

INFORMATION PROCESSING METHODOLOGIES TO COMBAT THE COVID-19 PANDEMIC

Abstract—Information and signal processing tools are crucial for interpreting coronavirus disease 2019 (COVID-19) pandemic data. These tools allow us to extract, synthesize, and interpret pandemic information, thus providing valuable support to the decision-making authorities. This paper presents an overview of recent advances in information processing methodologies to combat the COVID-19 pandemic. First, we describe the quickest detection procedure designed to detect an exponential growth of positive cases with a mean delay of only a few days and a low risk of erroneously declaring an outbreak. Second, we present a Bayesian approach designed to estimate some features of the pandemic, e.g., the infection rate, and reliably forecast the evolution of the contagion.

INTRODUCTION

Since the beginning of 2020 and up to the end of April 2021, the virus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), responsible of the coronavirus disease 2019 (COVID-19) respiratory illness, has infected more than 150 million individuals worldwide and caused the death of more than 3.2 million people. Because of its rapid human-to-human transmission and the presence of highly infectious asymptomatic individuals, the COVID-19 disease was declared by the World Health Organization to be a pandemic on March 11, 2020. Since then, many governments, pushed by the lack of an effective therapy and the imperative need of containing the contagion, have decided to undertake unprecedented extraordinary social measures that have changed many aspects of our lives. These measures, which included travel bans; closure of schools, universities, shops, and factories; and even national lockdowns, effectively slowed the spread of the virus; however, their early relaxation has been causing the recrudescence of the contagion almost everywhere. The implementation of massive vaccination campaigns represents the only way to definitely defeat the pandemic, as shown, e.g., from the evolution of the contagion in Israel, where 85% of individuals older than 60 years had been fully vaccinated after only 2 months into the vaccination campaign [1]. Nevertheless, as researchers are still debating whether new variants can undercut the effectiveness of these first-generation COVID-19 vaccines [2], it is of paramount importance to remain vigilant and assist the authorities in evaluating the implementation of pandemic countermeasures.

Information and signal processing tools, exploiting the vast amount of data collected since the beginning of 2020, can support decision makers in monitoring the contagion and predicting the evolution of the pandemic [3]. This article excerpts from [4]–[8], providing an overview of recent advances in information processing methodologies to combat the COVID-19 pandemic.

QUICKEST DETECTION OF PANDEMIC WAVES

One aspect to monitoring the COVID-19 pandemic is to detect, as quickly as possible, the outbreak of a new exponential growth of positive cases, which would allow governments and authorities to react in a timely manner [8]. Indeed, on the one hand, the early application of countermeasures, such as social distancing and closure of commercial activities, can save lives; in this context, the delay of intervention needs to be as short as possible. On the other hand, an incorrect detection of an outbreak (i.e., a false alarm), and the consequent imposition of unnecessary restrictive measures, may have huge trust, societal, and economic costs [9]. This risk, mathematically defined as the reciprocal of the mean time between two consecutive false alarms, needs thus to be extremely small.

Leveraging quickest detection theory [10], [11], we have developed in [4], [5], and [6] a variation of the celebrated Page's test [12], called the mean-agnostic sequential test (MAST), designed to detect the transition from a controlled regime of the pandemic, characterized by a limited number of daily new positive cases, to a critical regime, in which the infection spreads exponentially fast. MAST is based on the recursive computa-

Domenico Gaglione

NATO STO Centre for Maritime
Research and Experimentation
(CMRE)
Research Department
La Spezia, Italy
domenico.gaglione@cmre.nato.int

Paolo Braca

Giovanni Soldi

Nicola Forti

Leonardo M. Millefiori

NATO STO Centre for Maritime
Research and Experimentation
(CMRE)
Research Department
La Spezia, Italy

Stefano Marano

University of Salerno
Dipartimento di Ingegneria
dell'Informazione ed Elettrica e
Matematica Applicata (DIEM)
 Fisciano (SA), Italy

Peter K. Willett

Krishna R. Pattipati

University of Connecticut
Department of Electrical and
Computer Engineering
Storrs, CT, USA

tion of a decision statistic that depends only on the observed growth rate, computed daily as the ratio between two consecutive new positive counts (details in [5]). This statistic is then compared to a threshold—selected to trade off mean decision delay and risk—and, if it is crossed, an outbreak is declared.

