

Supplement material

**Socio-hydrologic data assimilation: Analyzing human-flood interactions by model-data integration**

Yohei Sawada<sup>1</sup>, Risa Hanazaki<sup>2</sup>

<sup>1</sup> Institute of Engineering Innovation, School of Engineering, the University of Tokyo,  
Tokyo, Japan

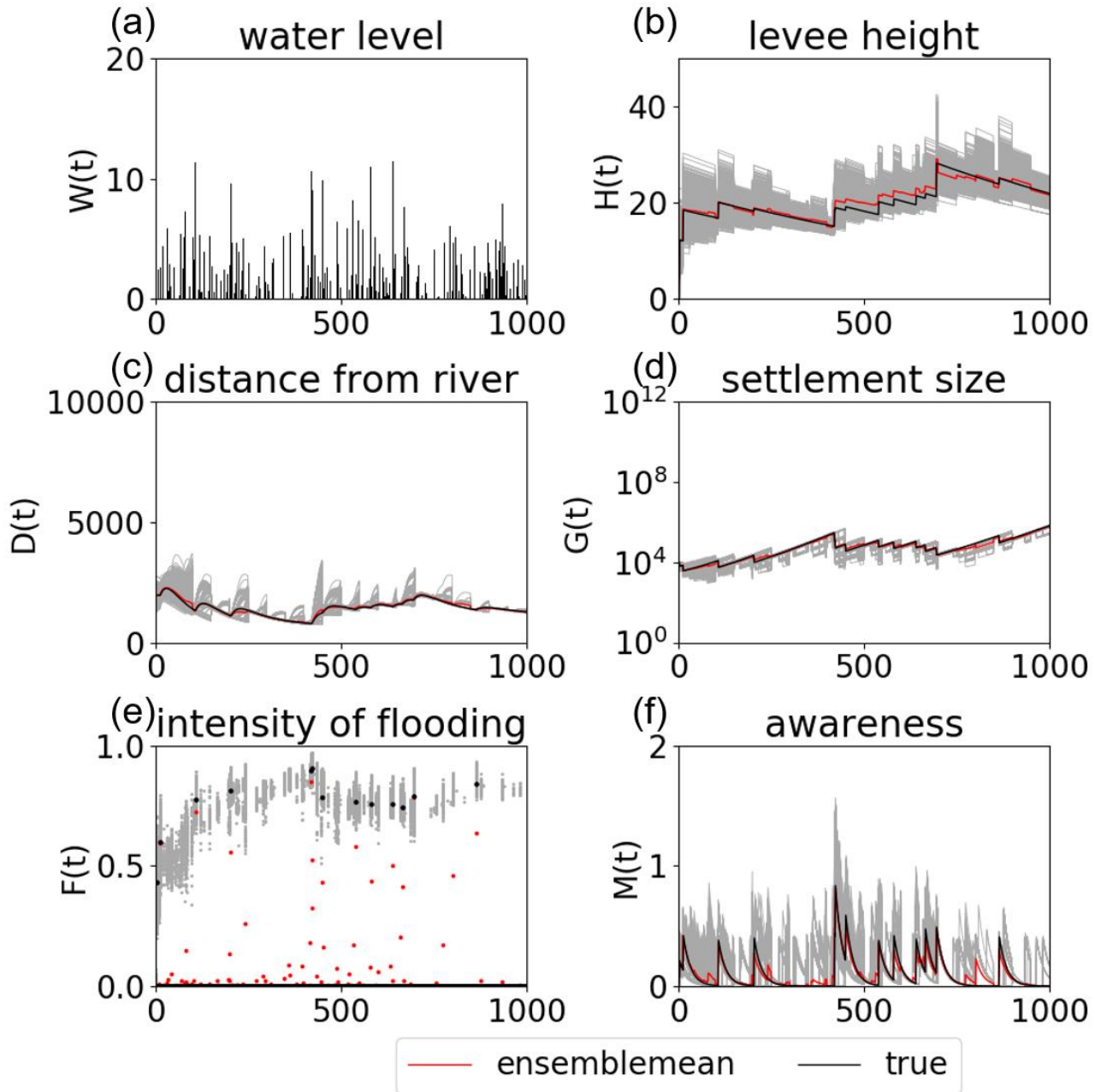
<sup>2</sup> Institute of Industrial Science, the University of Tokyo, Tokyo, Japan

