**Commentary**

# COVID-19 big data highlight new environmental research frontiers and current spotlights

Jie Han*, Han Dai, Yi Bai

Institute of Global Environmental Change, School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, Xi'an 710049, P.R. China

*Corresponding author. Email: jiehan@xjtu.edu.cn; ORCiD: https://orcid.org/0000-0001-6790-0650

In the past two years, environmental scientists around the world have put tremendous efforts into understanding the environmental persistence and transmission of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the infectious agent of COVID-19, and the reverse impact on the natural and living environments by the global pandemic. Many innovative, forward-thinking thoughts and methods have been put forward by environmental scientists, such as wastewater-based epidemiology surveillance (Farkas et al. 2020, Wilder et al. 2021), sewage archiving (Dolfing 2020), biobank-enabled tracing of early cases of infection (He and Han 2021), and aerosol-mediated airborne transmission as an under-investigated route of viral transmission in the current pandemic (Morawska and Cao 2020). Meanwhile, new environmental issues either caused or amplified by COVID-19 have been raised, such as the widespread use of single-use plastics (Hale and Song 2020), inappropriate disposal of personal protective equipment (e.g., face masks) (De-la-Torre et al. 2021; Di Maria et al. 2020), community-wide use of chemical disinfectants and potential transfer of microbial resistance to humans (Chen et al. 2021), and the overuse of antiviral and antibiotic drugs through the current pandemic (Kumar et al. 2021).

There are now more than 10,000 articles related to COVID-19 published in scholarly journals in environmental research disciplines, representing about 5% of scientific publications on COVID-19 from all research disciplines and showcasing the tireless efforts and swift actions taken by environmental researchers around the globe. The vast and rapidly growing body of scientific publications that appeared in such a short timeframe also makes it a daunting task for one to comb through the existing findings and opinions to identify major consensuses and knowledge gaps, for decision-making or prioritizing research efforts. Under such circumstances, new data-driven tools provide an efficient means of harnessing the wealth of information and perhaps more importantly, identifying critical knowledge gaps for further research (**Fig. 1**). In this domain, the recent work by Zhu et al. (2021) presents a novel perspective on analyzing the recent achievements, major trends, and research opportunities across the environmental research landscape represented by two decades

of research published in a leading journal in the environmental science discipline. By empowering researchers with automated tools on data mining and trend analysis, it breaks the barrier of traditional literature searching and analysis, a lengthy process predominantly relying on human intelligence in the past.

There are currently seven scientometric studies—most of which are of a preliminary nature—related to environmental research and COVID-19 after performing keyword searching in Clarivate® Web of Science and manually screening the search results. Many of these studies, however, focused on analyzing authors and affiliations, as well as the countries of origin, source titles, and the citations of journal articles published in this interdisciplinary and rapidly evolving research domain. Some apparent trends were readily drawn with respect to the bibliometric aspects of those publications (Usman and Ho, 2020, Zyoud and Zyoud, 2021). Other studies concluded that the existing research on COVID-19 in the environmental research disciplines had largely focused on environmental variables affecting the virus spread and the reverse impact of human behaviors on the environment such as lockdowns in the current pandemic (Aranda et al., 2021, Usman and Ho, 2020, Wang et al., 2021). Some authors also called for more collaborated research on emerging environmental issues or existing issues aggravated by the COVID-19 pandemic, as well as responses and policies by governments and public authorities (Chernysh and Roubik, 2020; Wang et al., 2021).

While these studies presented meaningful snapshots of the recent research efforts put by environmental researchers around the globe, there are apparent limitations in their scopes and methods from a scientometric study's perspective. First, there is a general lack of in-depth analysis on the specific topics, findings, opinions, and other contents presented in those publications that are of scientific merit. Apart from analyses on bibliometric data, existing scientometric studies generally do not conduct in-depth assessments on the topic relevance, data contents, body texts, and the inter-relations of findings from those publications. **Figs. 2–5** show excerpts of the best efforts in the latter.

