Real-world experiment

Experiment design

In addition to the OSSEs, we performed the real-world experiment in the city of Rome, Italy. Ciullo et al. (2017) collected real-world data and calibrated their flood risk model. Using the data collected by Ciullo et al. (2017), we performed the data assimilation experiment. It should be noted that the flood risk model of Ciullo et al. (2017) is different from our model (i.e. Di Baldassarre et al. 2013), although they are conceptually similar.

All the data were collected from Figure 1 of Ciullo et al. (2017) by WebPlotDigitizer (https://automeris.io/WebPlotDigitizer/). The observed high water level of Tiber River was used as input forcing data (W). The levee height (H) and population (G) were used as the observation data to be assimilated into the flood risk model. In Ciullo et al. (2017), population values within the Tiber's floodplain were normalized by the theoretical maximum Tiber's floodplain population which is estimated to the range between 10^6 and 2×10^6. Since our flood risk model needs the population values (not normalized values), we multiplied 1.5×10^6 and the normalized values shown in Figure 1 of Ciullo et al. (2017) to obtain population in the floodplain.

We added lognormal multiplicative noise to the observed high water level as we did in the OSSEs. The observation errors of levee height and population were set to 10% and 25% of the observed values, respectively. Since Ciullo et al. (2017) showed the large uncertainty in the estimation of the theoretical maximum population (see above), it is reasonable to assume that the estimation of population values also has relatively large uncertainty.

As the second and third OSSEs, we have 4 unknown parameters in this real-world experiment. We used the same settings of parameters as the OSSEs, which are shown in Table 1, except for ξ_H, proportion of additional high water level due to levee heightening. In this real-world experiment, we set $\xi_H = 0$ because the observed high water level includes the effects of levee heightening. This treatment is consistent to Ciullo et al. (2017) (see their Table 2).

The initial conditions of H and M were set to 0. The initial conditions of D were obtained from the uniform distribution between 1000 and 5000. The initial conditions of G were obtained from 1500 and 50000. Since we have no information of the initial conditions, we assumed the large uncertainties of them.

Results

Figure 1 shows the timeseries of the model variables calculated by 5000 ensembles with no data assimilation. The 5000-ensemble simulation reveals the two bifurcated social systems. One builds a high levee and maintains a course of stable economic growth. The other one has no levee and its economy is damaged by severe floods many times (ensemble mean shown in Figure 1b implies that there are many ensemble members with zero levee height).

In reality, the city of Rome constructed the levee responding to the severe flood occurred on 28 December 1870. After the construction of this levee, no major flood losses occurred, allowing the steady and undisturbed growth. Figure 2 indicates that our SIRPF successfully constrains the trajectory of the ensemble simulation to the real-world (i.e. high levee and stable economic growth) by assimilating the real data of H and G. Figure S1 shows the SIRPF-estimated unknown parameters. Our SIRPF suggests lower γ_E than the initial ensemble mean to promote the levee construction with lower costs. Lower κ_T is also obtained because the assimilated real data show no decay of levee from 1874 to

2009. Compared with the OSSE experiment 2, the large uncertainty in estimated parameters remains at the final timestep due to the limited number of assimilated observations.

We analyzed the impacts of the individual observation types (i.e. H and G) on the simulation skill as we did in the OSSEs. Figure 3 indicates that our SIRPF realistically simulates the socio-hydrologic dynamics in the city of Rome and provides the similar estimated state variables shown in Figure 2 by assimilating only population data. As we found in the OSSEs, observations of the size of the human settlement G are informative to effectively constrain the flood risk model. The dynamics of the parameter estimation is similar to the case in which data of both G and H are assimilated (Figure S2).

On the other hand, assimilating only levee height data cannot provide the similar results to those shown above. Figure 4 shows the timeseries of the model variables by the data assimilation experiment in which we assimilated the observation data of H only. Observations of the levee height cannot effectively constrain D, G, and M compared with the observations of G. This finding is consistent to the OSSEs. The uncertainty in estimated parameters becomes larger when we omit to assimilate observations of G (Figure S3).

