Dear Referee #1,

Thank you very much for reading our manuscript carefully and providing us the insightful comments.

Referee's comment:

> *Review of the manuscript 'Fortnight conditioning of historical data to improve short-term precipitation predictions' by Yoshito Hirata and Yoshinori Yamada*

> *The present manuscript tries to attribute short-range precipitation predictability in the large Tokyo megalopolis to the indirect effect of aerosols produced by anthropogenic activities, through their influence on the production of precipitation nuclei and optic effects.*

Our response:

> Thank you very much for your comments here. Our intention here is to show the fact that short-term time series prediction of precipitation may be improved by conditioning past data periodically. However, we do not mean to argue the underlying mechanisms of why this periodicity is generated.

Referee's comment:

> *The manuscript is very short not giving enough details for the appropriate reproducibility of the results. Moreover, the methodology, and the arguments in the discussion are very dubious and even not falsifiable, which is fundamental requirement in any scientific theory. Moreover, there are severe methodological shortcomings, described below. Giving those reasons, the manuscript is judged not reaching enough standards to be published in NPG.*

Our response:

Thank you very much for sharing your views. However, we do not agree with your views because (i) our message is quite simple: we can improve short-term time series prediction of precipitation by just conditioning past data periodically. This message itself should be practically important enough to be published so that we may be able to prevent causalities due to the heavy rains; (ii) we have described our method concisely in the Appendix. In addition, we have attached our codes so that readers can reproduce what we have done; (iii) In the reply to Dr. Kondrashov, we supplied the results of a toy example forced periodically. This example demonstrates that by conditioning past data periodically, we can take into account the periodicity of the underlying dynamics in our time series prediction.

Referee's comment:

The present study should be preceded by experiments with a toy minimal model, reproducing convection and precipitation mechanisms triggered by aerosol nucleation. Then, predictability experiments should be run by imposing some weekly periodicity to aerosol emissions to simulate the periodic anthropogenic forcing and seek whether any phase synchronizing is observed in precipitation. The predictability study described in the manuscript, obtained with timeseries only is far unsatisfactory due to the existence of a vast number of noncontrolled factors, beyond aerosols. It is thus very difficult to produce a convincing quantifiable attribution of the very-short term precipitation predictability to the aerosol's forcing.

Our response:

Thank you very much for your comments here. In the reply to Dr. Kondrashov, we have provided a toy example. Thus, we have shown that by the proposed approach, we could consider the underlying periodicity in short-term time series prediction. As mentioned above, we do not intend to discuss the underlying paths on why this periodicity occurs because (i) our finding itself has the applicational values and (ii) we do not have the expertise to discuss the underlying mechanisms. By establishing the fact

that time series prediction can be improved by conditioning past data periodically, then the underlying mechanisms should be further investigated by the other researchers who know better the underlying physics of precipitation.

Referee's comment:

The applied methodology is dubious and impacted by severe pitfalls such as:

- *The method of analogues is too little described; for instance, the analogs metric is not clear. Is it based on precipitation only? If yes, the analog's distance is too strict.*

Our response:

Thank you very much for sharing your view. Yes, our time series prediction is only based on past data of precipitation. In our prediction procedure, we do not have to run a big meteorological model or some so that we can issue time series predictions in 1 minute resolution up to 2 hours ahead. We will consider combining the other weather variables as a future research topic. Thanks.

Referee's comment:

- *It is not clear if analogs are sought in an independent period of the validation period.*

Our response:

Predictions were made by using the information up to the certain time and thus independent of the future values of precipitation.

Referee's comment:

- *The details of the AR model are not described. Other benchmark stochastic models should be tested.*

Our response:

We are sorry that we have not described the mathematical detail of the AR model, although we had explained by words in lines 44-45. In the AR model, we fit the following model by the least square fitting:

$$\hat{s}(t+p) \sim a_{p,-1} + \sum_{d=0}^{119} a_{p,d} s(t-d).$$

We used the dataset of 2006 to find the parameters $a_{p,d}$ and evaluated the prediction errors on the dataset between 2007 and 2015.