Without performing more detailed analyses on these, however, it would be difficult, if at all possible, to gain insights into the current consensus, major gaps, and evolving research trends from the vast body of information presented in those scholarly publications. Secondly, there is also significant room for improvement in the methods adopted in these studies. Generally, existing scientometric studies have used two strategies, namely, user-defined keywords and article categories provided by Clarivate® Web of Science to define their scope of searches and subsequent data extraction. Some user-defined keywords are inadequate and potentially biased (e.g., 'environment protection and covid'), and do not provide adequate coverage on scholarly publications on COVID-19 in different environmental research disciplines. Adding to these limitations, all scientometric studies published to date in this domain have only analyzed publications up to December 2020, nine months after the World Health Organization declared COVID-19 a pandemic (WHO 2020) and prior to the onset of the massive second- and third-waves of infections around the globe since the beginning of 2021. Since then, the number of scholarly publications in this domain has more than quadrupled. To be specific, the number of publications on COVID-19 in the environmental research disciplines between January–June 2021 had surpassed the total number of publications in this research domain throughout 2020, which again doubled in the second half of 2021. The sheer volume of new publications that appeared after these scientometric studies necessitates a more comprehensive and in-depth look at the current findings and scholarly opinions presented in scientific literature.

Nowadays, it is feasible to use freely available programming languages such as Python and *R* to extract, analyze, and visualize information relevant to any specific, user-defined topics of interest, in a highly efficient manner. Along with data searches and crawling, the screening, learning, treatment, and output of data can all be made into programmable tasks at users' requests. Algorithms—once they are built and made available—can be readily adopted by other researchers and public health authorities and made into flexible and robust tools to perform routine data mining and analyses, for instance, to aid in the discovery of significant findings and knowledge gaps or to capture early

research trends as they emerge in the scientific literature. We, therefore, call for the attention of researchers, organizations, and public health authorities to adopt and develop these new and potentially transformative tools to make the existing and prospective work by global environmental scientists more visible and understandable to policy-makers and the general public and ultimately, to serve for the conceptualization, development, and implementation of effective and sustainable solutions in efforts dealing with pandemics and other grand challenges in the future uncertain world.

**Conflict of interest.** The authors declare that they have no conflict of interest in this work.

**Reference**

Al-Gheethi, A., Al-Sahari, M., Malek, M. A., Noman, E., AI-Maqtari, Q., Mohamed, R., Talip, B. A., Alkhadher, S., Hossain, M. S., 2020. Disinfection Methods and Survival of SARS-CoV-2 in the Environment and Contaminated Materials: A Bibliometric Analysis. Sustainability, 12(18): 7378. doi: 10.3390/su12187378.

Casado-Aranda, L. A., Sánchez-Fernández, J., Viedma-Del-jesús, M. I., 2021. Analysis of the scientific production of the effect of COVID-19 on the environment: a bibliometric study. Environ. Res. 193: 110416. doi:10.1016/j.envres.2020.110416.

Chen, B., Han, J., Dai, H., Jia, P. Q., 2021. Biocide-tolerance and antibiotic-resistance in community environments and risk of direct transfers to humans: Unintended consequences of community-wide surface disinfecting during COVID-19? Environ. Pollut. 283: 117074. doi:10.1016/j.envpol.2021.117074.

Chernysh, Y., Roubik, H., 2020. International Collaboration in the Field of Environmental Protection: Trend Analysis and COVID-19 Implications. Sustainability.12(24): 10384. doi: 10.3390/su122410384.

De-La-torre, G. E., Rakib, M. R. J., Pizarro-Ortega, C. I., Dioses-Salinas, D. C., 2021. Occurrence of personal protective equipment (PPE) associated with the COVID-19 pandemic along the Coast of Lima, Peru. Sci. Total. Environ. 774: 145774. doi:10.1016/j.scitotenv.2021.145774.

Di Maria, F., Beccaloni, E., Bonadonna, L., Cini, C., Confalonieri, E., La Rosa, G., Milana, M. R., Testai, E., Scaini, F., 2020. Minimization of spreading of SARS-CoV-2 via household waste produced by subjects affected by COVID-19 or in quarantine. Sci. Total. Environ. 743: 140803. doi:10.1016/j.scitotenv.2020.140803.