References

Ciullo, A., Viglione, A., Castellarin, A., Crisci, M., and Di Baldassarre, G.: Socio-hydrological modelling of flood-risk dynamics: comparing the resilience of green and technological systems. *Hydrological Sciences Journal*, *62*(6), 880–891. https://doi.org/10.1080/02626667.2016.1273527, 2017

Table 1. Parameters of the flood risk model

	description	Values	Ranges in data assimilation	$\boldsymbol{\omega}$ in equation (17)
$\boldsymbol{\xi_H}$	proportion of additional high water level due to levee heightening	0.5	-	-
$\boldsymbol{\alpha_H}$	parameter related to the slope of the floodplain and the resilience of the human settlement	0.01	-	-
$\boldsymbol{\rho_E}$	maximum relative growth rate	0.02	-	-
$\boldsymbol{\lambda_E}$	critical distance from the river beyond which the settlement can no longer grow	5000	-	-
$\boldsymbol{\gamma_E}$	Cost of levee raising	0.5	0.2-5.0	0.01
$\boldsymbol{\lambda_P}$	distance at which people would accept to live when they remember past floods whose total consequences were perceived as a total destruction of the settlement	12000	-	
$\boldsymbol{\varphi_P}$	rate by which new properties can be built	10000	1000-50000	100
$\boldsymbol{\varepsilon_T}$	safety factor for levees rising	1.1	-	-
$\boldsymbol{\kappa_T}$	rate of decay of levees	0.001	0-0.0015	0.0000025
$\boldsymbol{\alpha_S}$	proportion of shock after flooding if levees are risen	0.5	-	-
$\boldsymbol{\mu_S}$	memory loss rate	0.05	0-0.4	0.0025

