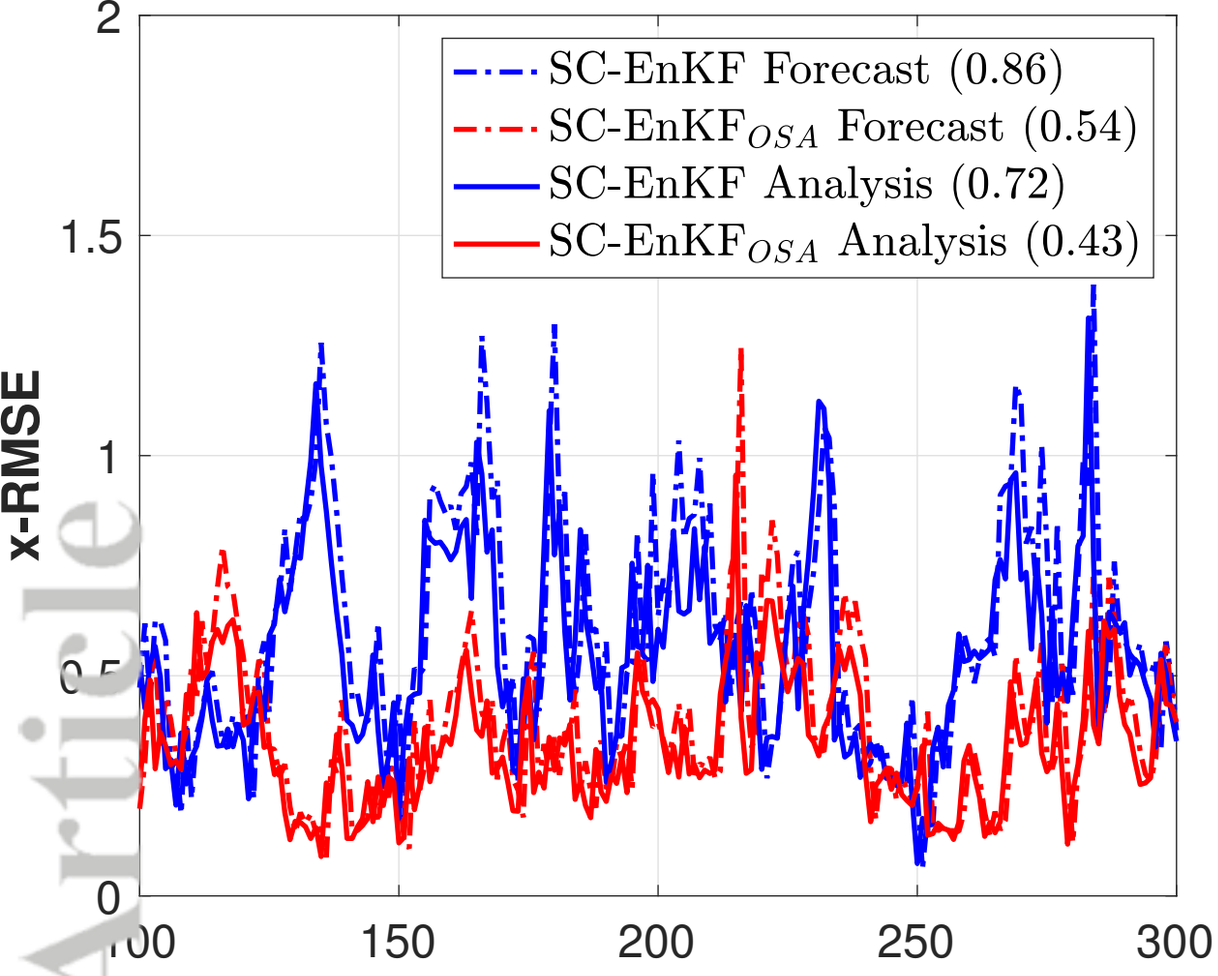


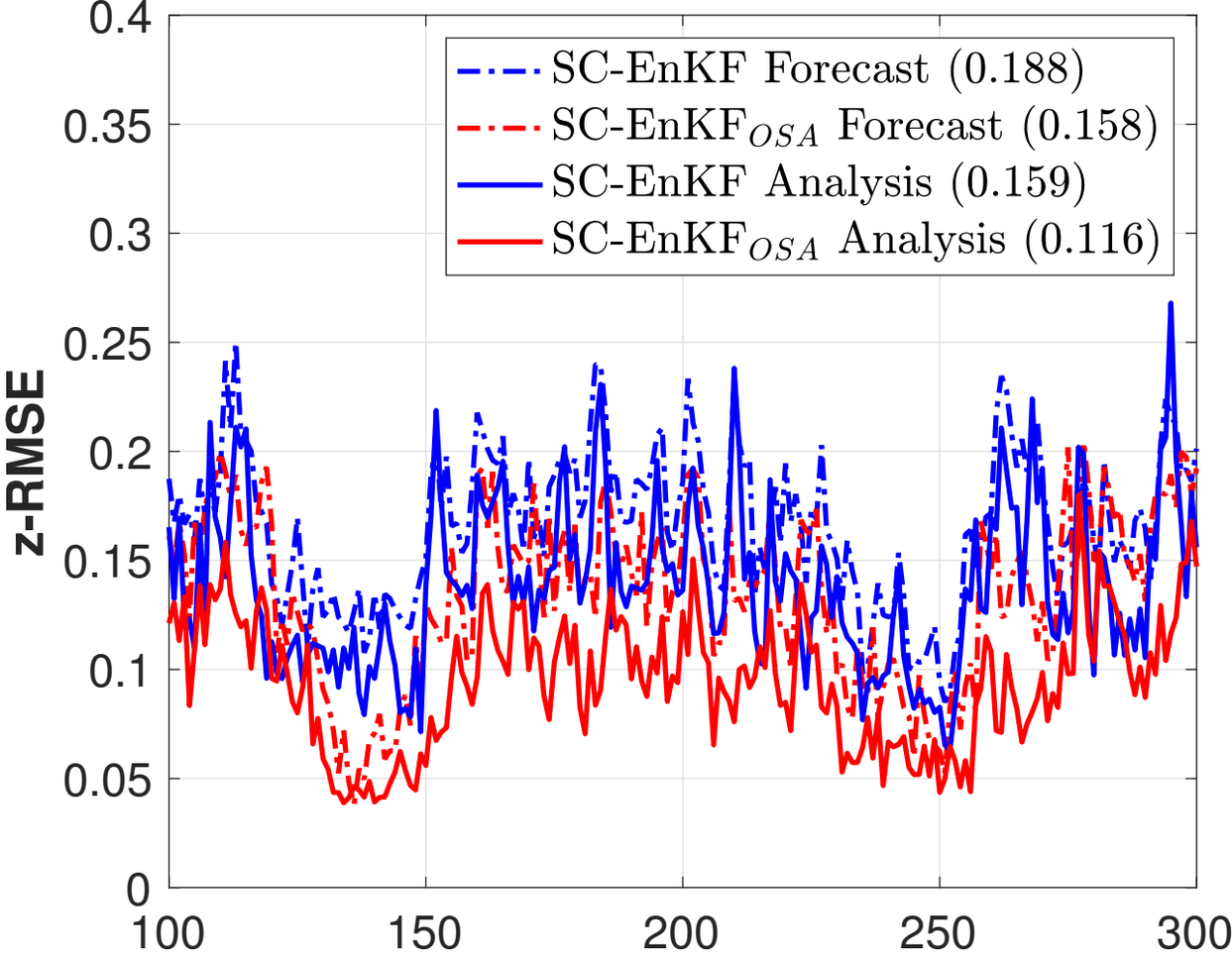


