

Pseudo-code for MDLINFER

For an epidemiological model O_M , we need a calibrate procedure to learn parameterizations. Here, we assume the existence of the calibrate procedure for epidemiological model O_M and term it as CALIBRATE. The CALIBRATE takes two inputs: O_M and the dataset. We further assume the dataset contains at least reported infections time sequence D_{reported} . This CALIBRATE uses loss function to learn parameterizations, usually root mean squared error (RMSE) or log likelihood [2, 1] loss functions. We then give the pseudo-code for MDLINFER:

Algorithm 1 MDLINFER

Input: Epidemiological model O_M , calibration procedure CALIBRATE, reported infections time sequence D_{reported} .

- 1: Calibrate baseline parameterization $\mathbf{p} = \text{CALIBRATE}(O_M, \{D_{\text{reported}}, \text{others}\})$
- 2: Find $\alpha_{\text{reported}}^* = \text{GETALPHA}(O_M, \text{CALIBRATE}, D_{\text{reported}}, \mathbf{p})$
- 3: Find $D^* = \text{GETTOTALINFECTIONS}(O_M, \text{CALIBRATE}, D_{\text{reported}}, \alpha_{\text{reported}}^*, \mathbf{p})$

Output: Total infections D^*

Next, we give the pseudo-code for GETALPHA and GETTOTALINFECTIONS:

Algorithm 2 GETALPHA

Input: O_M , CALIBRATE, D_{reported} , \mathbf{p} .

- 1: The array to save the MDL cost: $\text{CostArray} = []$
- 2: **for** α_{reported} in the grid search space from 0.01 to 1 with step 0.01 **do**
- 3: $D = \frac{D_{\text{reported}}}{\alpha_{\text{reported}}}$
- 4: $D_{\text{unreported}} = D - D_{\text{reported}}$
- 5: Calibrate candidate parameterization $\mathbf{p}' = \text{CALIBRATE}(O_M, \{D_{\text{reported}}, D_{\text{unreported}}, \text{others}\})$
- 6: Save the MDL cost for α_{reported} in $\text{CostArray}[\alpha_{\text{reported}}] = L(D_{\text{reported}}, D, \mathbf{p}', \mathbf{p})$
- 7: **end for**
- 8: Find the $\alpha_{\text{reported}}^* = \arg \min_{\alpha_{\text{reported}}} \text{CostArray}[\alpha_{\text{reported}}]$

Output: Reported rate $\alpha_{\text{reported}}^*$

Algorithm 3 GETTOTALINFECTIONS

Input: O_M , CALIBRATE, D_{reported} , $\alpha_{\text{reported}}^*$, \mathbf{p} .

- 1: Find the $D^* = \arg \min_D L(D_{\text{reported}}, D, \mathbf{p}', \mathbf{p})$. (using the Nelder-Mead algorithm*).

Output: Total infections D^*

*Note: Nelder-Mead is an optimization algorithm.

References

- [1] HAO, X., CHENG, S., WU, D., WU, T., LIN, X., AND WANG, C. Reconstruction of the full transmission dynamics of covid-19 in wuhan. *Nature* 584, 7821 (2020), 420–424.
- [2] KAIN, M. P., CHILDS, M. L., BECKER, A. D., AND MORDECAI, E. A. Chopping the tail: How preventing superspreading can help to maintain covid-19 control. *Epidemics* 34 (2021), 100430.