



BY SIVA ATHREYA, GAUTAM I. MENON,  
AND RAJESH SUNDARESAN

# COVID-19 Modeling for India and a Roadmap for the Future

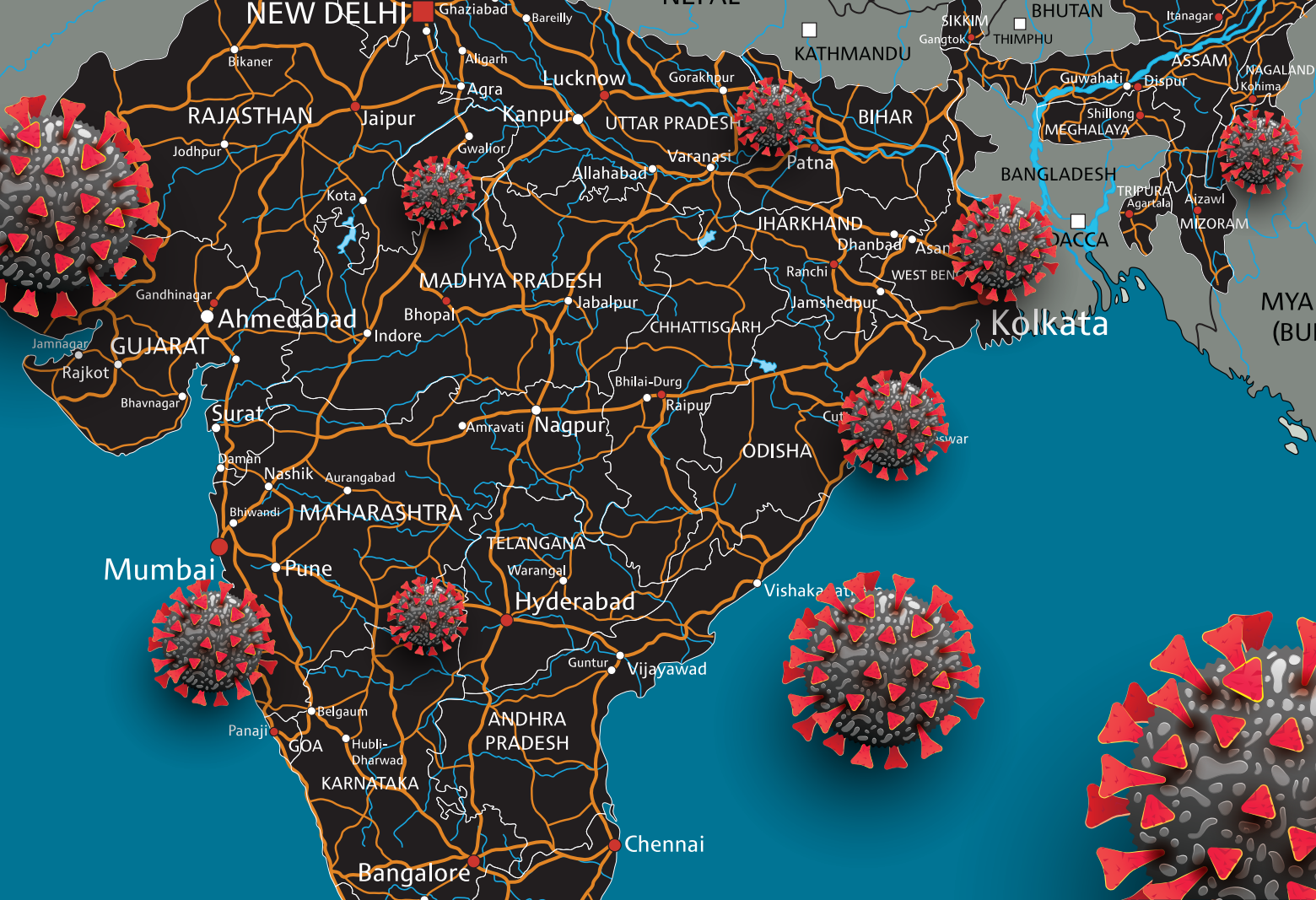
A NUMBER OF models have been developed in India to forecast the spread of the coronavirus disease or COVID-19 in the country. While these have largely been variants of the classical susceptible-exposed-infectious-recovered (SEIR) compartmental model, other approaches using time-series analysis, machine-learning, network models, and agent-based simulations have also helped to provide specific insights into questions of policy. Model building has had to incorporate our evolving knowledge of the disease, including the appearance of new variants, immune escape leading to reinfections, time-varying non-pharmaceutical interventions, the pace of the vaccination program, and breakthrough infections. The predictive power of these models has been hampered by the lack of availability of quality data