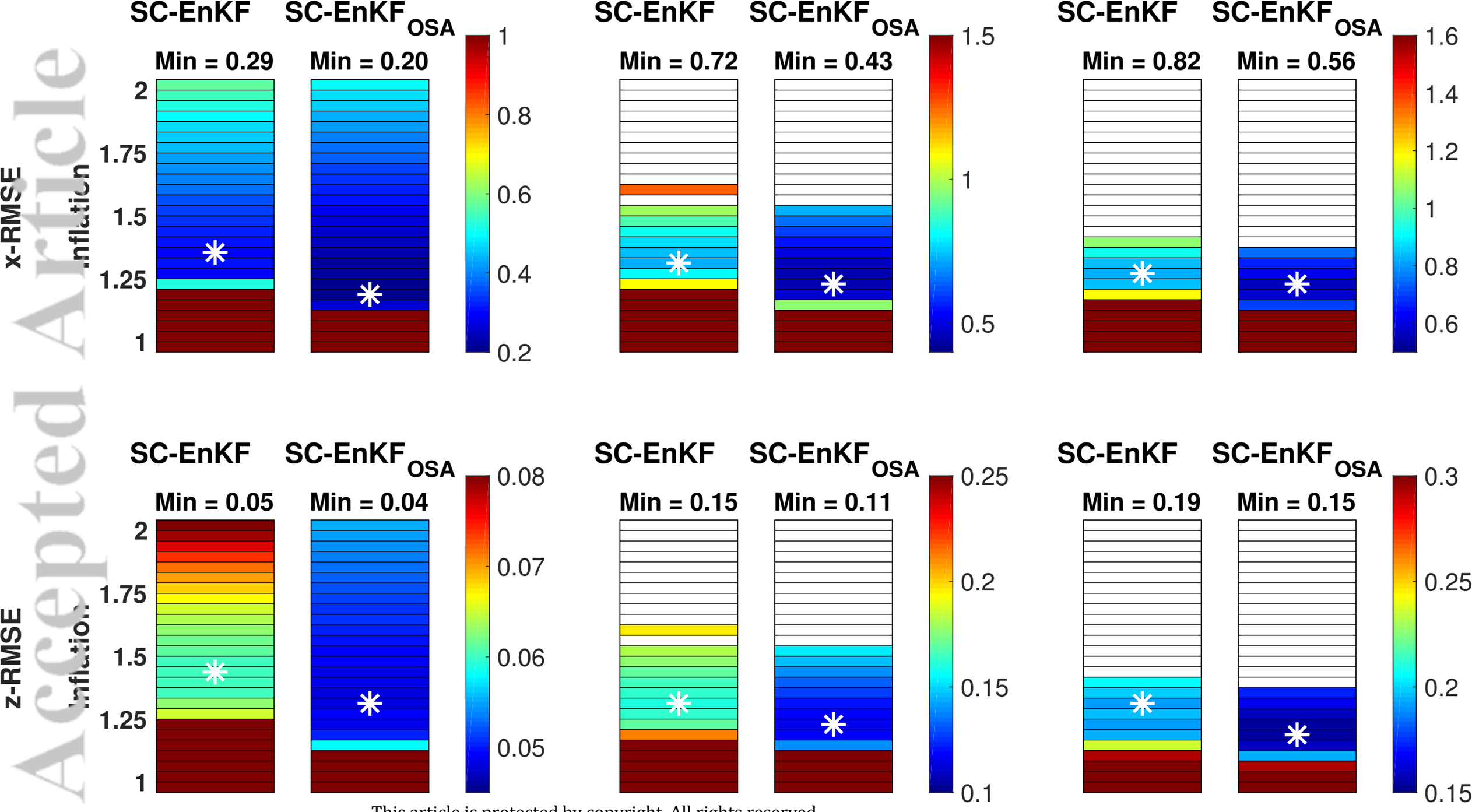


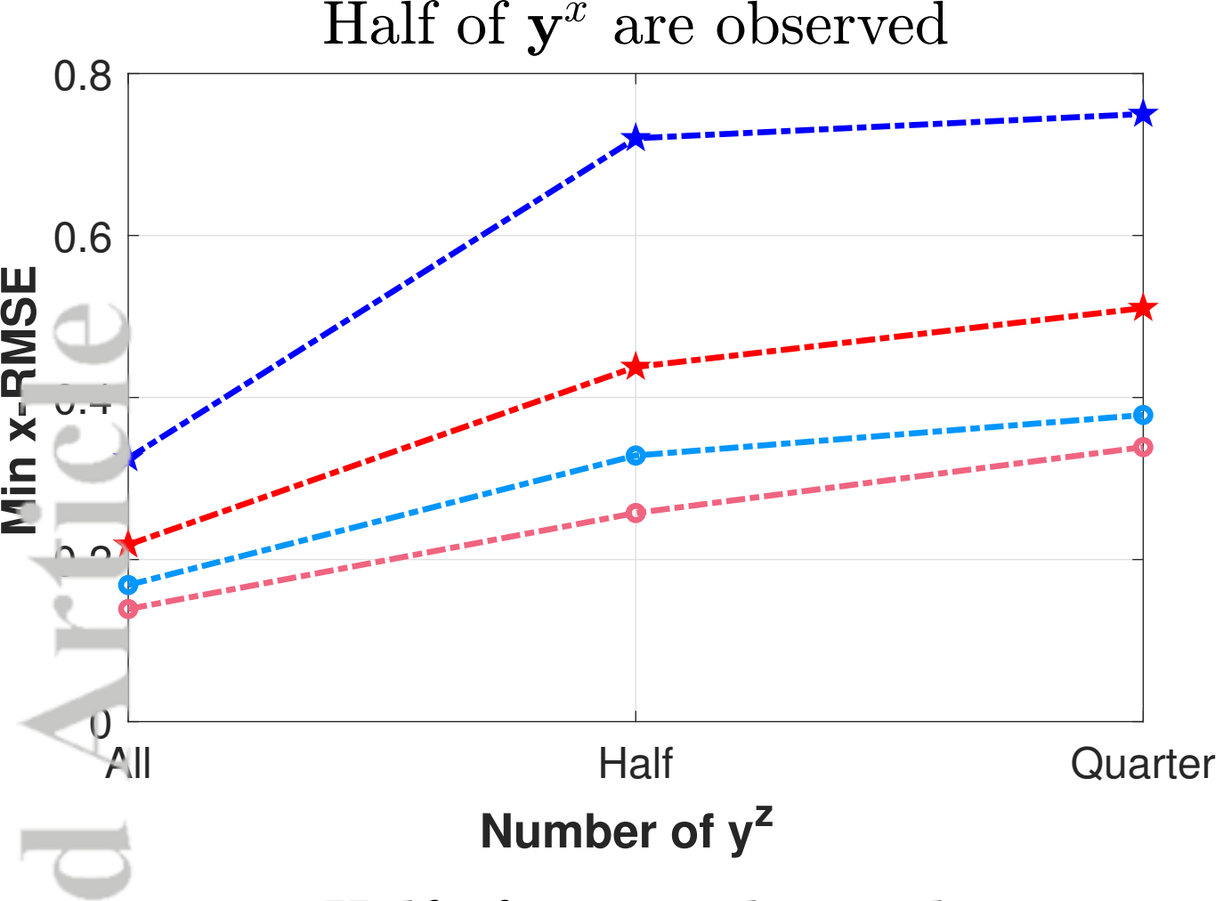