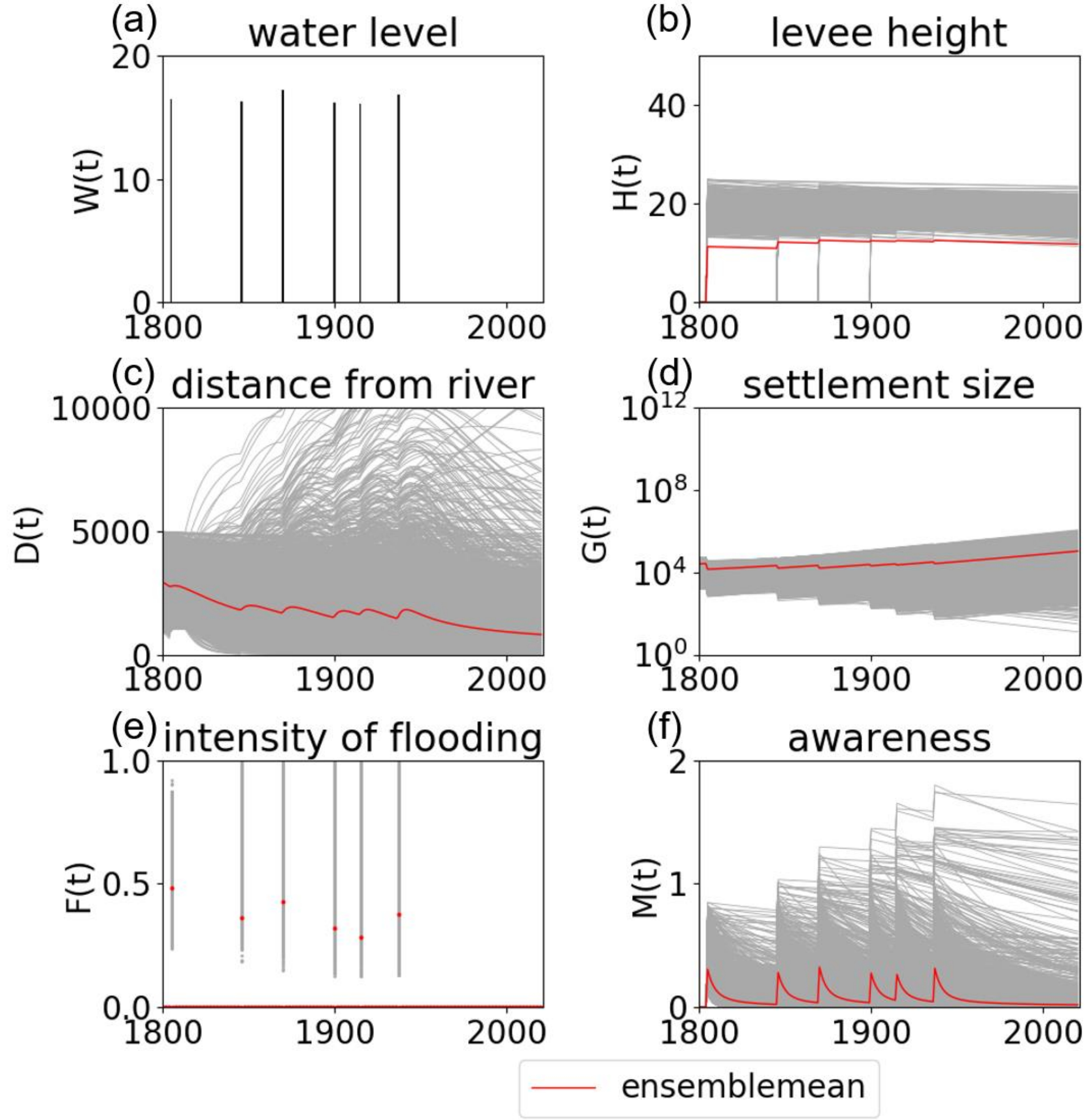

Figure 1. Timeseries of (a) high water level W(t), (b) the flood protection level (or levee height) H(t), (c) the distance of the center of mass of the human settlement from the river D(t), (d) the size of the human settlement G(t), (e) the intensity of flooding events F(t), and (f) the social awareness of the flood risk M(t) simulated by 5000 ensembles with uncertain high water levels and no data assimilation in the real-world experiment in the city of Rome. Grey, and red lines are the ensemble members and their mean, respectively.