Dolfing, J., 2020. The importance of sewage archiving in coronavirus epidemiology and beyond. Environ. Sci. Technol. 54 (13): 7740-7741. doi:10.1021/acs.est.0c02972.

Farkas, K., Walker, D. I., Adriaenssens, E. M., McDonald, J. E., Hillary, L. S., Malham, S. K., Jones, D. L., 2020. Viral indicators for tracking domestic wastewater contamination in the aquatic environment. Water Res. 181: 115926. doi:10.1016/j.watres.2020.115926.

Hale, R. C., Song, B., 2020. Single-use plastics and COVID-19: scientific evidence and environmental regulations. Environ. Sci. Technol. 54 (12): 7034-7036. doi:10.1021/acs.est.0c02269.

He, S.S., Han, J., 2021. Biorepositories (biobanks) of human body fluids and materials as archives for tracing early infections of COVID-19. Environ. Pollut. 274: 116525. doi:10.1016/j.envpol.2021.116525.

Ji, B., Zhao, Y. Q., Wei, T., Kang, P. Y., 2021. Water science under the global epidemic of COVID-19: Bibliometric tracking on COVID-19 publication and further research needs. J. Environ. Chem. Eng. 9(4). doi:10.1016/j.jece.2021.105357.

Kumar, M., Mazumder, P., Mohapatra, S., Kumar Thakur, A., Dhangar, K., Taki, K., Mukherjee, S., Kumar Patel, A., Bhattacharya, P., Mohapatra, P., et al. 2021. A chronicle of SARS-CoV-2: Seasonality, environmental fate, transport, inactivation, and antiviral drug resistance. J. Hazard. Mater. 405: 124043. doi:10.1016/j.jhazmat.2020.124043.

Morawska, L., Cao, J. J., 2020. Airborne transmission of SARS-CoV-2: The world should face the reality. Environ. Int. 139: 105730. doi:10.1016/j.envint.2020.105730.

Usman, M., Ho, Y. S., 2021. COVID-19 and the emerging research trends in environmental studies: a bibliometric evaluation. Environ. Sci. Pollut. Res. 28 (14): 16913-16924. doi:10.1007/s11356-021-13098-z.

Wang, Q., Zhang, C., 2021. Can COVID-19 and environmental research in developing countries support these countries to meet the environmental challenges induced by the pandemic? Environ. Sci. Pollut. Res.: 1-21. doi:10.1007/s11356-021-13591-5.

WHO 2020. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. World Health Organization. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

Wilder, M. L., Middleton, F., Larsen, D. A., Du, Q., Fenty, A., Zeng, T., Insaf, T., Kilaru, P., Collins, M., Kmush, B., et al. 2021. Co-quantification of crAssphage increases confidence in wastewater-based epidemiology for SARS-CoV-2 in low prevalence areas. Water Res. X 11: 100100. doi:10.1016/j.wroa.2021.100100.

Zhu, J. J., Dressel, W., Pacion, K., Ren, Z. J., 2021. ES&T in the 21st century: a data-driven analysis of research topics, interconnections, and trends in the past 20 years. Environ. Sci. Technol. 55 (6): 3453-3464. doi:10.1021/acs.est.0c07551.

Zyoud, S. H., Zyoud, A. H., 2021. Coronavirus disease-19 in environmental fields: a bibliometric and visualization mapping analysis. Environ. Dev. Sustain. 23 (6): 8895-8923. doi:10.1007/s10668-020-01004-5.