on infection and deaths as a function of age, the nature of social contacts, demography, and the clinical consequence of infection. An early emphasis on “ensemble models,” a thrust toward increased data availability, a greater engagement of modelers with the epidemiological and public health communities, and a more nuanced approach to communicating the limitations of modeling could have substantially increased the usefulness of models during the COVID-19 pandemic in India.

Most models were variants of the SEIR model where the individuals in the population move from S = susceptible to E = exposed to I = infectious to R = removed compartments. The time-evolution of the fraction of the population in each compartment is captured by a system of ordinary differential equations. Parameter estimation based on observables then enables the prediction of the future evolution of the pandemic.

Early in the pandemic, Bhatnager et al.<sup>19</sup> concluded port-of-entry restrictions would achieve modest delays in the introduction of the virus, but would not be enough to prevent the COVID-19 outbreak in India. By incorporating mobility across cities using airline data, it quantified the effectiveness of quarantining symptomatic patients. Using age-stratified social contact patterns, Singh et al.<sup>23</sup> highlighted the national three-week lockdown would not prevent a resurgence and advocated intermittent lockdowns and relaxations. Incorporating social behavioral patterns, Venkateswaran et al.<sup>25</sup> projected a similar resurgence and advocated increased testing, tracing, and isolation to mitigate the pandemic. Modeling asymptomatic transmission, a key characteristic of COVID-19, and migration, Ansumali et al.<sup>5</sup> concluded “herd immunity” is attained at 25% and that lockdown restricting migration is effective.

INDSCISIM<sup>15</sup> (see Figure 1), a more sophisticated model with nine compartments and finer stratifica-



tion for age-structured interaction, modeled lockdowns and other non-pharmaceutical interventions, testing variations, undercounting, and a range of age-specific infection fatality ratios. Recognizing significant spatial variation in the effective reproduction number, Ravinder et al.<sup>22</sup> estimated the regionwise asymptomatic and undetected symptomatic COVID-19 infected fractions. District-level infection counts were estimated in Gupta et al.<sup>14</sup> by taking into account mobility, disparate infrastructure capacities, and responses of the different states to the pandemic while Menon et al.<sup>20</sup> studied the effectiveness of a few non-pharmaceutical interventions and Ghosh et al.<sup>12</sup> provided 30-day shorter-term forecasts using a simpler SIS model.

The mechanistic ODE models in the works described here are large system limits of the stochastic agent-based models where the agents in the system interact with each other through their contact networks (household, workplace, schools, markets, transport spaces, among others).

The interactions lead to the spread of infection in the population. Focusing on city-scale models for Bengaluru and Mumbai, Agarwal et al.<sup>4</sup> explored the impact of various finer-grain unlocking scenarios. Considering a flu-like confounding illness coexisting with COVID-19, Gopalan et al.<sup>13</sup> studied the impact of several testing policies and showed that location-based random symptomatic testing can capture the ground truth quite closely with very few daily tests. Recognizing that cities are spatially organized, Bhattacharyya et al.<sup>7</sup> introduced a multilattice small world network for the city and showed local lockdowns restricted to the infected wards are effective compared to full-scale lockdowns. For such computationally intensive agent-based simulations, Agrawal<sup>4</sup> and Kshirsagar et al.<sup>18</sup> provide frameworks for parallelization and efficient implementation. A naïve implementation taking every edge's effect in account would have taken  $O(N^2)$  computations per time-step, where  $N$  is the size of the population. These frameworks, at the expense of not being able to tell from

whom an infected individual caught the infection, reduce the computations to a manageable  $O(N)$  complexity per time-step. The mechanistic ODE models have  $O(1)$  complexity independent of the population size, and are hence scalable. The constant of course will depend on the number of compartments.