**Contents of this file**

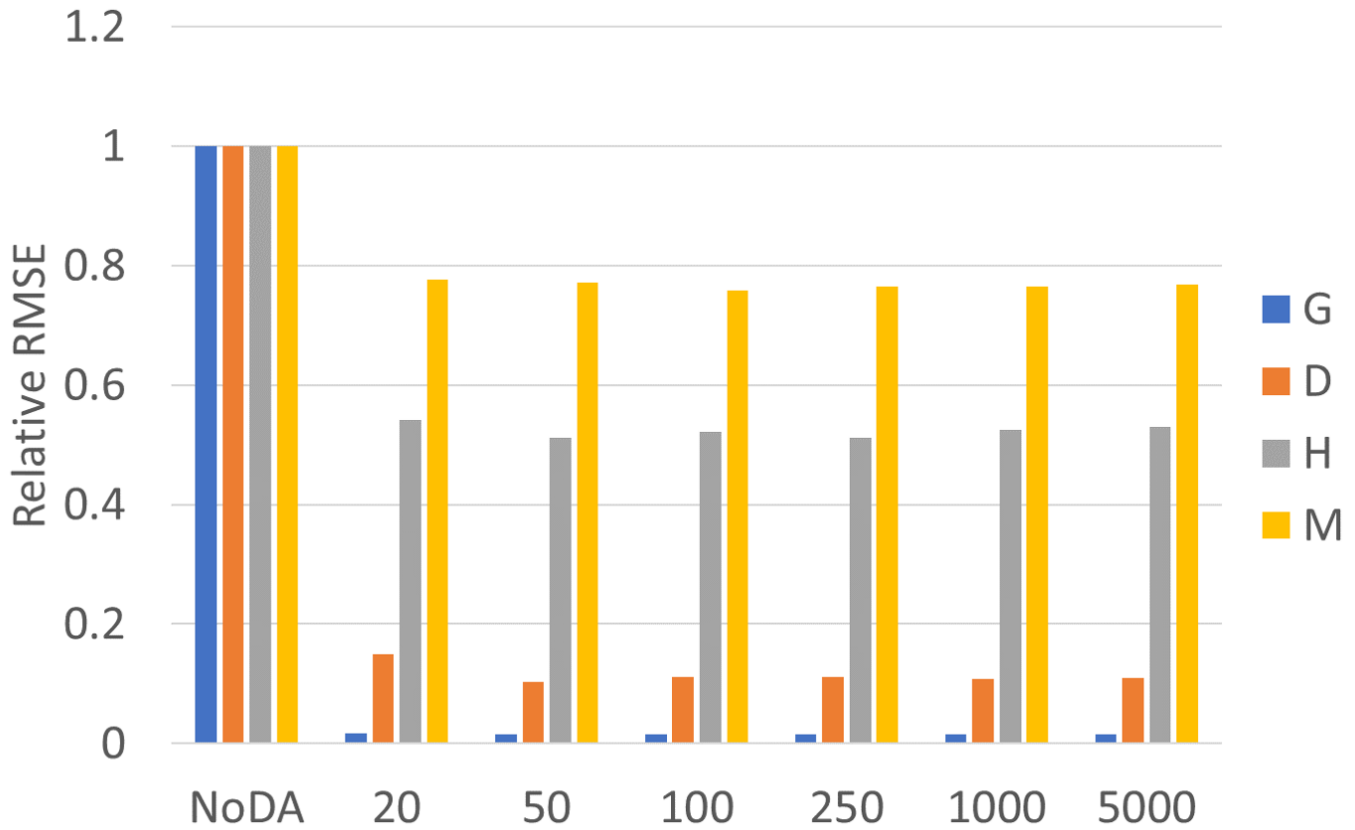
Figures S1 to S5

**Introduction**

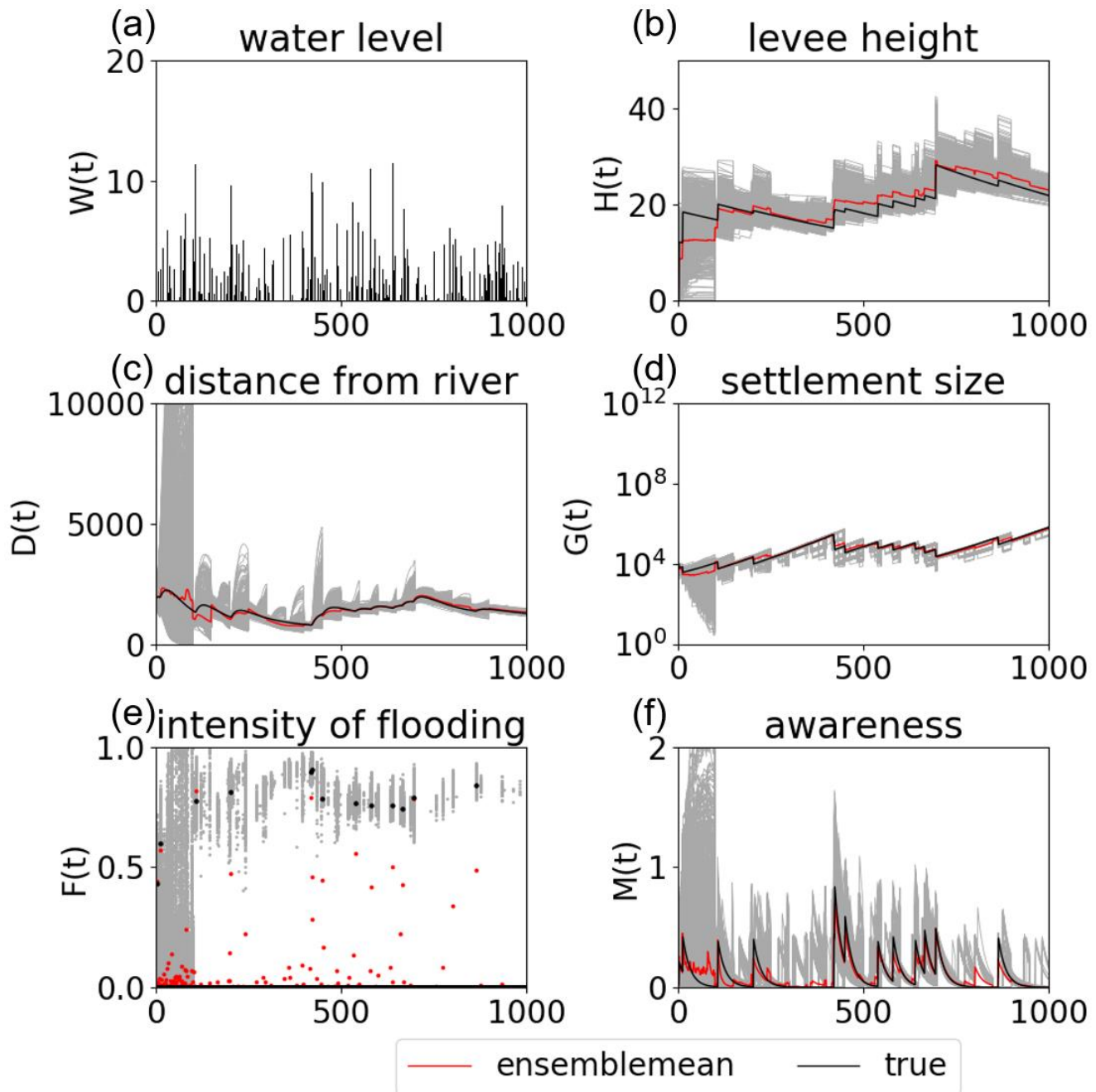
This supporting information provides the figures which support the findings shown in the main article.