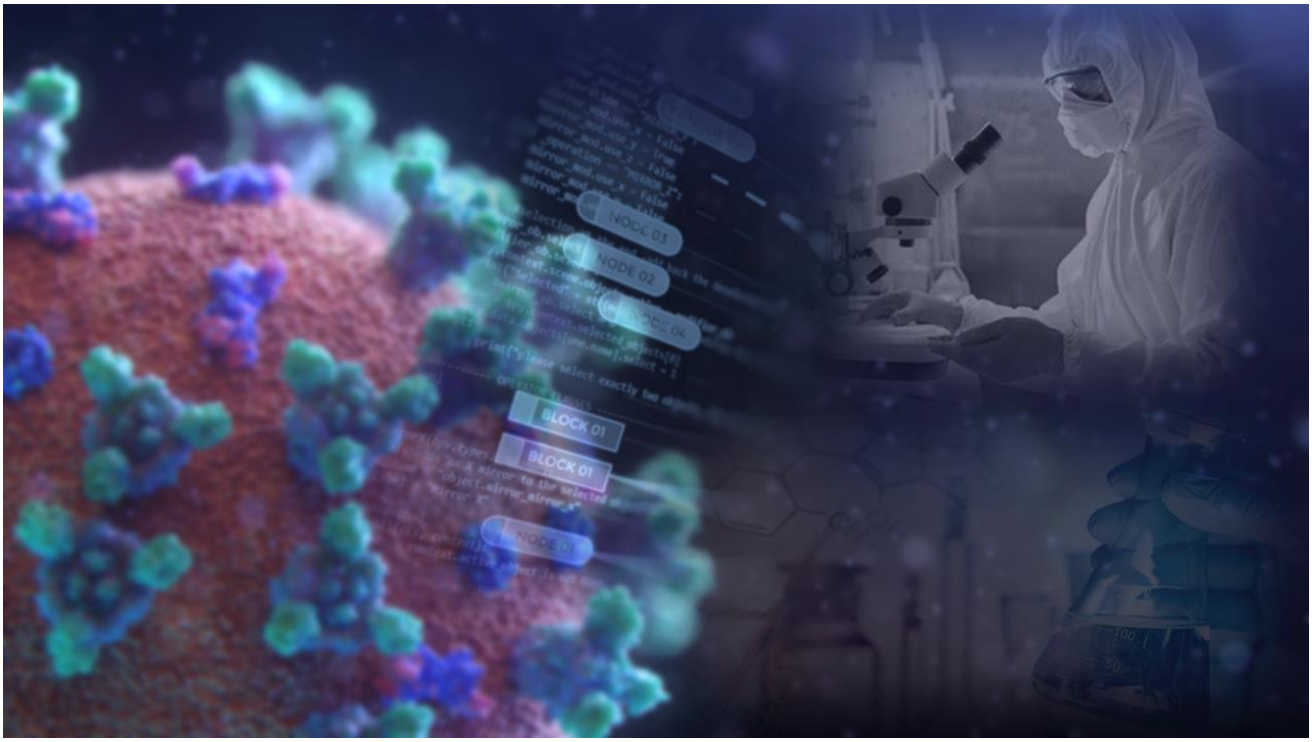


**Fig. 1** Data-driven tools provide an efficient approach for researchers and policy-makers to harness the wealth of information on COVID-19 related research, which enables researchers and policy-makers to quickly identify systematic and critical knowledge gaps for prioritizing future research efforts. Coronavirus image overlay adopted from brandeis.edu.