While these examples provide a description of the modeling work coming from India, some example models from the South Asian region include Chowdhury et al.,<sup>28</sup> which is a neural network-based ANFIS and LSTM model for Bangladesh, Ali et al.,<sup>29</sup> which is an autoregressive integrated moving average model for Pakistan, and Jayatilke et al.,<sup>30</sup> which is an SIR model for Sri Lanka with the contact rate modulated by a mapped stringency index. There are many more models for the region from researchers in other parts of the world that we have not discussed in this article.

### Challenges

All the models described here relied on published data of infected, recov-

ered, and deceased counts at city/district/state levels for calibration. Clearly, the model projections had significant variations among themselves and with reality. The key challenges for predictive power were:

► **Data availability:** Indian Council of Medical Research (ICMR) maintains a central repository with all cases data. Each state used this along with recovered and deceased data from hospital networks to publish district-wise daily numbers in varying formats. It was left to a voluntary effort, <https://www.covid19india.org> (see Figure 2), to provide the data in a standardized

machine-readable format.

► **Data quality:** The reported data were of varying granularity—some states reported at the district level and daily while some others did not. The metadata associated with the test cases were noisy because of significant variations in the nature of data collection (symptoms, vaccination, date and time of sample, test process variability, delay in reporting).

► **Evolving response:** The union government and the states' responses continuously evolved as the country-wide lockdowns during March–May 2020 were followed by decentralized

nonpharmaceutical and pharmaceutical interventions at the state level.

► **Evolving testing capacity and policy:** Detection of cases is highly dependent on testing capacity, which was seriously limited in the beginning but improved as the pandemic progressed. Further, testing policy resulted in varying tested numbers over time.

► **Evolving virus:** The virus continuously mutated resulting in newer variants with different transmission rates and immune escape properties.

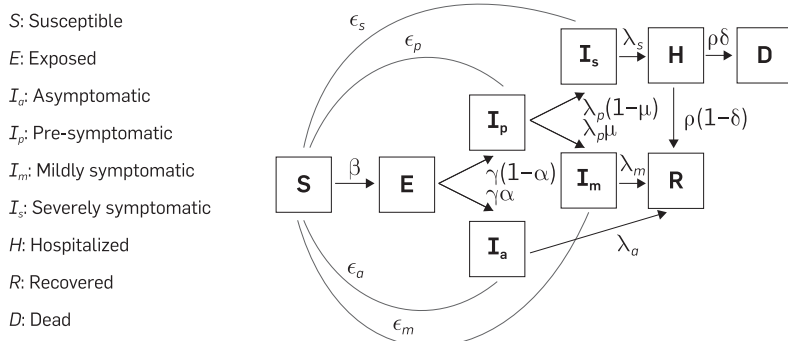
Given these challenges, models varied significantly in their projections and performances. The heightened public attention on modeling coupled with the difficult task of communicating uncertainty quantification to the general public raised the expectations of an accurate match with reality to unrealistic levels. All these adversely impacted the adoption of models for devising effective public health responses.

**Approaches**

**Selection.** The variety of models with varying assumptions and projections led the Department of Science and Technology, Government of India, to consolidate the plethora of models into one robust 'supermodel' that could be "subjected to rigorous tests required for evidence-based forecasting, routinely practiced in weather forecasting communities."<sup>9</sup> The coordination team, see Vidyasagar et al.,<sup>26</sup> zeroed in on one model,<sup>21</sup> which was a variation on Ansumali et al.<sup>5</sup> The 'supermodel' predicted the pandemic had peaked at the all-India level in late September 2020. The first wave and the number of modelers waned over the next few months.