An extensive analysis of MAST when applied to data of the COVID-19 second wave from several countries is presented in [4]. Here, we report in Figure 1 the operational curve—risk versus mean delay—for the 14 countries considered. We observe that for a reasonable risk, e.g., 10^{-4} days⁻¹, which means that an erroneous decision is made, on average, once every 27 years, the mean delay in declaring the onset of the second wave is always less than 20 days. Additional analyses are available in [13].

MAST was shown in [6] to also be effective when used on data from a smaller community (e.g., a region or a province) and, after minor modifications, for detecting the termination of a pandemic wave [14]; this information may be crucial, e.g., to safely relax the restrictive measures. As an example, we report in Figure 2 the growth rate computed for the Lombardia region of Italy from February 24, 2020, to February 26, 2021, and in Figure 3 the MAST statistics used to detect, in the same time interval, the onset of the second and third pandemic waves and the termination of the second wave. Here, the value of the threshold corresponds to a risk of 10^{-5} days⁻¹, which means that a false change of regime is declared, on average, once every 270 years. We observe that the onset of the second wave is declared on August 20, 2020, with its termination on December 3, 2020. A third wave is detected on February 25, 2021. Exact dates on which the second and third waves began are clearly not available. Nevertheless, we can argue that the detection delay is reasonably small for the third wave [15], whereas the second wave is largely anticipated by MAST, with the restrictive measures implemented only more than 2 months after August 20, 2020 [16].

ADAPTIVE FORECAST OF THE INFECTION

Once an exponential growth of positive cases is detected, to understand how the pandemic will evolve is essential information for policymakers to plan their future actions, e.g., increase hospital bed capacity and relocate health care personnel. Equally important is the forecast of the infection in a controlled regime that can support planning for the gradual reopening of commercial, industrial, and social activities.

In the introduction to the third volume of *ISIF Perspectives on Information Fusion* in May 2020 [17], Dr. Streit rightly foresaw that “two areas of research will naturally beckon for our attention in the coming days and months. One area is the mathematical modeling of the spread of infectious diseases” that “began in the 1920s with differential equation compartmental models of the numbers of Susceptible/Infectious/Recovered (SIR) individuals. [...] The other area concerns spatial-temporal data modeling”. Compartmental epidemiological models are—still today—commonly used to study the spread of infectious diseases. They assume that a given population is partitioned into a predefined number of compartments (population subgroups), in which each compartment represents a pandemic state that an individual can

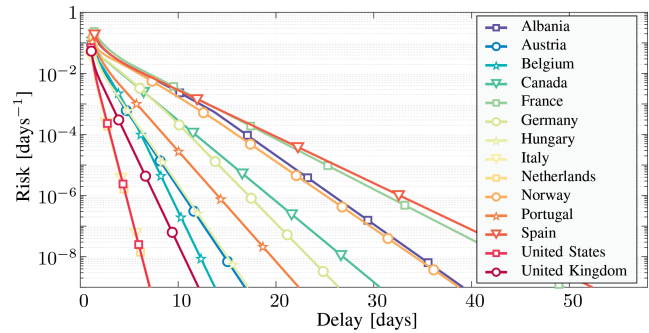


Figure 1
From [4], operational curve—risk versus mean delay for decision—for 14 countries.

