Improving pandemic forecasts by assimilating observations

Geir Evensen

NORCE

Motivation

- Provide realistic predictions with uncertainty estimates.
- Scenario modeling (e.g., importance of interventions).
- Inform public and decision makers.

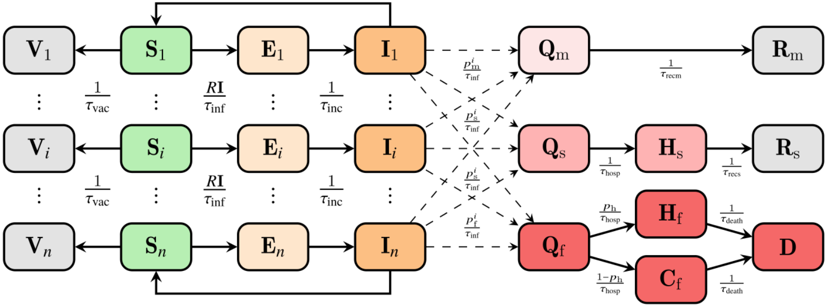
Approach

- Use an extended SEIR model.
- Condition on observed hospitalizations and deaths.
- Use ensemble data-assimilation methods for parameter estimation.

- Estimate effective reproductive number $R(t)$ as a function of time.
- The “control parameter” $R(t)$ drives the model.
- Interventions two weeks ago determines today’s deaths and hospitalizations.

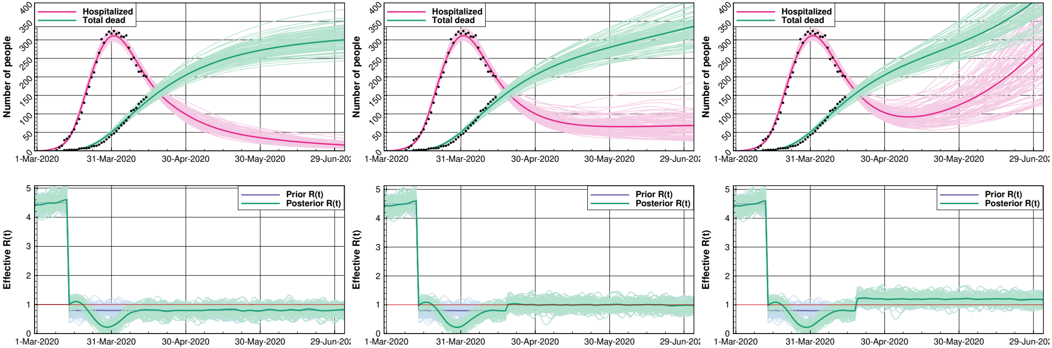
- Meterological centers use ensemble DA (EnKF) for updating weather prediction models.
- Petroleum companies use ensemble DA (IES) for history matching reservoir models.

Extended SEIR model

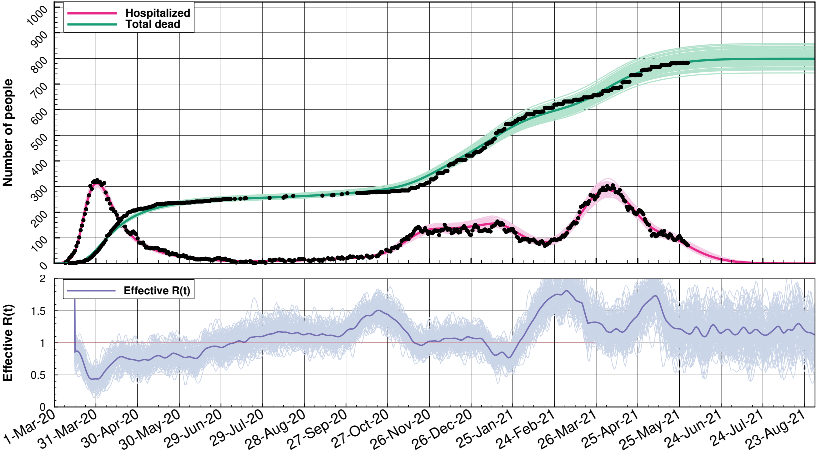


- We add age classes to model age-specific infection and death rates.
- We differentiate between mild, severe, and fatal symptoms.
- We model those with fatal symptoms who die in care homes.