**Ensemble.** Elsewhere, particularly in the U.S. and Europe, the respective centers for disease control took a different approach that fostered the development of 'forecast hubs'<sup>16</sup> and leveraged the modeling community's expertise and enterprise. Such hubs served as platforms for modelers to submit models and projections for both the short-term and the long-term. The hub, which was launched in April 2020, collected forecasts from over 82 modeling teams into a data repository and made them easily accessible for comparison, evaluation, and guid-

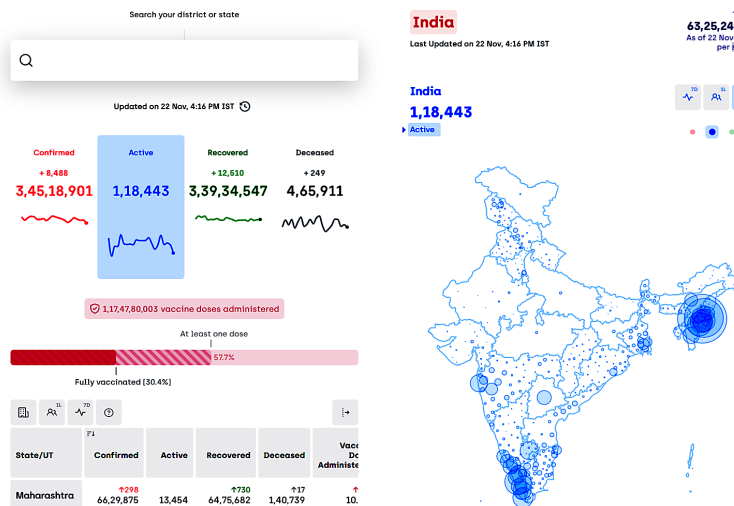
**Figure 1. The INDCISIM model with several compartments, transitions, and associated transition rates. Given the large number of parameters, a Bayesian method was used to obtain optimal fits to reported cases and deaths.**



Rates are indicated against the arrows. The  $\epsilon$ s represent relative intensity of contacts.

**Figure 2. States of India reported via daily media bulletins.**

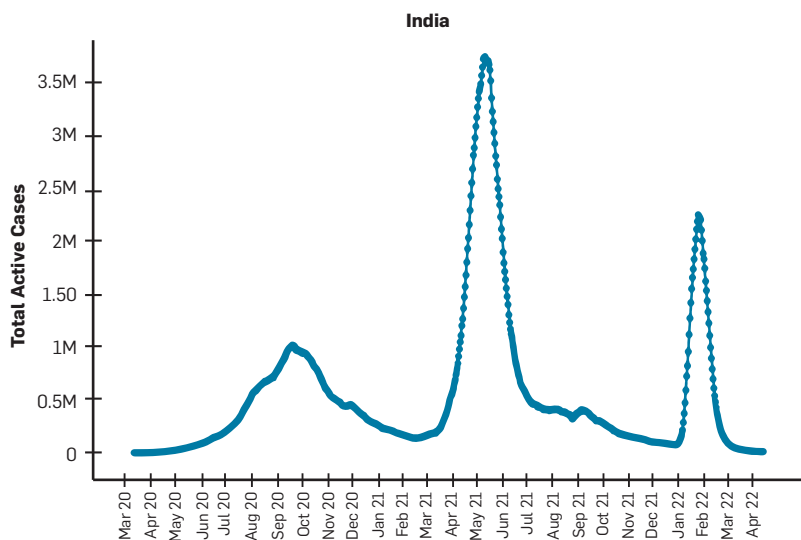
Due to varying formats a separate effort was needed to scrub and present the data in a standardized machine-readable format. This was done by volunteers at <https://www.covid19india.org> that provided a very valuable resource of district-level data. Currently, <https://www.incovid19.org> continues this valuable effort.