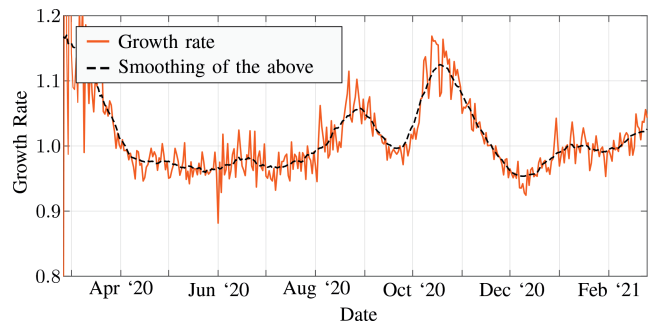


Figure 2
From [6], growth rate of the pandemic in the Lombardia region of Italy, computed from the averaged daily new positive cases (orange solid line) from February 24, 2020, to February 26, 2021; for easier visualization, we also show its smoothed version obtained through a noncausal moving average filter with uniform weights of a length of 21 days (black dashed line).

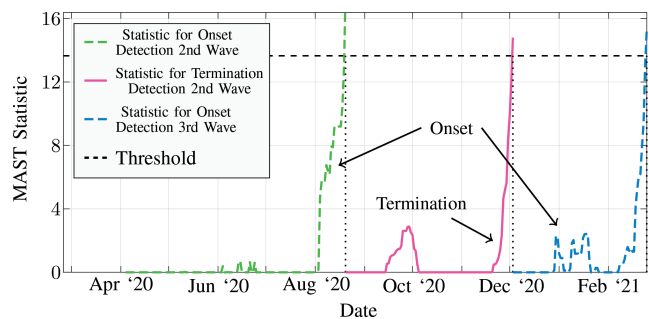


Figure 3
From [6], MAST statistics computed for the Lombardia region of Italy, starting from April 4, 2020, for the onset detection of the second wave (green dashed line) and the third wave (blue dashed line) and for the termination detection of the second wave (magenta solid line). The threshold (black dashed line) corresponds to the risk of 10^{-5} days⁻¹. The onset and termination of the second wave are declared on August 20 and December 3, 2020, respectively. The onset of the third wave is declared on February 25, 2021.

occupy, and the flow dynamics from one compartment to another are modeled as a set of differential equations [18]. The pioneering study on mathematical theory of epidemics mentioned by Streit dates back to 1927 and proposes an epidemiological model in which the entire population of, e.g., a city, a region, or a nation is constant and divided into three mutually exclusive compartments, namely, susceptible (S), infected (I), and recovered (R) individuals, and known as the SIR model [19], [20]. An infected individual infects a susceptible one at a given infection rate β . Once infected, the individual is removed from the compartment of susceptible individuals and enters the infected compartment. Each infected person runs through the course of the disease and eventually is removed from the number of those who are still infected, either by recovery or death, thus exiting the system at recovery rate γ ; the recovered people are considered permanently immune. More complex extensions of the SIR model have been developed over the years, and the COVID-19 pandemic has motivated researchers to further investigate the topic [21]–[24].

The parameters that rule the dynamics from one compartment to another, e.g., infection rate β and recovery rate γ in the SIR model, are usually time invariant, and several approaches have been proposed for tuning or estimating them [25], [26], [27]. We have developed in [7] a Bayesian approach that sequentially estimates the compartmental model's parameters by exploiting data made publicly available daily by national authorities, such as the number of new positive cases, the number of recovered people, and the number of fatalities. The approach is based on discretization of the continuous stochastic differential equations that describe the compartmental epidemiological model [28] and on basic principles of Bayesian sequential estimation that involve a prediction step and an update step. The estimated model's parameters are then used to forecast the evolution of the COVID-19 pandemic via ensemble forecasting, i.e., a Monte Carlo approach that produces a set (or ensemble) of forecasts. This forecasting approach also requires hypothesizing about the future—i.e., not observed yet—evolution of the infection rate (see details in [7]). The effectiveness of the proposed method has been evaluated in [7] through its application on data from the first pandemic wave in the Lombardia region of Italy and in United States. Here, we report in Figure 4 the estimated infection rate in Lombardia from February 24 to June 30, 2020; the decrease in the infection rate, which represents the slowdown of the pandemic, clearly reflects the restrictive measures established by the Italian government on March 8, 2020. Figure 5 instead shows the forecast of the pandemic evolution performed on April 13, 2020; we observe that accurate estimation of the time-varying infection and recovery rates facilitates reliable prediction of the evolution of the infection, with a forecasted number of infected individuals that closely follows future observations.