We presume that the AR model should be one of our benchmark stochastic models. We also have included the persistence model as well as the mean prediction model as the other benchmarks. Therefore, if you think that we should include the other benchmark stochastic models, could you raise an example of such methods or some to be included so that we can evaluate more concretely whether we should consider the other benchmarks or not.

Referee's comment:

- *By forecast rank, authors mean error, so authors should precise that.*

Our response:

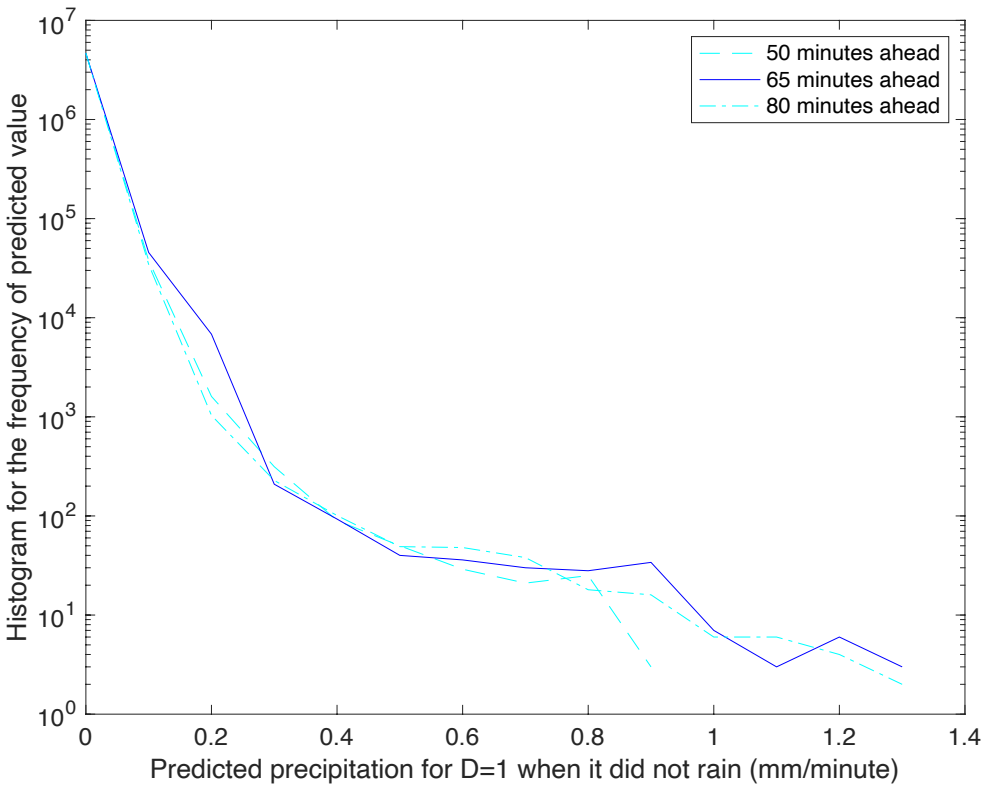
We ranked the mean absolute error for each method for each prediction steps. Then, we took the mean of the ranks over the prediction steps. We hope that our meaning is clearer.

Reviewer's comment:

The unique figure presented is not fully discussed. There are results which are not understandable neither discussed such as: the bump in rank around the forecast delay 60-70 minutes for D=1; the reason why the predictability is larger when analogous are sought with D=14 than D=7. Authors present a very speculative unproven reason for that: 'there is a period doubling bifurcation in the precipitation and that a week periodicity, if it exists, could be unstable'.

Our response:

Thank you very much for your comments here. For D=1, we have observed that 65 minutes ahead prediction forecasted more rains than 50 or 80 minutes ahead prediction when it did not rain actually as shown in Additional Figure 4. In addition, we decided to withdraw our statement related to the above period doubling bifurcation. Based on the toy example of the periodically forced Rössler model, we can say that we could do better in the medium-term time series prediction when we turned the periodicity of the conditioning to the right period. As far as we tested the period doubling could not be observed in the toy model. We deeply appreciate your critical comments, by which we could remove this wrong statement.



Additional Figure 4 | Histogram for predicted precipitation for 50, 65, and 80 minutes ahead predictions when we set D=1.