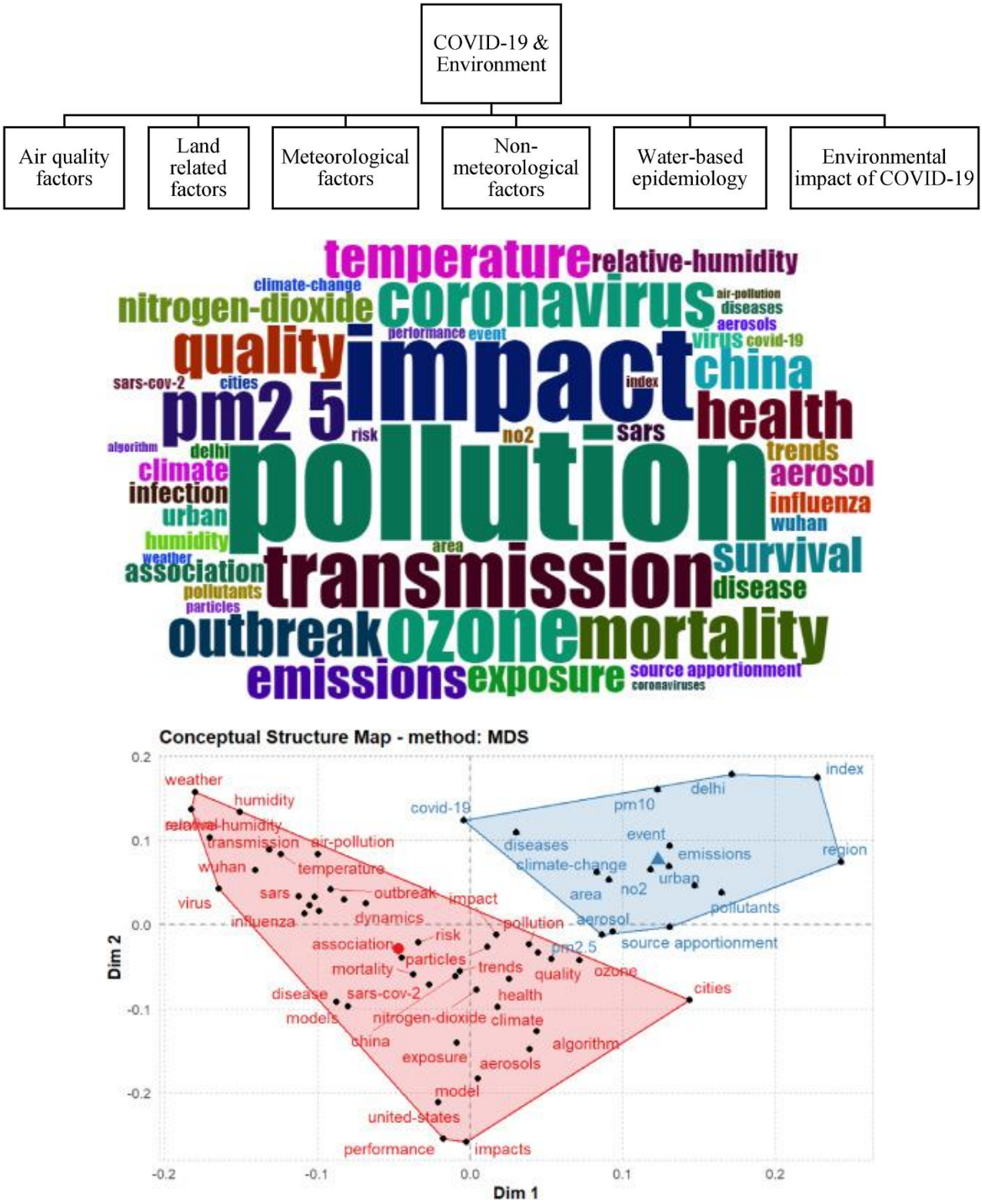


**Fig. 2** Research themes (top), word cloud of prominent keywords (middle), and multi-dimensional scaling of keyword clusters (bottom) in 328 publications relevant to COVID-19 and the environment. Publications were searched in the Web of Science database and MedRxiv preprint platform. Adapted from Sharma et al. (2021).