As mentioned above, the proposed forecasting approach requires hypothesizing about the future evolution of the infection rate. The infection rate models the interaction between people; therefore, its future evolution depends on how authorities react to the progress of the contagion and how people respond to the imposed restrictions. Thus, modeling its future evolution is still an open issue. In [8], we proposed a solution that employs MAST,

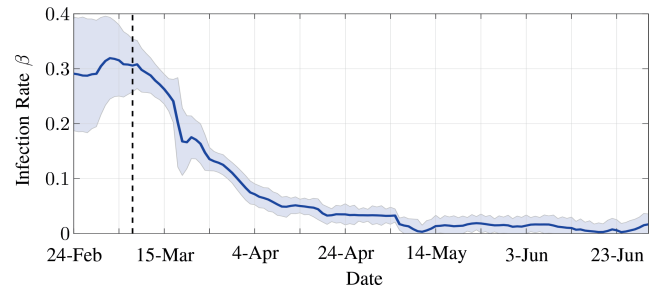


Figure 4

From [7], estimated infection rate β for the Lombardia region of Italy from February 24 to June 30, 2020. The vertical dashed line indicates March 8, 2020, the beginning of the lockdown imposed by the Italian government during the first pandemic wave. The shaded areas represent the 90% confidence interval.

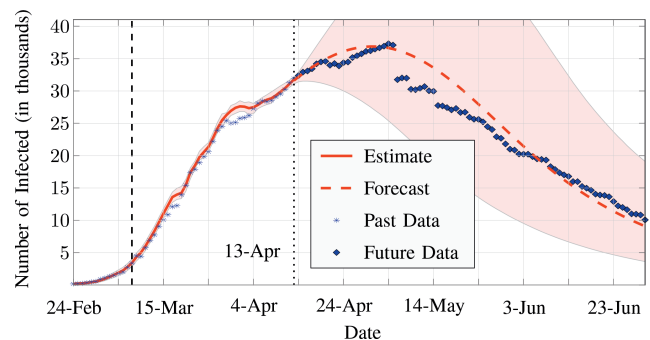


Figure 5

From [7], estimation and forecasting in solid and dashed lines, respectively, of the number of infected individuals in the Lombardia region of Italy. The date corresponding to the end of the estimation and the beginning of the forecast, that is, April 13, 2020, is marked by a vertical dotted line. The leftmost vertical dashed line marks March 8, 2020, the beginning of the lockdown imposed by the Italian government during the first pandemic wave. The shaded area represents the 90% confidence interval. The large step in the number of infected individuals observed on May 6, 2020, is due to inaccurate reporting of the data by local authorities.

hence providing a comprehensive, decision-directed estimation-detection-forecasting tool. Specifically, when an outbreak is declared through MAST, the hypothesized infection rate slope (i.e., the derivative of the infection rate) is positive (or zero), whereas when the termination of a pandemic wave is declared, the hypothesized infection rate slope is negative (or zero).

Detection and forecast of the second and third waves in United States are analyzed in [8]. Here, we report in Figure 6 the mean absolute percentage error (MAPE) of the forecast computed on different days from June 22, 2020 (day of detection of the second wave in the United States), and for two time horizons, i.e., 2 and 4 weeks. The results are compared with an alternative approach that employs a nonlinear least squares fitting algorithm that, using the number of infected and recovered individuals, computes the parameters of an epidemiologi-