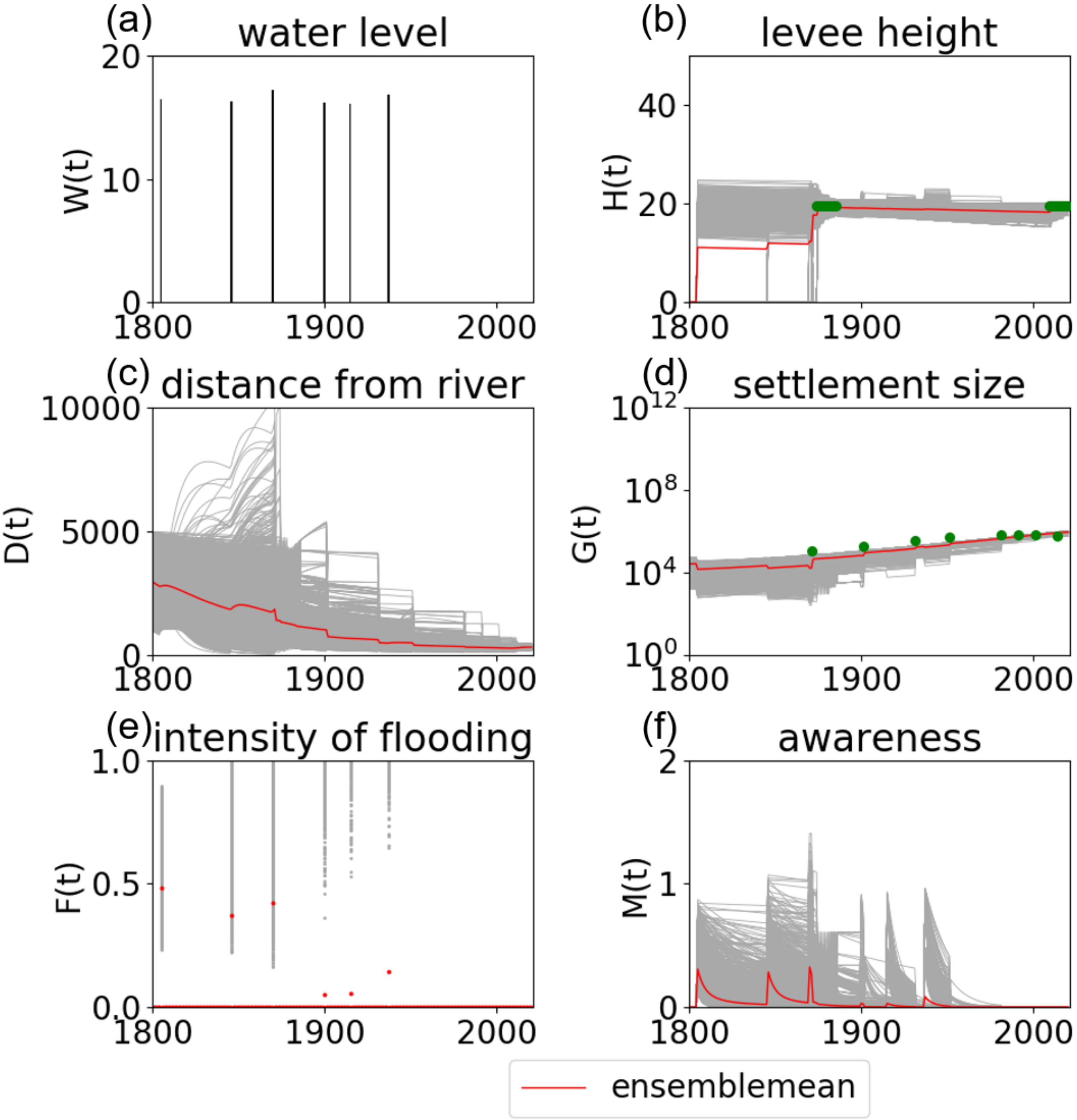

Figure 2. Timeseries of (a) high water level W(t), (b) the flood protection level (or levee height) H(t), (c) the distance of the center of mass of the human settlement from the river D(t), (d) the size of the human settlement G(t), (e) the intensity of flooding events F(t), and (f) the social awareness of the flood risk M(t) simulated by the data assimilation experiment in which the real-world observations of G and H (green dots) are assimilated

into the model with 5000 ensembles in the real-world experiment in the city of Rome. Grey, and red lines are the ensemble members and their mean, respectively.

