A Python package for data assimilation in the eWatercycle program – a hydrological framework

Zhenwu Wang, Rolf Hut, Nick van de Giesen

Department of Water Management, Delft University of Technology, Delft, 2628 CD, the Netherlands

Abstract

The **eWaterCycle** program (https://www.ewatercycle.org/) provides a collaborative environment for hydrological modelers, developed by the Netherlands eScience Center together with the Delft University of Technology. It aims to build a community of scientists in hydrology who use different programming languages for their specific models. Python is the lingua franca of the **eWaterCycle** platform and requires no modification to a particular model, making the platform user-friendly and flexible. Therefore, it can readily be applied in other geoscientific models. Currently, the Python data assimilation package includes ensemble-type methods, particle filters, and their variants, which are all sequential data assimilation algorithms. The implementation of techniques related to localization and inflation methods is included in this package. Localization and inflation are effective ways to avoid the collapse of a filter, which happens commonly in high dimensional models. The package gives access to all tunable parameters by configuration files quickly. To evaluate the performance of data assimilation comprehensively, a series of metrics is provided. In addition, the package offers a set of visualization tools to explore the results of data assimilation and the improvement of models.

How to use it –Part 1

Import necessary packages

```
# import all required packages
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt
import yaml
import io
import docker

# import an BMI model
from grpc4bmi.bmi_client_docker import BmiClientDocker
# import a Data assimilation algorithm
import LETKF
```

Settings

```
# set the model
pcrg = BmiClientDocker(image="ewatercycle/pcrg-grpc4bmi:latest", image_port=5555
5,
                       input_dir="./data/RhineMeuse30min",
                       output_dir="./output")

# configure the DA algorithm
set_DA = configure_DA(N = 20, local_scale = 1, ...)

# prepare observations
Observations = load(Observations)

# initial an evaluation object to collect metics during the process of DA
stats_res = evaluation(set_DA, set_metrics)
```

How to use it –Part 2

Build an ensemble

```
ensemble = []

# N is the number of ensemble members
for n in range (N):
    #add an ensemble methods
    ensemble.append(pcrg)
    ensemble[n].initialize('settings.yaml')
```

Run data assimilation

```
# M is the number of model states and obsSize is the number of observations
foreCastEnsemble = np.zeros([M, N])
observationEnsemble = np.zeros([obsSize, N])

while current_time < end_time:

    for n in range (N):
        ensemble[n].update()
        foreCastEnsemble[:,n] = ensemble[n].get_value('state')

    # H is an operator, which is used to transfer model states into observation
space
    if current_time > updateTime:

        stats_res.ens_valuate(current_time,truth=truth_now,type='forecast',E=for
eCastEnsemble)
        updateTime = updateTime + updateInterval
        analysesEnsemble = LETKF(foreCastEnsemble,observationEnsemble,H,set_DA)

        for n in range (N):
            ensemble[n].set_value('state',analysesEnsemble[:,n])

        stats_res.ens_valuate(current_time,truth=truth_now,type='analysis',E=ana
lysesEnsemble)
```

Plot results

```
# plot RMSE of the prior ensemble
plt.plot(stats_res.rmse.f)
# plot the spread of the prior ensemble
plt.plot(stats_res.rmv.f)
# plot the rank histogram of the prior ensemble
plt.plot(stats_res.rh.f)
```