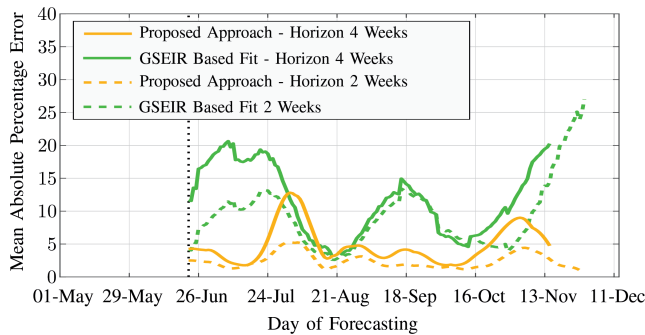


Figure 6 From [8], MAPE of the forecast of the pandemic evolution in the United States performed with the proposed algorithm and the GSEIR-based fit approach on different days (abscissa) and for different time horizons, i.e., 2 and 4 weeks (depicted with solid and dashed lines, respectively).

Table 1
Time-Averaged MAPE of the Forecast of the Epidemic Evolution

Algorithm	2 Weeks (%)	4 Weeks (%)
Proposed	2.39	4.82
SIR-based fit	101.83	137.17
SIR-X-based fit	25.86	27.53

cal model, known as generalized SEIR (GSEIR) [29], with four more compartments than SIR that account for insusceptible people, exposed (but not infectious) people (E), quarantined people, and deaths. We observe that the proposed forecast algorithm outperforms the GSEIR-based fit approach for both forecast horizons of 2 and 4 weeks, and that, apart from the time interval of roughly between July 19 and August 13, the proposed approach presents a MAPE that is always below 10%.

An additional comparison is reported in Table 1. The SIR- and SIR-X-based fits represent approaches similar to the one described for GSEIR that use, respectively, the classical SIR model and the SIR-X model [30]; the latter directly accounts for restrictive measures by removing susceptible individuals from the disease-spreading process. The results show significant performance improvements by the proposed forecast algorithm in terms of time-averaged MAPE.

CONCLUSION

Leveraging known concepts from related fields, we provided an overview of the recent advances on information processing methodologies to combat the COVID-19 pandemic. First, we described a quickest detection procedure, known as MAST, designed to detect an exponential growth of positive cases with a mean delay of few days and, at the same time, with a low risk of erroneously declaring an outbreak. In addition, MAST was shown to be suitable—with proper adjustments—for the detec-

tion of the termination of a pandemic wave. The effectiveness of MAST has been demonstrated through extensive analysis of COVID-19 data of second and third waves from different countries, as well as from smaller communities.

Second, we reported a Bayesian approach that estimates the features of the pandemic, e.g., the infection rate, and reliably forecasts the evolution of the contagion. This estimation-forecasting approach has been demonstrated on COVID-19 data of first and second waves, achieving low MAPEs for forecasts of up to 4 weeks and favorable comparison with alternative approaches.