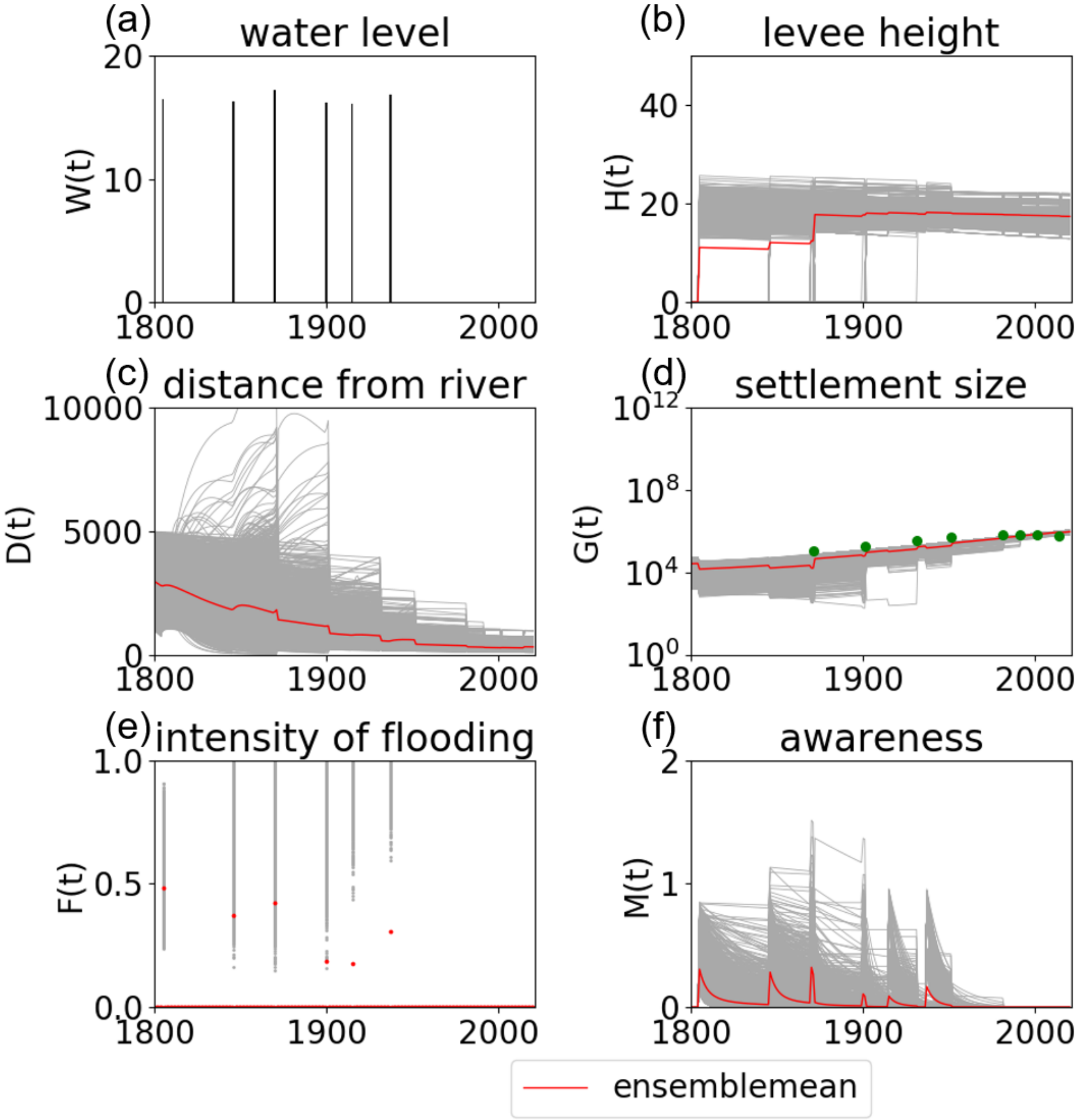

Figure 3. Same as Figure 2 but only real data of G are assimilated.