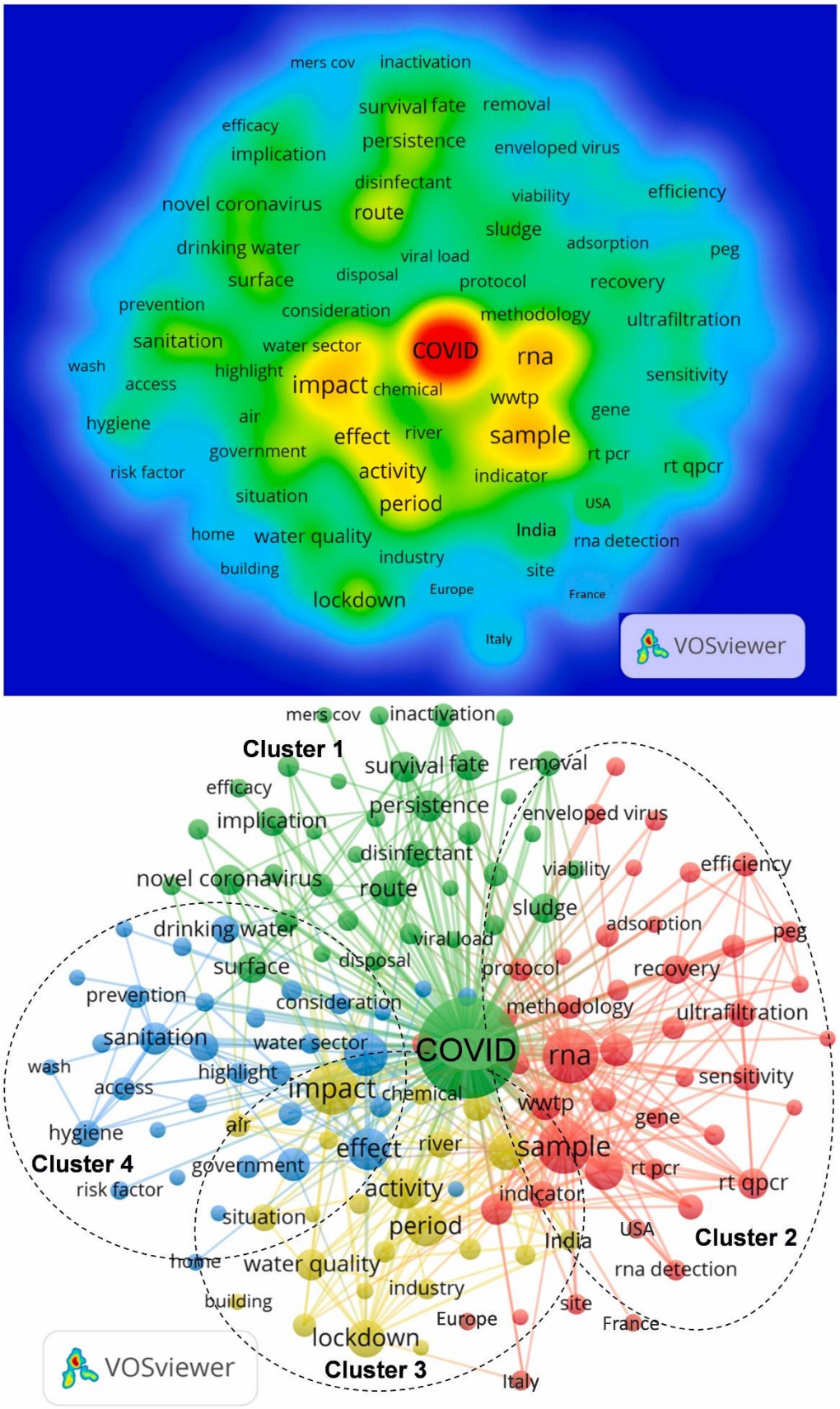


**Fig. 3** Density visualization of co-occurrence keywords and key terms occurrences network in 262 COVID-19 publications related to water science and technology. Adapted from Ji et al. (2021).