**Figure 3. COVID-19 active cases in India.**

The sharp rise in the second wave due to the Delta variant was missed by models. For the third wave due to the Omicron variant, models did predict the sharp rise.<sup>17</sup>



**Models varied significantly in their projections and performances.**

ance on response efforts. This “collect and aggregate” ensemble approach encouraged modelers to stay engaged and adapt their models to the evolving pandemic. (See the Scientific Advisory Group for Emergencies<sup>27</sup> in the U.K. where modeling groups provided good quality monthly projections and influenced policy.)

*New models November 2020–March 2021.* Despite the zeroing in on the ‘supermodel,’ a few others emerged with a focus on the short-term. CSIR-4PI<sup>8</sup> used machine learning techniques, IISC–ISI used a rudimentary log-linear fit model, and Ansumali et al.<sup>5</sup> adapted to provide short-term forecasts. For the longer-term projections, the DST ‘supermodel’ also evolved into SUTRA<sup>3</sup> relaxing the assumption that all symptomatic persons are detected and introducing a reach parameter to account for a growing “effectively involved” population. A PDE model was used in Ganesan et al.<sup>11</sup> to address spatial heterogeneity.

**Failure.** Even though the alpha variant was active in the U.K. by the end of 2020, with cases tapering off in India, talk of hygiene hypothesis, and innate immunity, complacency had set in. When the second wave hit the country in April 2021 (see Figure 3), there was no warning from the modeling community. We did not anticipate

the impact of variants, did not predict the greater severity of illness that affected hospitalization estimates, did not consider antibody waning, and did not fully comprehend the size of the susceptible population. It must be noted the immunization efforts in India began only in January 2021.

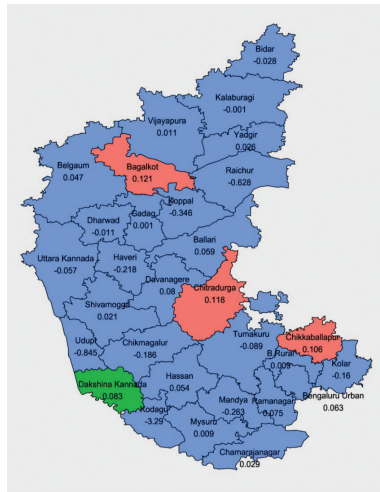
**Newer efforts.** Spurred by the failure, the modeling effort was renewed. Several mitigation strategies for Karnataka with inclusion of mobility, antibody waning, vaccination, and its interplay with non-pharmaceutical interventions, were studied in Adiga et al.<sup>1</sup> to provide longer-term projections including the possibility of a third wave and strategies for allocating vaccines across districts. INDSCISIM, SUTRA, the PDE model,<sup>11</sup> and CSIR-4PI<sup>8</sup> were revived, along with Dukkupati et al.,<sup>10</sup> another machine-learning model based on the Hawkes process. However, the ensemble framework remained to be done.

### Some Newer Approaches

In view of the severity of the second wave in April 2021, the need for an early warning system was acutely felt. Toward this, Athreya et al.<sup>6</sup> built an early-warning system that considered district and state healthcare infrastructure capacity and provided warnings of upcoming surges when the

**We should have a National Forecast Hub featuring an ensemble of models with regular updates from the modelers.**

**Figure 4. Early warning system for Karnataka built in early January 2022. See Athreya et al.<sup>6</sup> for details of the method and <https://www.isibang.ac.in/~incovid19/> for current status in Karnataka.**



cases were still low. It used an elementary first-order method that relied only reported data from a district and called out a warning when the local growth rate exceeding the recovery rate. Since the goal was to report upsurges quickly, the averaging window was kept small which led to false alarms. Nevertheless, the approach provided useful guidance to the state of Karnataka on the surging districts (see Figure 4). An alternative projection method was based on a comparison of the speed of the omicron variant's surge with the delta variant's

surge in South Africa. While South Africa and India have very different social contact structures, the advantage of omicron over delta could be taken to be an invariant biological factor applicable to both South Africa and India. This suggested a way to project the omicron peaks in India (see Figure 5) and provided a means to assess the impact of weeknights and weekend curfews.<sup>2</sup> Alongside these efforts, an Indian version of the forecast hub<sup>16</sup> collated projections from CSIR-4PI, IISc-ISI log-linear models, machine learning models, time-series models, and a version of SUTRA (see Figure 6).