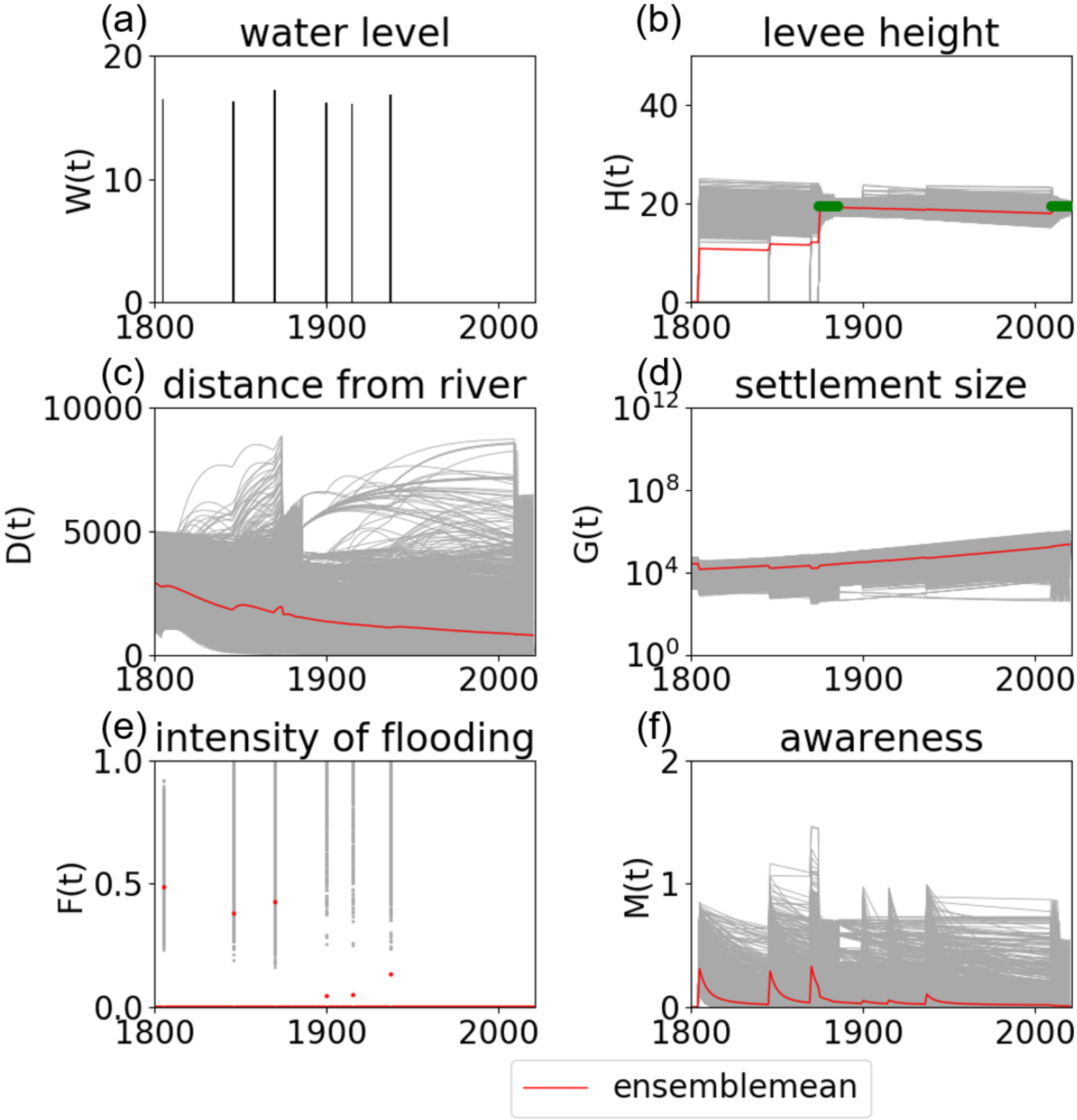

Figure 4. Same as Figure 2 but only real data of H are assimilated.