REFERENCES

- Rossmann, H., Shilo, S., Meir, T., Gorfine, T., Shalit, U., and Segal, E. COVID-19 dynamics after a national immunization program in Israel. *Nature Medicine*, Vol. 27, 3 (Apr. 2021), 1055–1061.
- Fontanet, A., Autran, B., Lina, B., Kieny, M. P., Karim, S. S. A., and Sridhar, D. SARS-CoV-2 variants and ending the COVID-19 pandemic. *Lancet*, Vol. 397, 10278 (Mar. 2021), 952–954.
- Pretz, K. Using modeling to understand how COVID-19 preventive measures work: washing hands, wearing masks: does it help?—[spectral lines]. *IEEE Spectrum*, Vol. 57, 10 (Oct. 2020), 4.
- Braca, P., Gaglione, D., Marano, S., Millefiori, L. M., Willett, P., and Pattipati, K. Decision support for the quickest detection of critical COVID-19 phases. *Scientific Reports*, Vol. 11, 1 (Apr. 2021), 8558.
- Braca, P., Gaglione, D., Marano, S., Millefiori, L. M., Willett, P., and Pattipati, K. Quickest detection of COVID-19 pandemic onset. *IEEE Signal Processing Letters*, Vol. 28 (Mar. 2021), 683–687.
- Braca, P., Gaglione, D., Marano, S., Millefiori, L. M., Willett, P., and Pattipati, K. MAST: a quickest detection procedure for COVID-19 epidemiological data to trigger strategic decisions. In *Proceedings of EUSIPCO-21*, Dublin, Ireland, Aug. 2021.
- Gaglione, D., Braca, P., Millefiori, L. M., Soldi, G., Forti, N., Marano, S., Willett, P., and Pattipati, K. R. Adaptive Bayesian learning and forecasting of epidemic evolution—data analysis of the COVID-19 outbreak. *IEEE Access*, Vol. 8 (Sep. 2020), 175244–175264.
- Soldi, G., Forti, N., Gaglione, D., Braca, P., Millefiori, L. M., Marano, S., Willett, P., and Pattipati, K. Quickest detection and forecast of pandemic outbreaks: analysis of COVID-19 waves. *IEEE Communications Magazine*. [Online]. Available: <https://arxiv.org/abs/2101.04620> Jul. 2021.
- Millefiori, L. M., Braca, P., Zissis, D., Spiliopoulos, G., Marano, S., Willett, P. K., and Camiel, S. COVID-19 impact on global maritime mobility. *Scientific Reports* (under review). [Online]. Available: <https://arxiv.org/abs/2009.06960> 2021.
- Basseville, M., and Nikiforov, I. V. *Detection of Abrupt Changes: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, 1993.
- Poor, H. V., and Hadjiladis, O. *Quickest Detection*. Cambridge, United Kingdom: Cambridge Univ. Press, 2009.
- Page, E. Continuous inspection schemes. *Biometrika*, Vol. 41, 1/2 (Jun. 1954), 100–115.
- Braca, P., Gaglione, D., Marano, S., Millefiori, L. M., Willett, P., and Pattipati, K. MAST: COVID-19 pandemic onset test—multi-country analysis and visualization. Available: <https://covid-mast.github.io> 2020.
- Marano, S., and Sayed, A. H. Decision-making algorithms for learning and adaptation with application to COVID-19 data. [Online]. Available: <https://arxiv.org/abs/2012.07844> May 2021 (submitted).
- Winfield, N. Italy’s Lombardy again in virus crisis as Brescia sees surge. *AP News*. [Online]. Available: <https://apnews.com/article/europe-italy>

coronavirus-pandemic-coronavirus-vaccine-national-governments-e4d4a b04ed7f0de06f454cdf718dfd48 Feb. 2021.