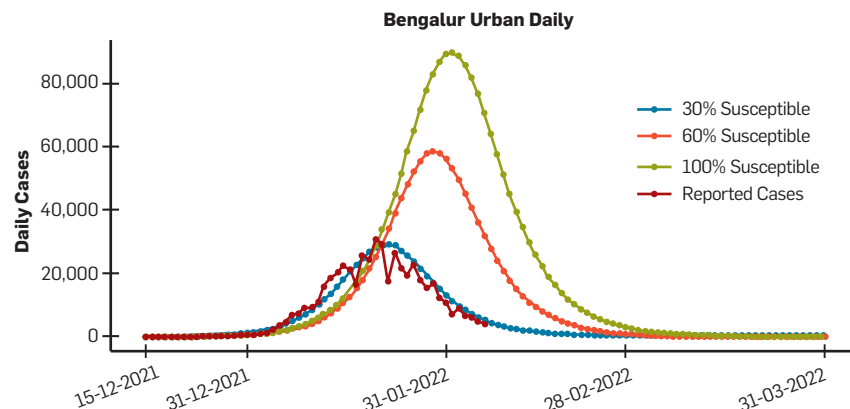
**Roadmap for the Future**

We have surveyed a subset of computational epidemiological modeling efforts for COVID-19 arising out of India. Our main message is we should have a National Forecast Hub featuring an ensemble of models with regular updates from the modelers. The differing assumptions and the varying predictions, all made available in one place, will enable better communication of uncertainty and a greater understanding of the applicability and the limitations of the individual models.

On another front, states/union-territories share testing data with ICMR. However, access to this (600+ million tests) is limited, resulting in only a few research outcomes and publications. The data is anticipated to be noisy. But the country has considerable expertise

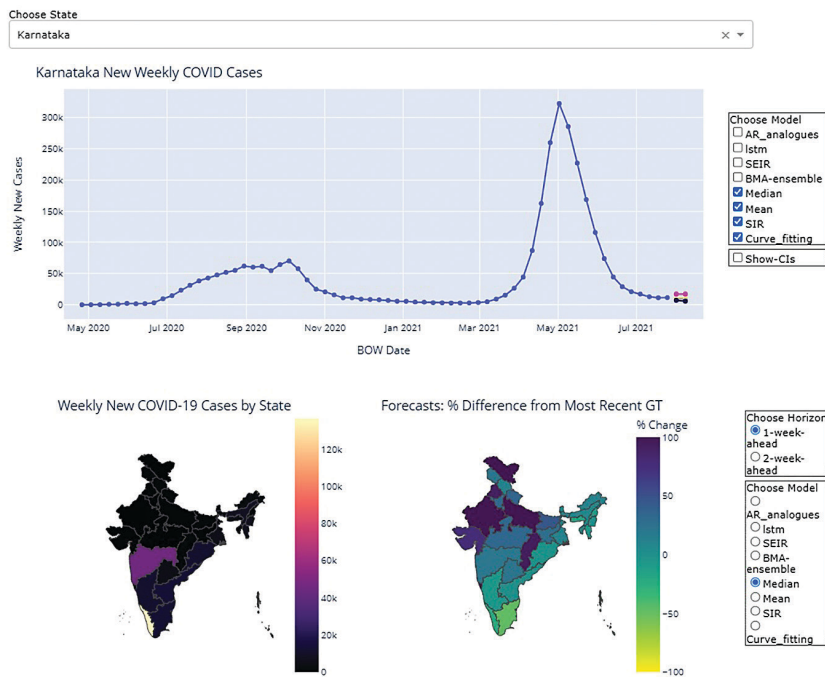
**Figure 5. The model for COVID-19 cases due to Omicron variant in Bengaluru Urban district<sup>17</sup> using a simple compartmental model.**

The entire analysis is based on the spread of Omicron variant in South Africa during the weeks 47–52 of 2021. The same data was used to provide the estimated requirements of hospital beds and numbers among those that require intensive ventilation.