Visualization

Figure 1. Comparison

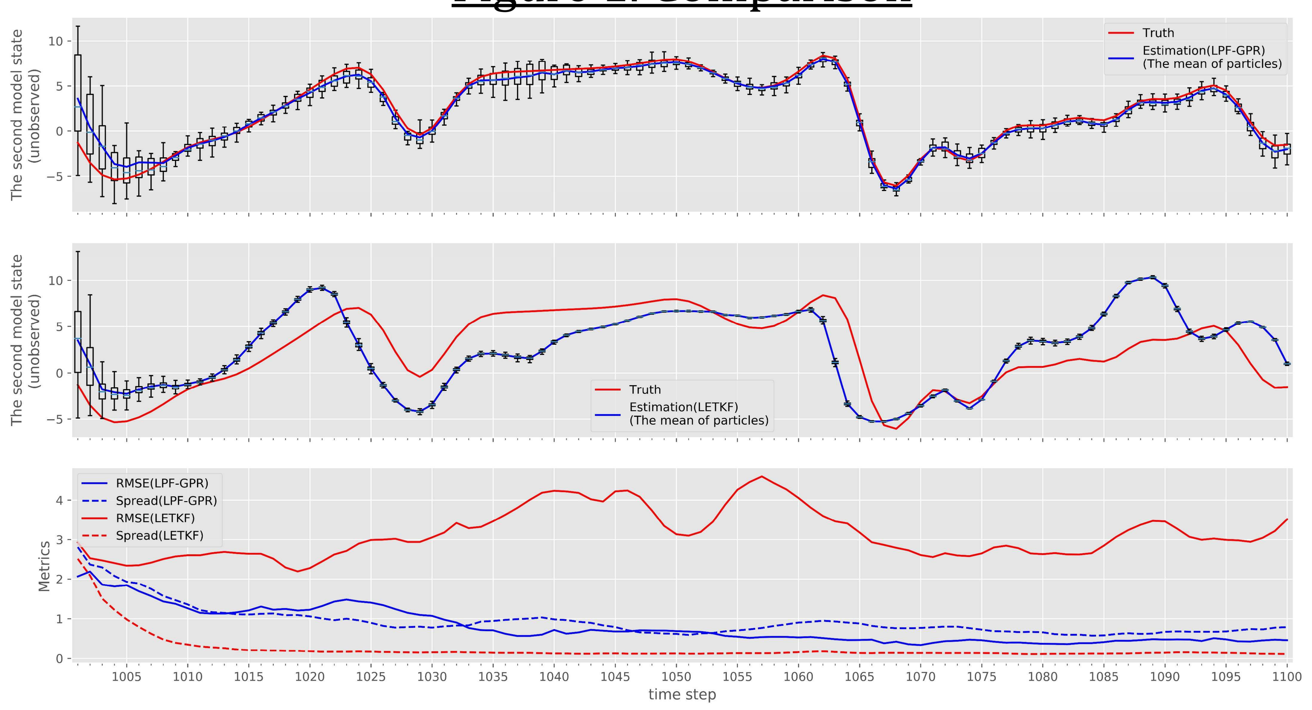
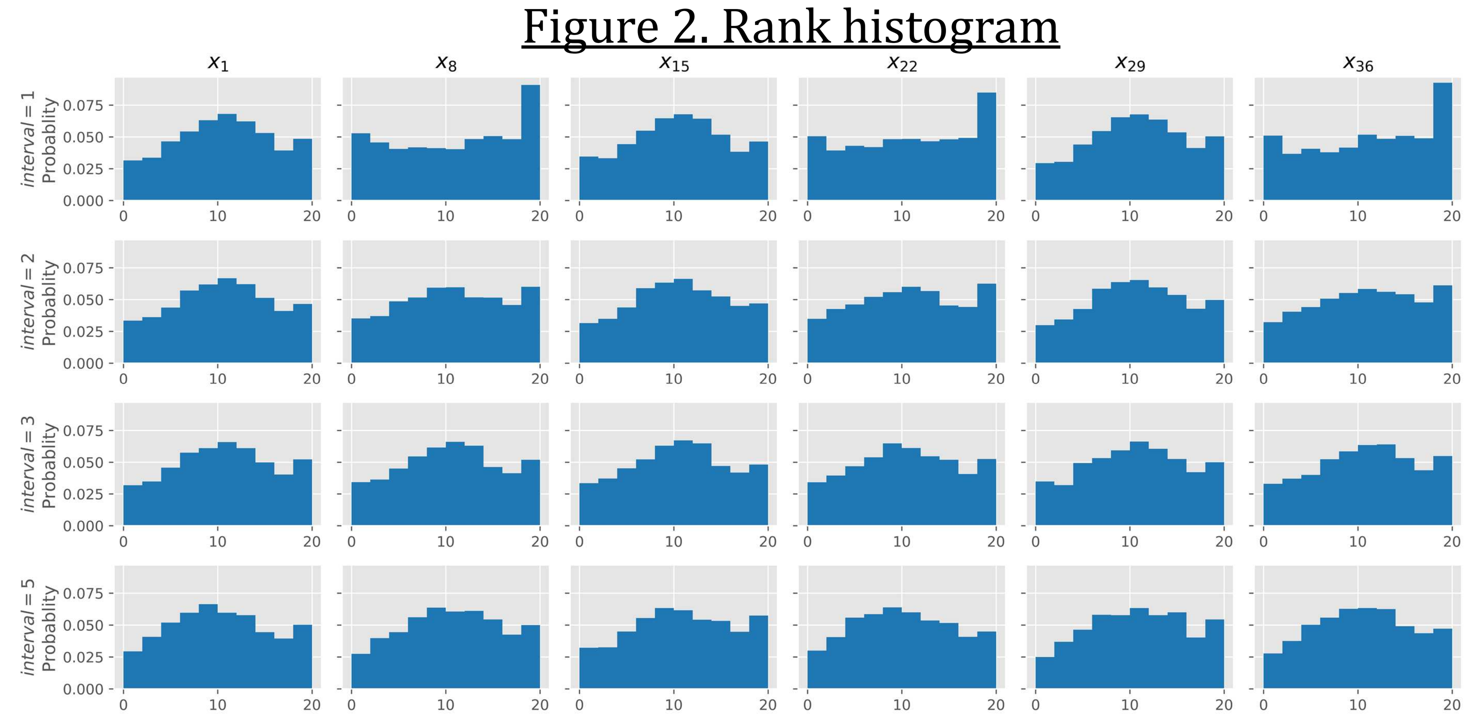


Figure 2. Rank histogram



eWaterCycle

References

[1] Hut, R., Drost, N., van Meersbergen, M., Sutanudjaja, E., Bierkens, M., & van de Giesen, N. (2016). eWaterCycle: a hyper-resolution global hydrological model for river discharge forecasts made from open source pre-existing components. Geoscientific Model Development.

[2] Hut, R., van de Giesen, N., & Drost, N. (2018, April). The future of global is local. eWaterCycle II: bridging the gap between catchment hydrologists and global hydrologists. In EGU General Assembly Conference Abstracts (Vol. 20, p. 10614).

Contact

Email: zhenwu.wang@tudelft.nl
Address: Faculty of Civil Engineering and Geosciences (Building 23), Stevinweg 1, Delft, the Netherlands