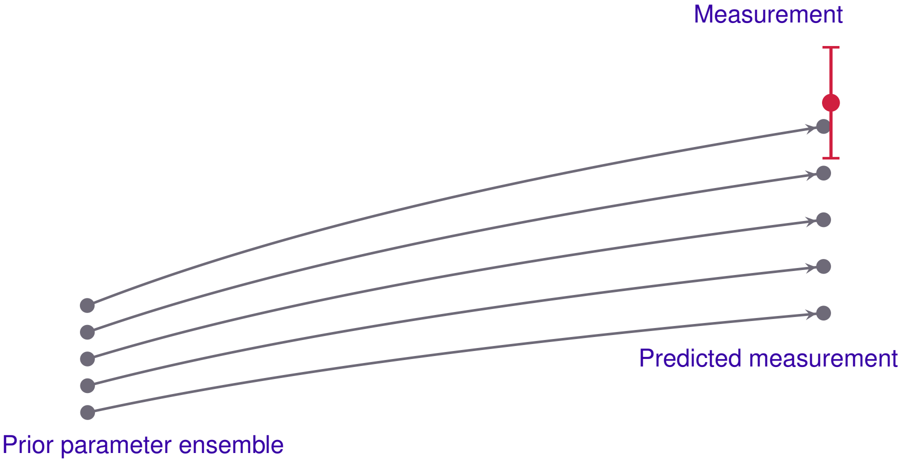
Back-to-school scenarios for Norway



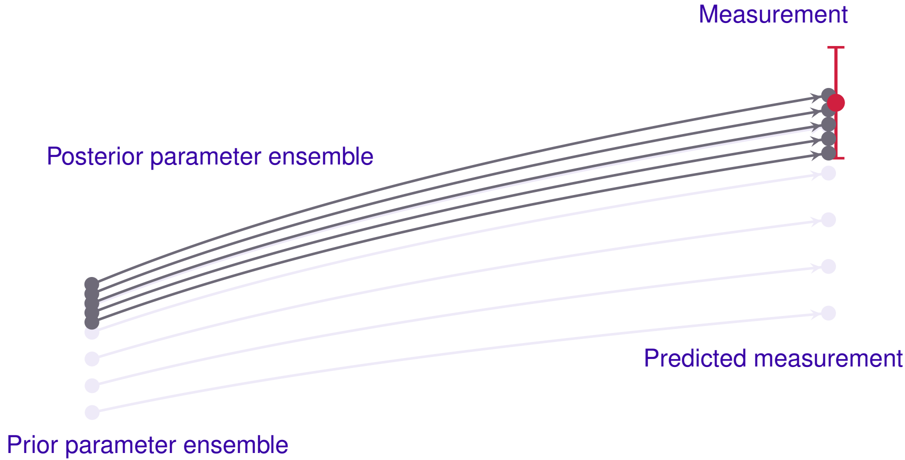
Norway: prediction including vaccinations



Prior ensemble and measurement



Regression update using ensemble correlations



Summary EnKF_seir

- The DA system tracks the epidemic accurately by estimating the past $R(t)$.
- Short-term forecasting using R-persistence works well.
- Predictions include uncertainty estimates.
- Long-term scenario forecasting with specified future R.
- Code: https://github.com/geirev/EnKF_seir
- The code supports multiple interacting "compartments."
- Plug and play model.
- Paper: (Evensen et al., 2020)
  http://www.aimsciences.org/article/doi/10.3934/fods.2021001
- Book: (Evensen et al., 2022b)
  https://link.springer.com/book/10.1007/978-3-030-96709-3

Bibliography

Evensen, G., J. Amezcua, M. Bocquet, A. Carrassi, A. Farchi, A. Fowler, P. L. Houtekamer, C. K. Jones, R. J. de Moraes, M. Pulido, C. Sampson, and F. C. Vossepoel. An international initiative of predicting the sars-cov-2 pandemic using ensemble data assimilation. *Foundations of Data Science*, page 65, 2020. doi:10.3934/fods.2021001.

Evensen, G., J. Amezcua, A. N. Carrassi, and A. Fowler. Improving pandemic forecasts: Assimilating observations and simulations. In Dhersin, J.-S., H. Kaper, W. Ndifon, F. Roberts, C. Rousseau, and G. M. Ziegler, editors, *Mathematics for action: supporting science-based decision-making*, pages 9–10. Springer, 2022a. ISBN 978-92-3-100517-6. URL `https://unesdoc.unesco.org/ark:/48223/pf0000380883.locale=en`.

Evensen, G., F. C. Vossepoel, and P. J. Van Leeuwen. *Data Assimilation Fundamentals: A Unified formulation for State and Parameter Estimation*. Springer, 2022b. ISBN 978-3-030-96708-6. doi:10.1007/978-3-030-96709-3. Open access.

Fleurantin, E., C. Sampson, D. Maes, J. Bennett, T. F. Nunez, S. Marx, and G. Evensen. A study of disproportionately affected populations by race/ethnicity during the SARS-COV-2 pandemic using multi-population SEIR modeling and ensemble data assimilation. *Foundations of Data Science*, page 72, 2021. doi:10.3934/fods.2021022.

Constant model parameters

1. Relative fractions p_{m}^{i}, p_{s}^{i}, p_{f}^{i} per age group.
2. Fractions dying in a Hospital p_{h} versus in a Care home $1 - p_{\mathrm{h}}$.

Age group	1	2	3	4	5	6	7	8	9	10	11
Age range	0–5	6–12	13–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–105
p–mild	1.00	1.00	0.99	0.99	0.97	0.96	0.93	0.90	0.84	0.81	0.81
p–severe	0.00	0.00	0.00	0.00	0.02	0.02	0.05	0.08	0.11	0.11	0.11
p–fatal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.06	0.06

Model parameters estimated by DA

Parameter	First guess	Description
τ_{inc}	5.5	Incubation period
τ_{inf}	3.8	Infection time
τ_{recm}	14.0	Recovery time mild cases
τ_{recs}	5.0	Recovery time severe cases
τ_{hosp}	6.0	Time until hospitalization
τ_{death}	16.0	Time until death
p_{f}	0.009	Case fatality rate
p_{s}	0.039	Hospitalization rate (severe cases)
I_0		Initial number of infectious
E_0		Initial number of exposed
$R(t)$		Effective reproductive number

Effective reproductive number

$$\boldsymbol{R}(t) = R(t)\hat{\boldsymbol{R}}$$

$\boldsymbol{R}(t)$ is a function of time (steered by how people isolate or interact).

- $R(t)$ is a scalar function of time.
- $\hat{\boldsymbol{R}}$ a constant matrix of transmissions between age classes..
- Behavior two weeks ago determines today's deaths and hospitalizations.
- We can estimate $R(t)$ for the past.
- We assume the value $R(t)$ for the future.

We used ESMDA

- Simple implementation and use.
- Efficient for large ensemble sizes.
- 5000 realizations and 32 ESMDA steps.