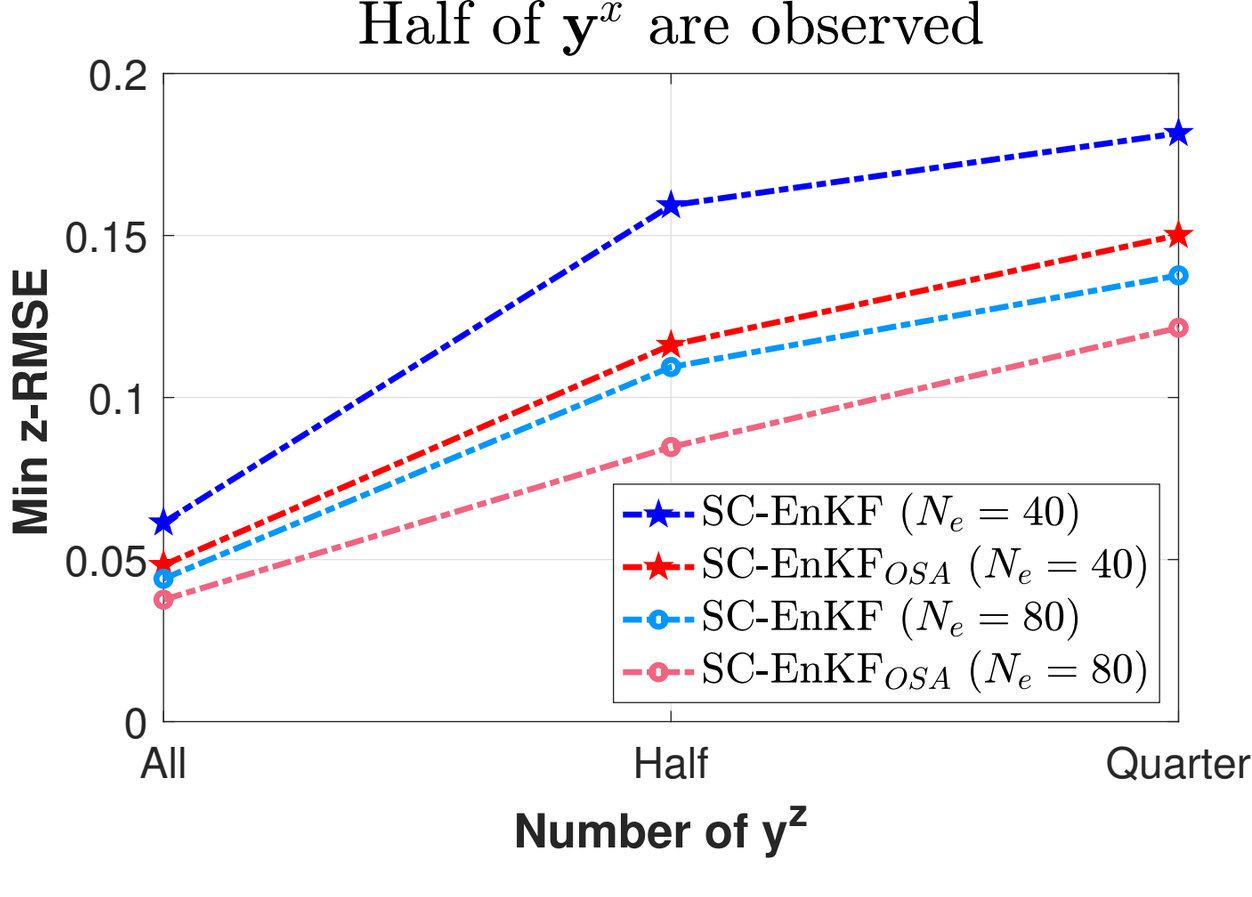


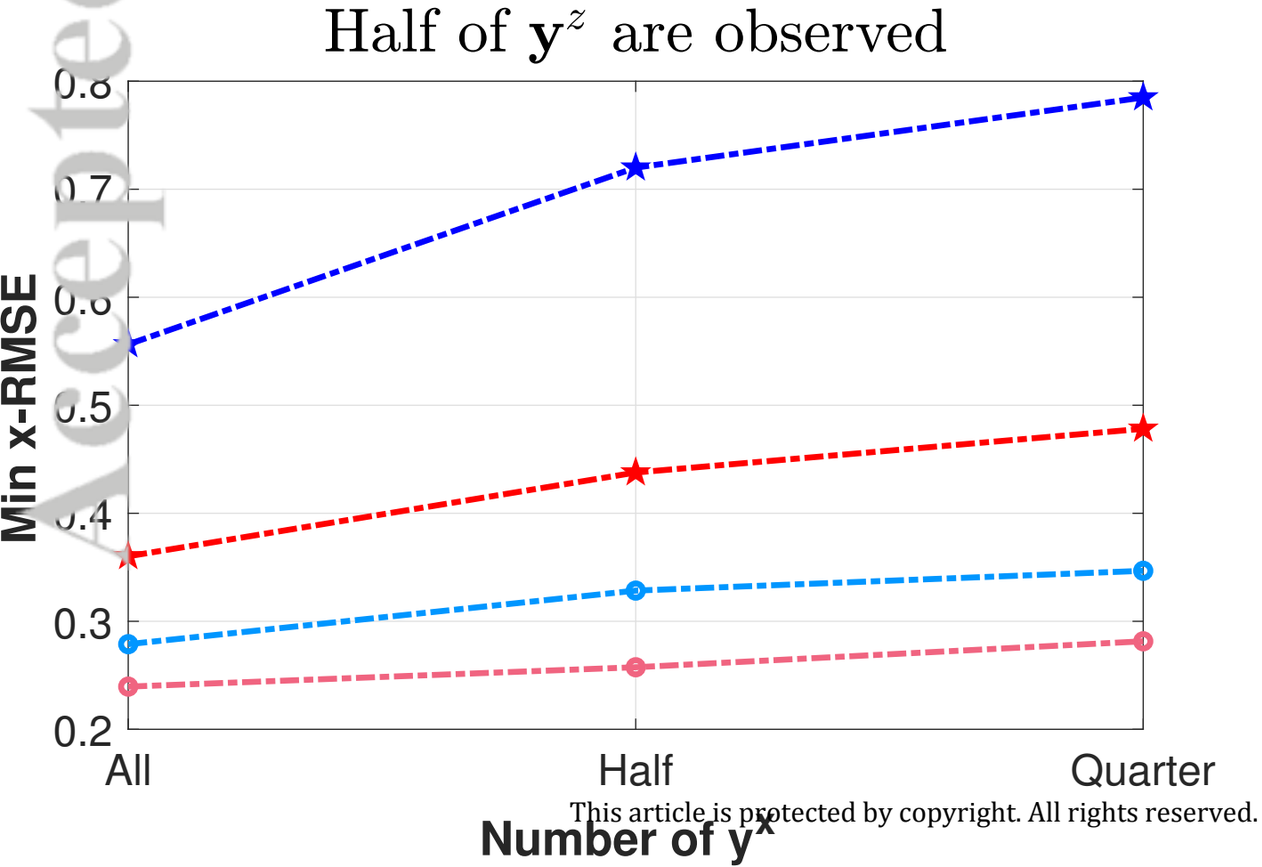