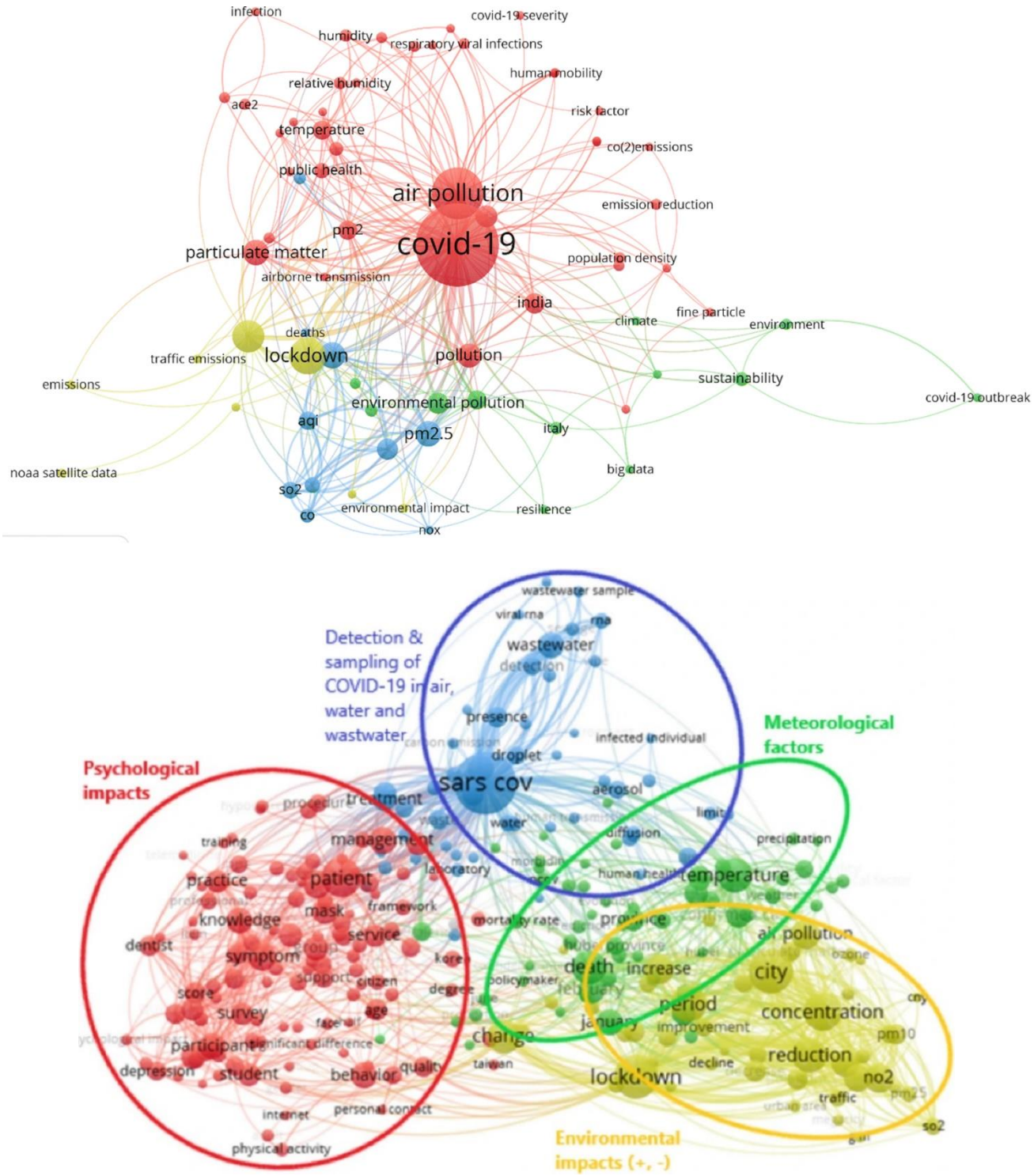


**Fig. 4** (Top) Network diagram of keywords in 255 publications related to COVID-19 and environmental pollution indexed in Web of Science Core Collection between Dec 2019 and Oct 2020. Adapted from Wang and Zhang (2021). (Bottom) Network visualization map of co-occurrence of terms in the abstracts of 729 publications on COVID-19 from environmental science journals in the Scopus database. Adapted from Zyoud and Zyoud (2021).