**Figure 6. India forecast hub<sup>16</sup> has been developed to provide a common platform for modeling and forecasting teams to contribute short-term COVID-19 incident case forecasts for the states of India. The goal is to enable effective communication of available forecasts to both the public and the policymakers.**

#### COVID-19 Forecast Dashboard



in statistical and machine learning to mine this data and gather insights. Further, this can be integrated with serosurvey data, genome sequencing, clinical data from the National Center for Disease Control, and mobile health data from platforms such as Aarogya Setu to build meaningful predictions and design better-targeted data-driven and evidence-based public health responses. These could include the treatment protocol, sampling strategy for sequencing and testing, non-pharmaceutical interventions, and vaccination strategies. We should also revisit previous waves to validate the models. To enable all this, it is imperative that India commits to data independence and to making data publicly available in a standardized format post anonymization.

We have not addressed many important items: assessing deaths due to COVID-19, sampling, testing, and sequencing protocols for sentinel surveys for enabling timely alerts, and biosurveillance through continuous wastewater testing. Modeling will be of enhanced quality if these are incorporated.

**Acknowledgments.** This work was partially supported by the Centre for Networked Intelligence, the Institute of Eminence grant at the Indian Institute of Science to support an Indo-U.S. COVID-19 response effort, the SERB-MATRICES grant, and the CPDA grant at the Indian Statistical Institute. 

#### References

- Adiga, A. et al. Strategies to Mitigate COVID-19 Resurgence Assuming Immunity Waning: A Study for Karnataka, India. *medRxiv*, (Jan. 1, 2021).
- Adiga, A. et al. Impact of weeknight and weekend curfews using mobility data: A case study of Bengaluru Urban. *medRxiv*, (Jan. 28, 2022).
- Agrawal, M., Kanitkar, M., Vidyasagar, M. SUTRA: A Novel Approach to Modelling Pandemics with Applications to COVID-19. Jan. 22, 2021; arXiv:2101.09158.
- Agrawal, S. et al. City-scale agent-based simulators for the study of non-pharmaceutical interventions in the context of the COVID-19 epidemic. *J. Indian Institute of Science 100*, 4 (Oct. 2020), 809–847.
- Ansumali, S., Kaushal, S., Kumar, A., Prakash, M.K., Vidyasagar, M. Modelling a pandemic with asymptomatic patients, impact of lockdown and herd immunity, with applications to SARS-CoV-2. *Annual Reviews in Control 50* (Jan. 1, 2020) 432–447.
- Athreya, S., Sarkar, D., Sundaresan, R. Estimated growth rate of active infections in Karnataka by district (Feb. 16, 2022); <https://www.isid.ac.in/~deepayan/covid-prediction/active-prediction-karnataka.html#1>.
- Bhattacharyya, C., Vinay, V. Suppress, and not just flatten: Strategies for Rapid Suppression of COVID-19 transmission in Small World Communities. *J. Indian Institute of Science 100*, 4 (Oct. 2020), 849–862.
- Bhimata, K.R., Patra, G.K., Mopuri, R., Muthneni, S.R. Prediction of COVID-19 cases using the weather integrated deep learning approach for India. *Transboundary and Emerging Diseases* (Apr. 10, 2021).