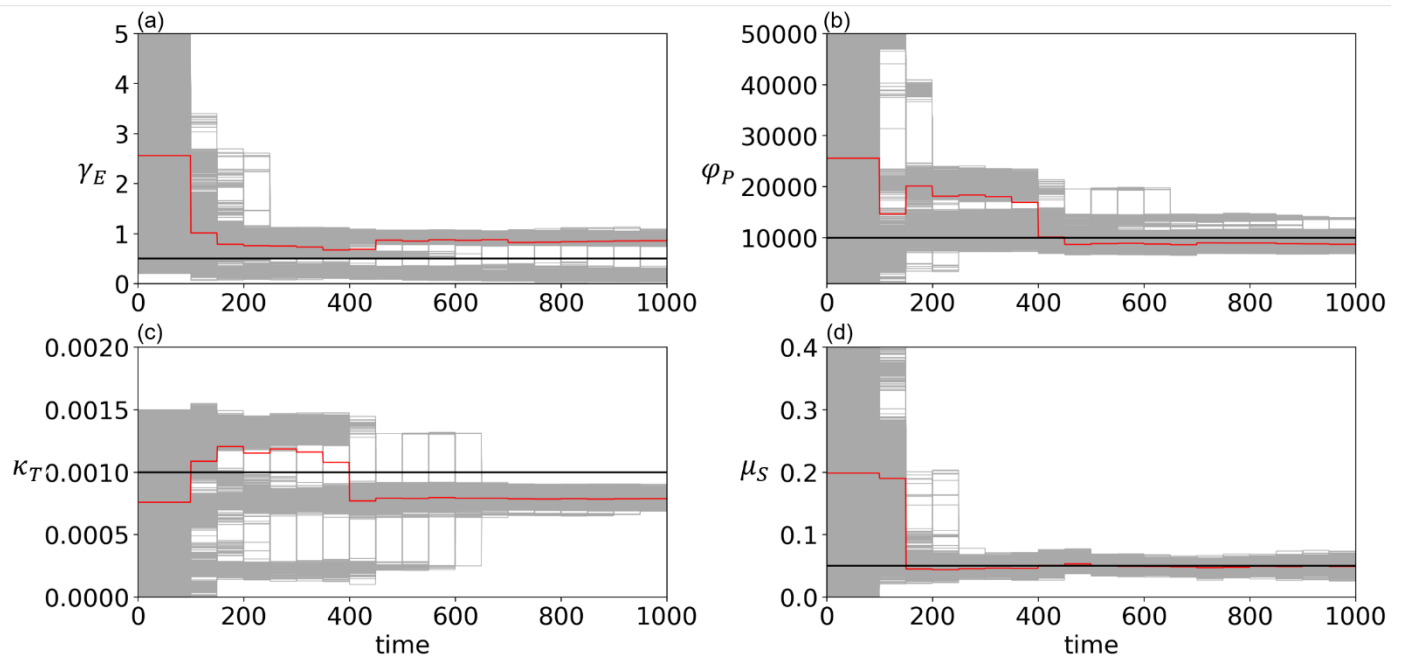
**Figure S1.** Timeseries of (a) high water level  $W(t)$ , (b) the flood protection level (or levee height)  $H(t)$ , (c) the distance of the center of mass of the human settlement from the river  $D(t)$ , (d) the size of the human settlement  $G(t)$ , (e) the intensity of flooding events  $F(t)$ , and (f) the social awareness of the flood risk  $M(t)$  simulated by the data assimilation experiment in which the observations of  $F$ ,  $G$ ,  $D$ ,  $H$ , and  $M$  are assimilated into the model every 50 years with 5000 ensembles in the experiment 1 (see section 3.1). Grey, red, and black lines are the ensemble members, their mean, and the synthetic truth, respectively.