16. BBC. Coronavirus: Italy imposes regional lockdown as Europe battles surges. [Online]. Available: <https://www.bbc.com/news/world-europe-54839429> Nov. 2020.
17. Streit, R. Introduction to the issue. *ISIF Perspectives on Information Fusion*, Vol. 3, 1 (May 2020), 3.
18. Daley, D., and Gani, J. *Epidemic Modelling: An Introduction*. Cambridge, United Kingdom: Cambridge Univ. Press, 2001.
19. Kermack, W. O., McKendrick, A. G., and Walker, G. T. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London*, Vol. 115, 772 (Aug. 1927), 700–721.
20. Bjørnstad, O. N., Shea, K., Krzywinski, M., and Altman, N. Modeling infectious epidemics. *Nature Methods*, Vol. 17, 5 (Apr. 2020), 455–456.
21. Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., and Colaneri, M. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, Vol. 26, 6 (Jun. 2020), 855–860.
22. Bertozzi, A. L., Franco, E., Mohler, G., Short, M. B., and Sledge, D. The challenges of modeling and forecasting the spread of COVID-19. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 117, 29 (Jul. 2020), 16732–16738.
23. Adiga, A., Dubhashi, D., Lewis, B., Marathe, M., Venkatramanan, S., and Vullikanti, A. Mathematical models for COVID-19 pandemic: a comparative analysis. *Journal of the Indian Institute of Science*, Vol. 100, 4 (Oct. 2020), 793–807.
24. Xiang, Y., Jia, Y., Chen, L., Guo, L., Shu, B., and Long, E. COVID-19 epidemic prediction and the impact of public health interventions: a review of COVID-19 epidemic models. *Infectious Disease Modelling*, Vol. 6 (2021), 324–342.
25. Skvortsov, A., and Ristic, B. Monitoring and prediction of an epidemic outbreak using syndromic observations. *Mathematical Biosciences*, Vol. 240, 1 (Nov. 2012), 12–19.
26. Sheinson, D. M., Niemi, J., and Meiring, W. Comparison of the performance of particle filter algorithms applied to tracking of a disease epidemic. *Mathematical Biosciences*, Vol. 255 (Sep. 2014), 21–32.
27. Shahtori, N. M., Scoglio, C., Pourhabib, A., and Sahneh, F. D. Sequential Monte Carlo filtering estimation of Ebola progression in West Africa. *2016 American Control Conference (ACC)* Jul. 2016, 1277–1282.
28. Allen, L. J. A primer on stochastic epidemic models: formulation, numerical simulation, and analysis. *Infectious Disease Modelling*, Vol. 2, 2 (May 2017), 128–142.
29. Peng, L., Yang, W., Zhang, D., Zhuge, C., and Hong, L. Epidemic analysis of COVID-19 in China by dynamical modeling. *arXiv*. [Online]. Available: <https://arxiv.org/abs/2002.06563> Feb. 2020.
30. Maier, B. F., and Brockmann, D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*, Vol. 368, 6492 (May 2020), 742–746.

Domenico Gaglione is currently a Research Scientist with NATO STO Centre for Maritime Research & Experimentation (CMRE), La Spezia, Italy. His research interests include statistical signal processing with emphasis on state estimation, data fusion, and multisensor multitarget tracking, and radar (sonar) signal processing.



Paolo Braca is a Senior Scientist with the Research Department and a Project Manager of the Data Knowledge and Operational Effectiveness (DKOE) Program. He conducts research in the general area of statistical signal processing with emphasis on detection and estimation theory, wireless sensor network, multi-agent algorithms, target tracking and data fusion, adaptation and learning over graphs, and radar (sonar) signal processing.



Giovanni Soldi is a Scientist with the Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy. His current research interests include statistical signal processing with a focus on state estimation, data fusion, and multisensor-multitarget tracking techniques.



Nicola Forti is currently a Scientist with the Research Department, NATO Science & Technology Organization Centre for Maritime Research and Experimentation (CMRE). His current research interests include state estimation and control theory, statistical signal processing, machine learning, target tracking, and data fusion.



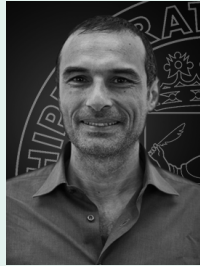
Leonardo M. Millefiori is a Research Scientist with the Research Department with the NATO Science and Technology Organization Center for Maritime Research and Experimentation (CMRE), La Spezia, Italy. His research interests include target motion modeling, statistical learning and signal processing, and target tracking and data fusion



Peter K. Willett is a faculty member in the Electrical and Computer Engineering Department at the University of Connecticut. His primary areas of research have been statistical signal processing, detection, machine learning, communications, data fusion, and tracking.



Stefano Marano is a Professor with DIEM, University of Salerno, Italy. His current research interests include statistical signal processing with emphasis on distributed inference, sensor networks, and information theory.



Krishna R. Pattipati is with the Department of Electrical and Computer Engineering, UConn, where he is currently the Board of Trustees Distinguished Professor and the UTC Chair Professor in systems engineering. His research interests include proactive decision support, uncertainty quantification, smart manufacturing, autonomy, knowledge representation, and optimization-based learning and inference.