- COVID-19 India National Supermodel; <https://dst.gov.in/dst-initiates-covid-19-india-national-supermodel-monitoring-infection-transmission-aid-decision>.
- Dukkipati, A., Gracious, T., Gupta, S. CoviHawkes: Temporal Point Process and Deep Learning based Covid-19 forecasting for India. Sept. 8, 2021; arXiv:2109.06056.
- Ganesan, S., Subramani, D. Spatio-temporal predictive modeling framework for infectious disease spread. *Scientific Reports 11*, 1 (Mar. 24, 2021) 1–8.
- Ghosh, P., Ghosh, R., Chakraborty, B. COVID-19 in India: statewise analysis and prediction. *JMIR Public Health and Surveillance 6*, 3 (Aug. 12 2020), e20341.
- Gopalan, A., Tyagi, H. How reliable are test numbers for revealing the COVID-19 ground truth and applying interventions? *J. Indian Institute of Science 11*, 4 (Oct. 2020), 863–884.
- Gupta, S. et al. India-specific compartmental model for Covid-19: projections and intervention strategies by incorporating geographical. *Infrastructural and Response Heterogeneity*; arXiv:2007.14392.
- Hazra, D.K. et al. The INDSCI-SIM model for COVID-19 in India. *medRxiv* (Jan. 1, 2021).
- India COVID-19 Forecast Hub; <https://www.isibang.ac.in/~incovid19/dash.php>.
- India COVID-19 Omicron Projections; <https://www.isibang.ac.in/~incovid19/>.
- Kshirsagar, J., Dewan, A., Hayatnagar, H. Epirust: Towards a Framework for Large-Scale Agent-Based Epidemiological Simulations Using Rust Language. EasyChair, Sept. 9, 2020.
- Mandal, S. et al. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach. *The Indian J. Medical Research 151*, 2–3 (Feb. 15, 2020), 190.
- Menon, A., Rajendran, N.K., Chandrachud, A., Setlur, G. Modelling and simulation of COVID-19 propagation in a large population with specific reference to India. *medRxiv*. (Jan. 1, 2020).
- National Supermodel Committee. Indian supermodel for Covid-19 pandemic; [https://www.iith.ac.in/~m\\_vidyasagar/arXiv/Super-Model.pdf](https://www.iith.ac.in/~m_vidyasagar/arXiv/Super-Model.pdf).
- Ravinder, R., Singh, S., Bishnoi, S., Jan, A., Sharma, A., Kodamana, H., Krishnan, N.A. An adaptive, interacting, cluster-based model for predicting the transmission dynamics of COVID-19. *Heliyon 6*, 12 (Dec. 1, 2020), e05722.
- Singh, R., Adhikari, R. Age-structured impact of social distancing on the COVID-19 epidemic in India. Mar. 26, 2020; arXiv:2003.12055.
- Talekar, A. et al. Cohorting to isolate asymptomatic spreaders: An agent-based simulation study on the Mumbai Suburban Railway. ArXiv, (Dec. 23, 2020).
- Venkateswaran, J., Damani, O. Effectiveness of testing, tracing, social distancing and hygiene in tackling COVID-19 in India: a system dynamics model. Apr. 19, 2020; arXiv:2004.08859.
- Vidyasagar, M., Agrawal, M., Kanitkar, M., Bagchi, B., Bose, A., Kang, G., Pal, S.K. Progression of the COVID-19 pandemic in India: Prognosis and lockdown impacts.
- Scientific Advisory Group for Emergencies. SPI-M-O Medium Term Projections (Feb. 9, 2022); <https://www.gov.uk/government/publications/spi-m-o-medium-term-projections-9-february-2022>.
- Chowdhury, A.A. et al. Analysis and prediction of COVID-19 pandemic in Bangladesh by using ANFIS and LSTM network. *Cognitive Computation 13*, 3 (May 02), 761–770.
- Ali, M. et al. Forecasting COVID-19 in Pakistan. *PLoS One 15*, 11 (Nov. 30, 2020), e0242762.
- Jayatilake, A. et al. COVID-19 case forecasting model for Sri Lanka based on Stringency Index. *medRxiv* (Jan. 1, 2020).

**Siva Athreya** is a professor at the Indian Statistical Institute, Bengaluru, India.

**Gautam I. Menon** is a professor at The Institute of Mathematical Sciences in Chennai, at the Homi Bhabha National Institute in Mumbai, and at Ashoka University in Haryana, India.

**Rajesh Sundaresan** is a professor in the ECE Department, an associate faculty in the Robert Bosch Centre for Cyber Physical Systems, and a member of the Cisco-IISc Centre for Networked Intelligence at the Indian Institute of Science in Bengaluru, India.

© 2022 ACM 0001-0782/22/11 \$15.00