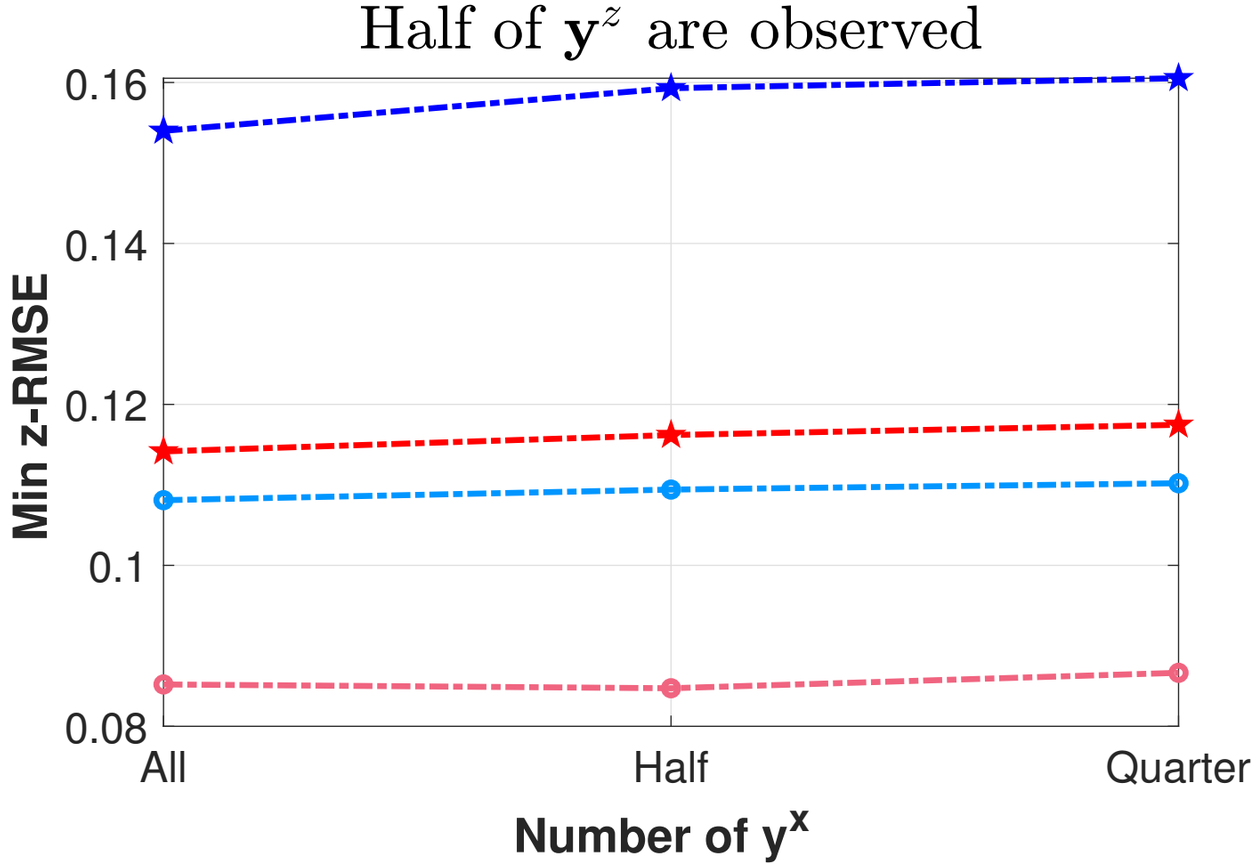


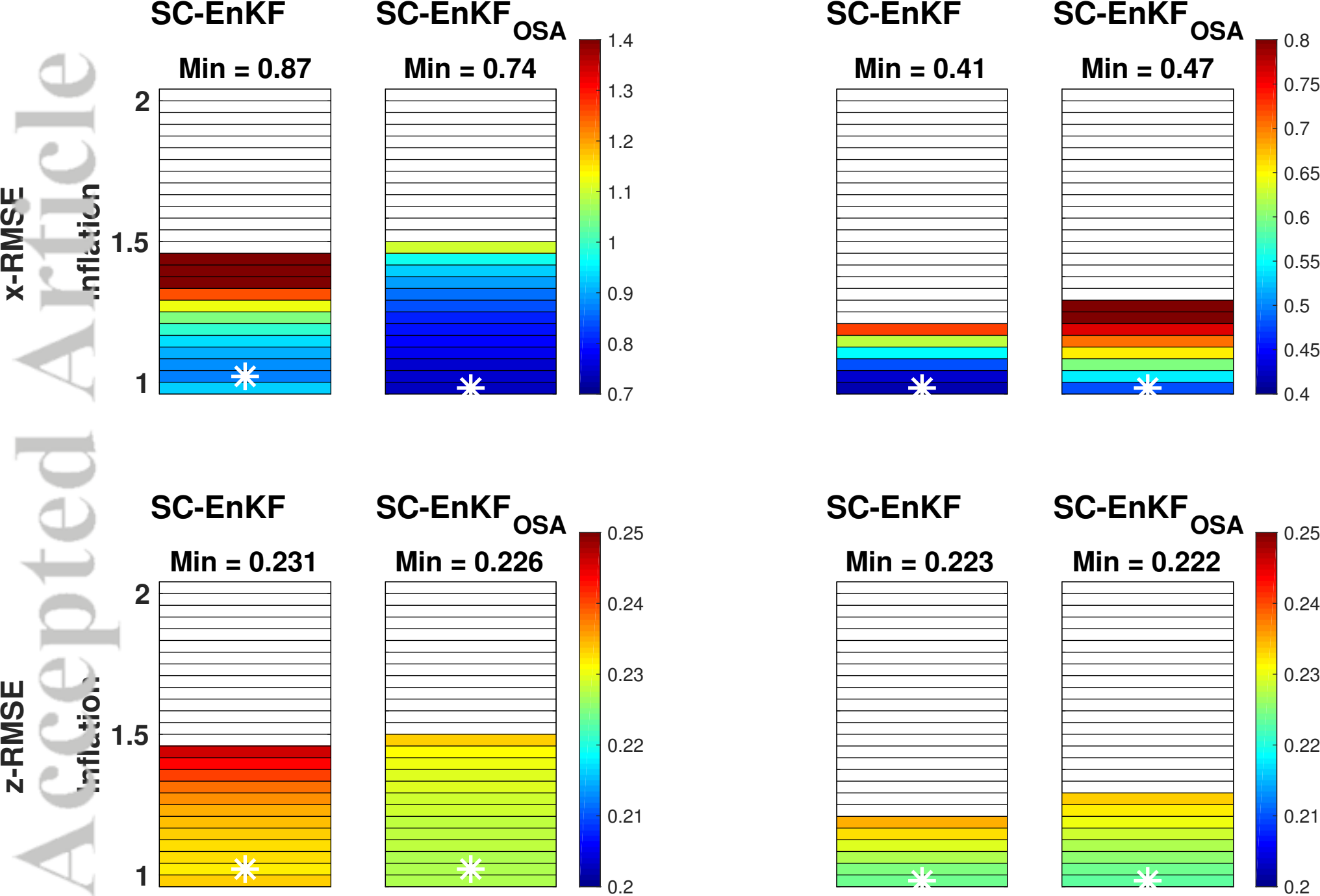