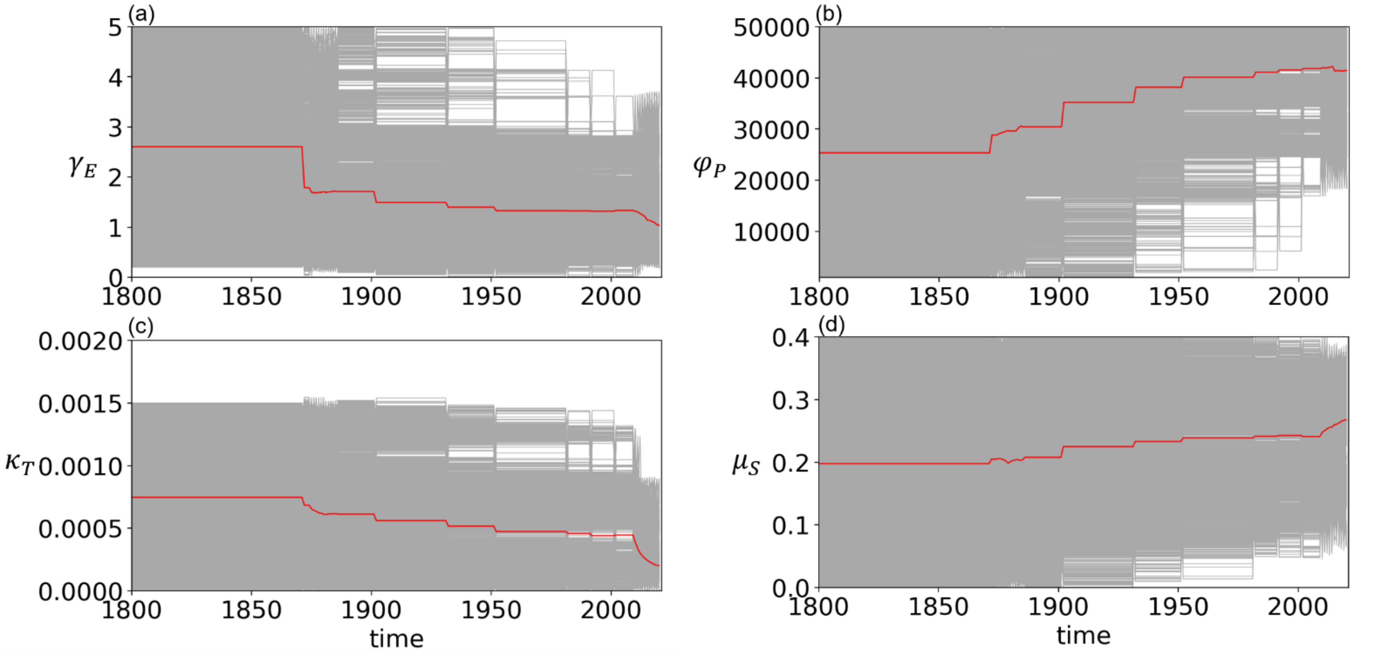

Figure S1. Timeseries of (a) the cost of levee raising γ_E, (b) the rate by which new properties can be built φ_P, (c) the rate of decay of levees κ_T, (d) memory loss rate μ_S estimated by the data assimilation of observations of G and H with 5000 ensembles in the real-world experiment in the city of Rome. Grey and red lines are the ensemble members and their mean, respectively.