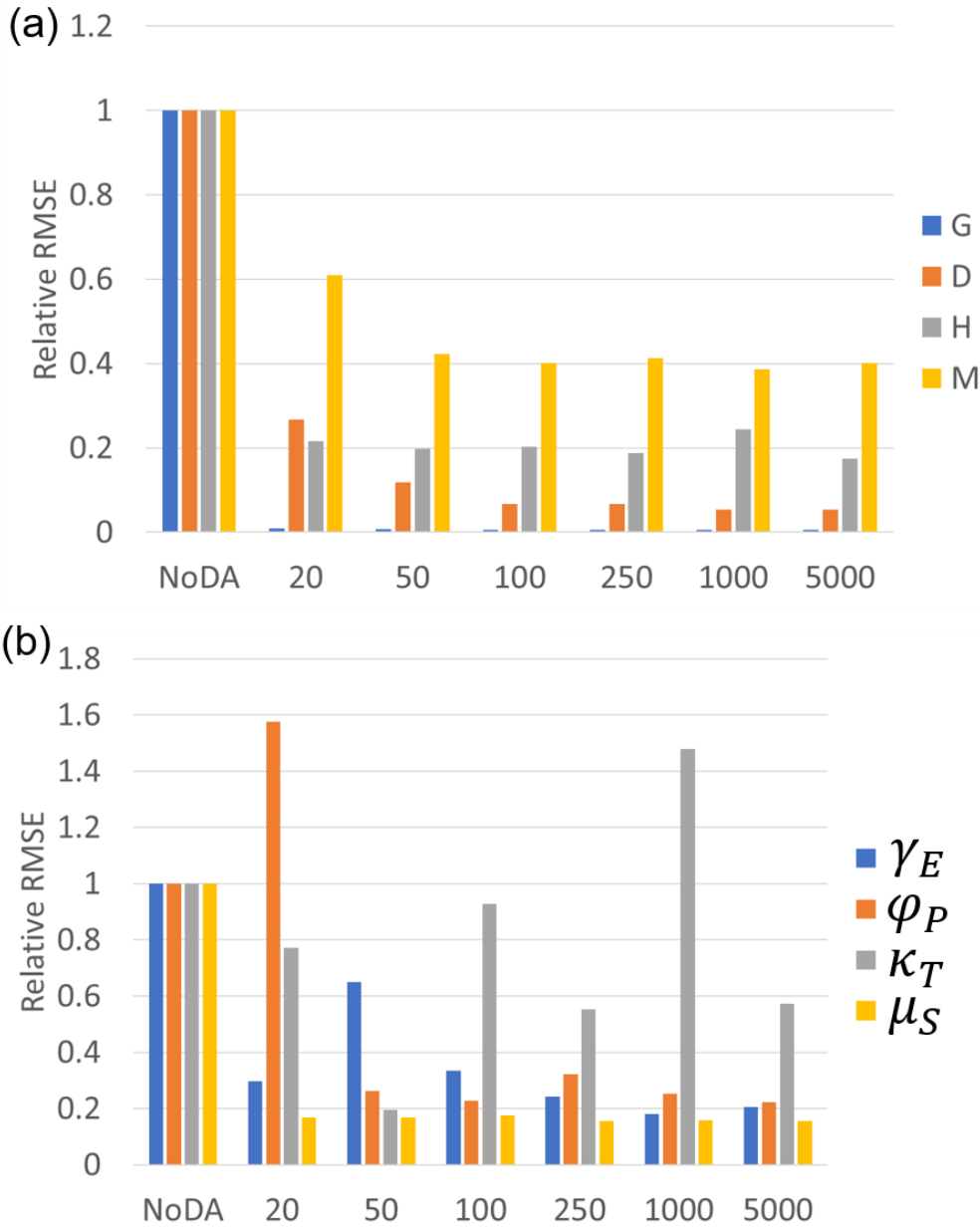
**Figure S2.** The ratio of RMSEs of the no data assimilation experiment (NoDA) to those of the data assimilation experiments in which all of observations (F, G, D, H, and M) are assimilated with 20, 50, 100, 250, 1000, and 5000 ensembles in the experiment 1 (see section 3.1). Blue, orange, gray, and yellow bars are RMSEs of the size of the human settlement  $G(t)$ , the center of mass of the human settlement from the river  $D(t)$ , the flood protection level (or levee height)  $H(t)$ , and the social awareness of the flood risk  $M(t)$ .



**Figure S3.** Timeseries of (a) high water level  $W(t)$ , (b) the flood protection level (or levee height)  $H(t)$ , (c) the distance of the center of mass of the human settlement from the river  $D(t)$ , (d) the size of the human settlement  $G(t)$ , (e) the intensity of flooding events  $F(t)$ , and (f) the social awareness of the flood risk  $M(t)$  simulated by the data assimilation experiment in which the observations of  $F$ ,  $G$ ,  $D$ ,  $H$ , and  $M$  are assimilated into the model every 50 years with 5000 ensembles in the experiment 2 (see section 3.2). Grey, red, and black lines are the ensemble members, their mean, and the synthetic truth, respectively.



**Figure S4.** Timeseries of (a) the cost of levee raising  $\gamma_E$ , (b) the rate by which new properties can be built  $\varphi_P$ , (c) the rate of decay of levees  $\kappa_T$ , (d) memory loss rate  $\mu_S$  estimated by the data assimilation of all observations (F, G, D, H, and M) with 5000 ensembles every 50 years in the experiment 2 (see section 3.2). Grey, red, and black lines are the ensemble members, their mean, and the synthetic truth, respectively.



**Figure S5.** The ratio of RMSEs of the no data assimilation experiment (NoDA) to those of the data assimilation experiments in which all of observations (F, G, D, H, and M) are assimilated with 20, 50, 100, 250, 1000, and 5000 ensembles in the experiment 2 (see section 3.2). (a) Blue, orange, gray, and yellow bars are RMSEs of the size of the human settlement  $G(t)$ , the center of mass of the human settlement from the river  $D(t)$ , the flood protection level (or levee height)  $H(t)$ , and the social awareness of the flood risk  $M(t)$ . (b) Blue, orange, gray, and yellow bars are RMSEs of the cost of levee raising  $\gamma_E$ , the rate by which new properties can be built  $\varphi_P$ , the rate of decay of levees  $\kappa_T$ , memory loss rate  $\mu_S$