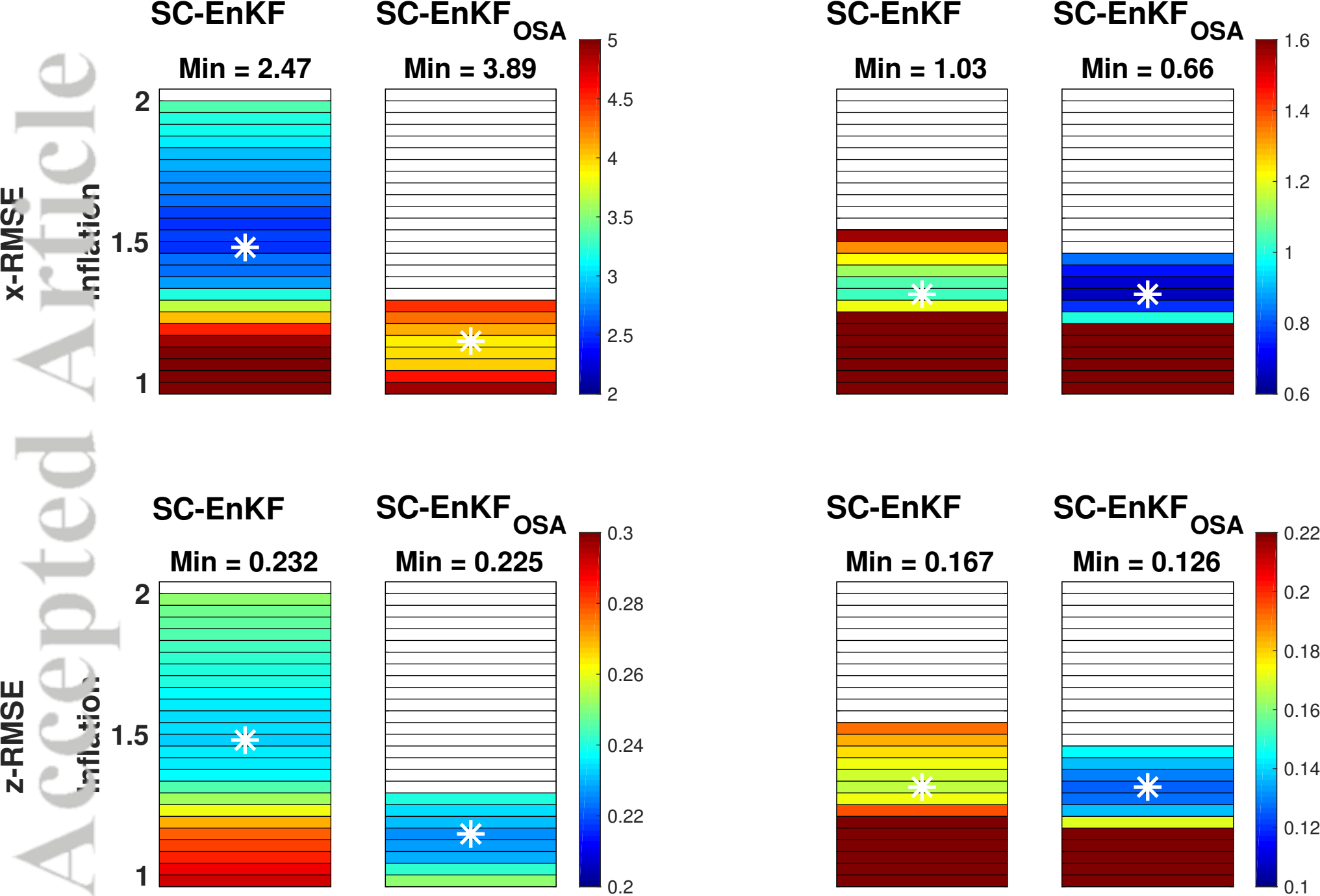






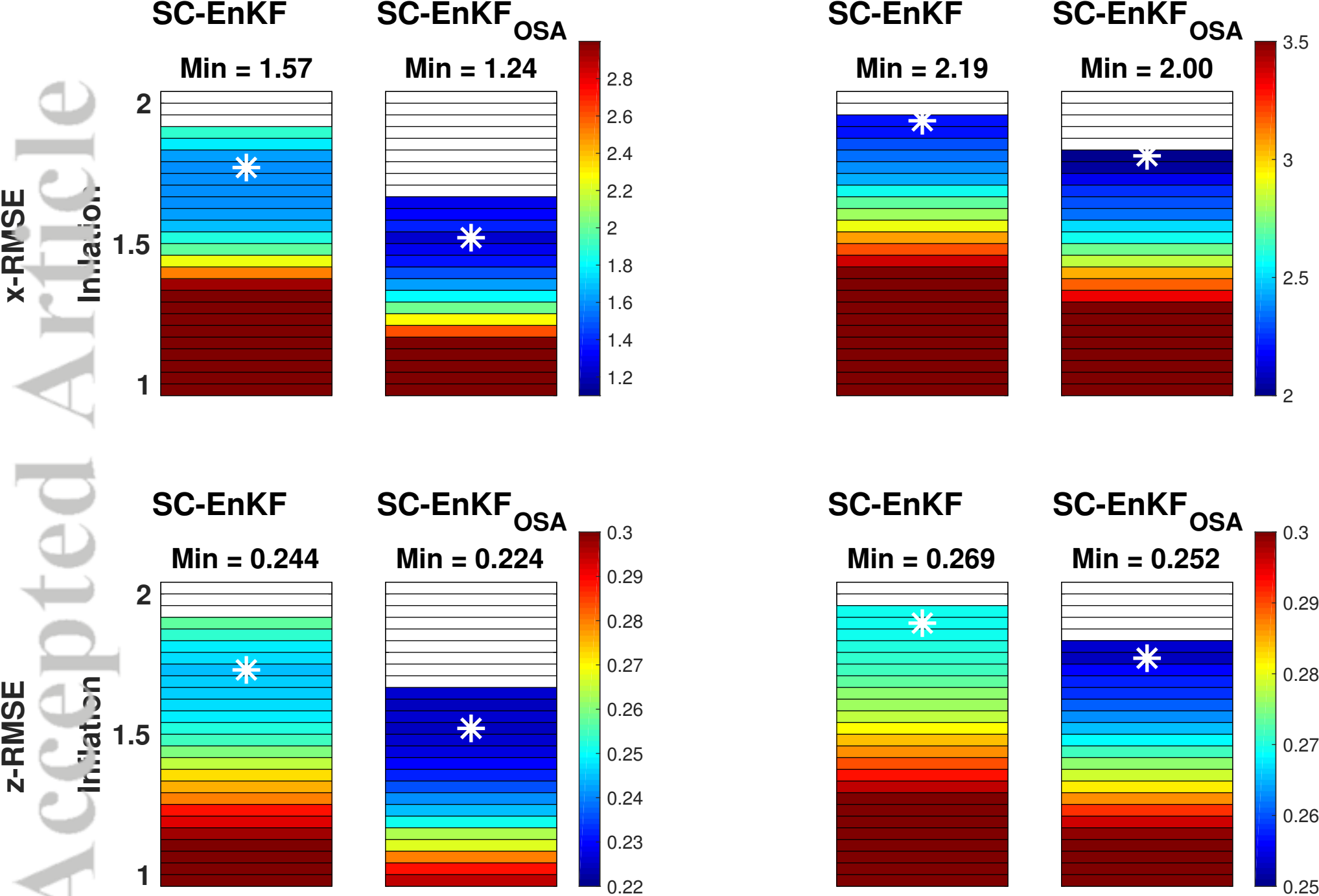
Only $\mathbf{y}^x$ observed every 40 ms

Only $\mathbf{y}^x$ observed every 10 ms



Only $\mathbf{y}^z$ observed every 40 ms

Only $\mathbf{y}^z$ observed every 10 ms



$c = b = 8$
$h = 0.8$
$F = 8$

$c = b = 6$
$h = 0.6$
$F = 6$