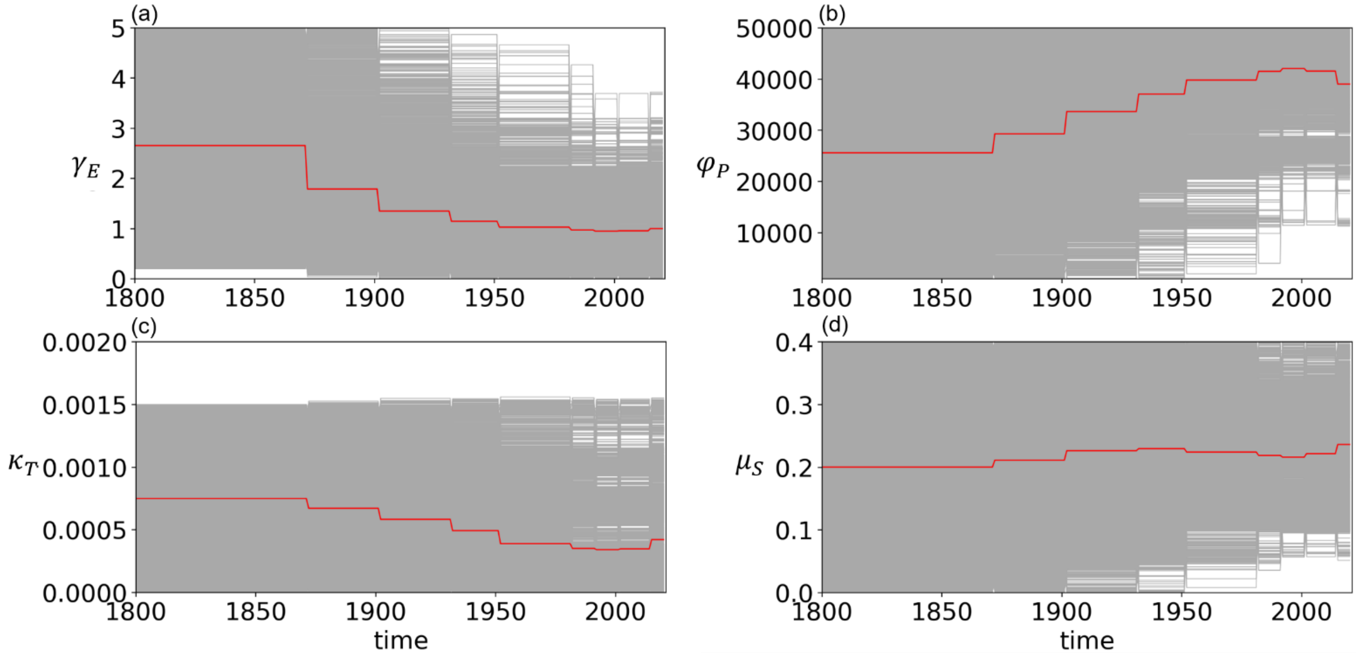

Figure S2. Same as Figure S1 but only real data of G are assimilated.

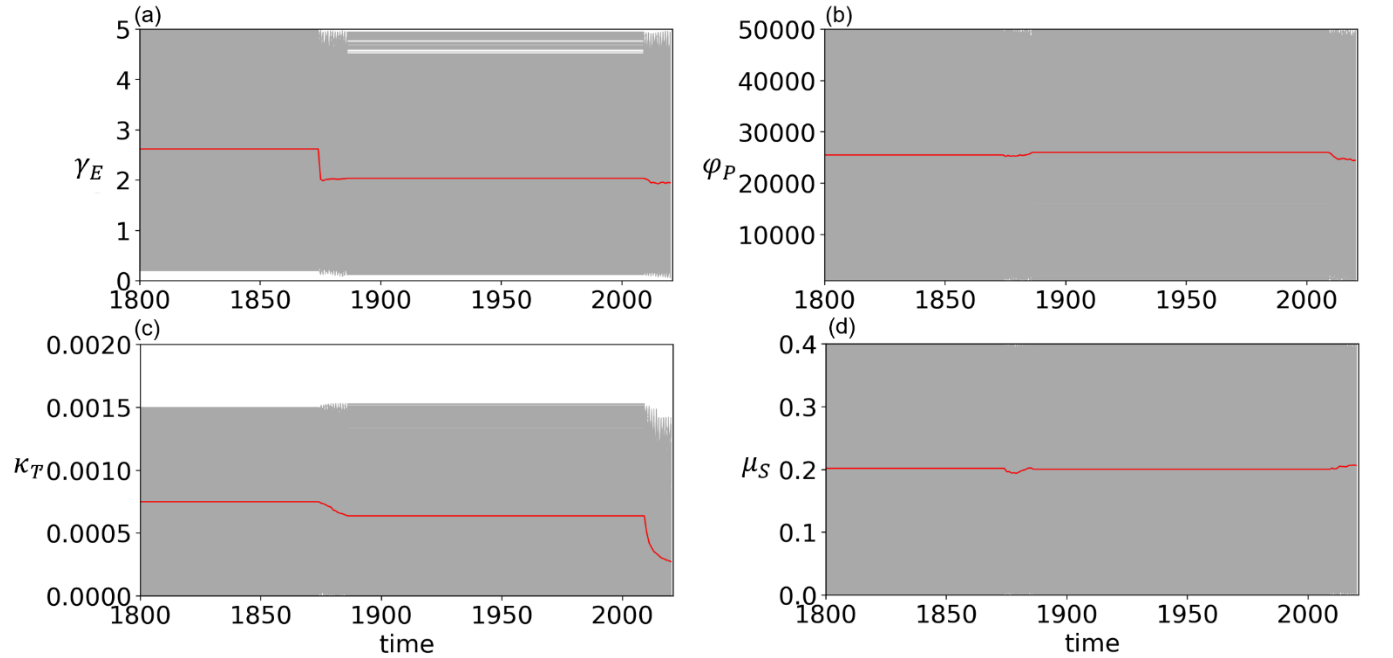

Figure S3. Same as Figure S1 but only real data of H are assimilated.