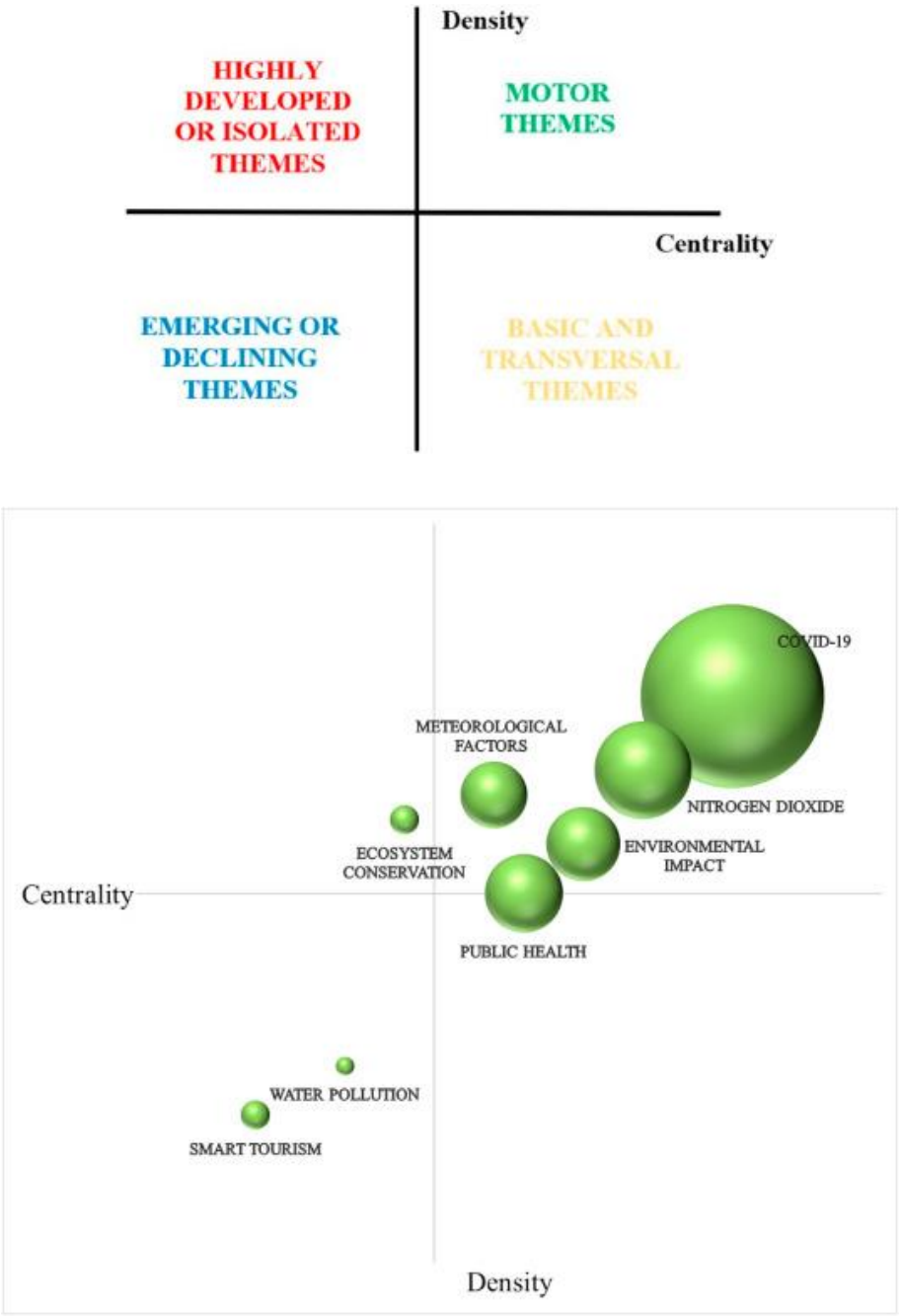


**Fig. 5** (Top) Quadrants of a strategic diagram; (bottom) Strategic diagram depicting the number of documents combining the themes of COVID-19 and the environment. In an attempt to discern active research topics and emerging fields related to environmental issues in COVID-19 research, the authors used SciMAT, an open-source science mapping analysis tool, and visualized the co-occurrence of keywords into a simple strategic diagram. A total of 440 publications between 1 December 2019 and 6 September 2020 were searched from Web of Science Core Collection and the Scopus database. Adapted from Casado-Aranda et al. (2021).