Unreflective use of old data sources produced echo chambers in the water-electricity nexus

Vaca-Jimenez, S.; Gerbens-Leenes, P. W.; Nonhebel, S.; Hubacek, K.

Published in:
Nature sustainability

DOI:
10.1038/s41893-021-00686-7

nature sustainability

ANALYSIS

https://doi.org/10.1038/s41893-021-00686-7

Check for updates

Unreflective use of old data sources produced echo chambers in the water–electricity nexus

S. Vaca-Jiménez [1,2] ✉, P. W. Gerbens-Leenes [1], S. Nonhebel [1] and K. Hubacek [1]

Echo chambers in science describe the amplification and repetition of information within closed networks. Frequently used data sources can cause echo chambers as scientists keep reading similar outputs from different sources, creating false perceptions of certainty and variety of data sources. We show this effect by studying the scientific and grey literature on water use by electricity systems. The power sector is the largest contributor to anthropogenic carbon emissions and the second largest water consumer. We have assessed the scope and references of 2,426 papers and created a citation network to trace original data sources. Most data sources used for the last 30 years originate from a few old US publications, recently also Chinese, that echo through publications. This echo effect, also reflected in recent scientific publications, creates a confirmation bias, also facilitating double counting of the water intensities of electricity generation. This example from sustainability science warns of the risk of echo chambers in other scientific disciplines.

Water and energy consumption are highly intertwined. For example, electricity is often required to provide freshwater, treat municipal water or clean wastewater. Water is an essential input in hydropower plants, the irrigation of biofuels and thermal power plants, and their fuels require (fresh) water for operation and mining. The power sector is the largest contributor to climate change[1]. To reduce greenhouse gas emissions, it aims to replace fossil energy by renewables[1,2]. Globally, the sector is, after agriculture, the largest freshwater user[3]. However, some mitigation options are water-intensive, for example, bioenergy[4,5] or carbon capture and storage systems[6,7], and might not be feasible due to freshwater constraints[8]. Information on water intensities, that is, the volumes of freshwater per unit of electricity produced, is paramount as demand increases and water becomes scarcer in the future.

Scientists require access to relevant data to provide suggestions for the reduction of water intensities. However, data sources repeatedly used, sometimes for decades, or inappropriately used in different contexts, can cause echo chambers and have adverse effects on scientific outcomes. Echo chambers describe information amplification and repetition within closed networks[9]. Echo refers to repetition, chamber to the medium in which the repetition occurs[10]. Echo chambers lead to biases, intensification of viewpoints and analyses based on incompatible data. Today, echo chambers are frequently identified in social media[11] and political affiliation[12], but also appear elsewhere. The Supplementary Information has more information regarding echo chambers.

In scientific communities, data sources repeatedly used through the years form a suitable environment for echo chambers to appear as scholars keep reading similar outputs from different sources, creating false perceptions of certainty, accuracy and confirmation. For example, the sustainability science community relies on relevant data to analyse the water requirements of electricity generation. The appearance of echo chambers has substantial effects on the contribution this community provides to sustainability research.

Water intensities are available from grey literature, studies and review papers. Grey literature contains both first- and second-hand data. First-hand data from energy companies provide the water used for their activities. Second-hand data usually come from government agencies, which have oversight over energy companies and log, or estimate, operational data. These sources rely on third-party reporting that is rarely published in scientific journals, and the information is frequently not peer-reviewed, or at least they do not disclose their peer-review processes, for example, the US Department of Energy (DOE)[13]. Nonetheless, there are cases in which there might not be better data available.

Moreover, data sources have spatial and temporal limitations. Old data sources are, in theory, less valid today as power plants, and their fuels, have improved technology, increasing outputs and operating time, decreasing greenhouse gas emissions and water intensities[14,15]. Likewise, electricity-generating technologies experience important spatial water intensity variation[16,17]. Existing case studies with specific system boundaries do not provide relevant information for present and future case studies that use other system boundaries or focus on other countries.

In this paper we identify echo chamber effects in science by studying the water–electricity nexus (WEN) literature, focusing on the water intensities of electricity generation. The WEN is a relevant case study as it involves a topic relevant to the Sustainable Development Goals (SDGs)[18]. According to Liu et al.[18], nexus approaches may provide a framework for the integrated planning and management required to achieve certain SDGs. In this case, the SDGs are related to energy and water. We have assessed the scope and references of 2,426 articles (including 854 peer-reviewed papers), classifying them by topic to identify articles assessing water intensities for different electricity-generating technologies and their fuels. Next, we used bibliometric methods and network analysis to study their underlying data sources through citation networks. Finally, we assessed the propagation of the original data sources. This example from sustainability science serves as an instructive case and a warning to pay attention to echo chambers in other scientific disciplines.

WEN most influential data sources

Table 1 shows the top 20 metric positions with the most important papers echoing throughout the WEN, indicating the papers with the

[1]Integrated Research on Energy, Environment and Society (IREES), University of Groningen, Groningen, The Netherlands. [2]Departamento de Ingeniería Mecánica, Escuela Politécnica Nacional, Quito, Ecuador. ✉e-mail: santiago.vaca@epn.edu.ec

Table 1 | Top 20 positions in the water–electricity nexus showing the papers with most citations (in-degrees) and the papers that connect papers of different clusters (betweenness)

No.	In-degrees	Betweenness
	Number of citations to this paper from other papers in the sample	Papers that bridge papers together
1	Macknick et al.[19] [S1]	Sanders[26] [S2]
2	Macknick et al.[20] [S3]	Macknick et al.[19] [S1]
3	Gleick[21] [S4]	DeNooyer et al. (2016)[S5]
4	Meldrum et al.[22] [S6]	Macknick et al.[20] [S3]
5	Byers et al. (2014)[S7]; Kenny et al. (2009)[S8]	Zhang and Anadon (2013)[S9]
6	IEA[23] [S10]	Cooper and Sehlke[27] [S11]
7	Zhang and Anadon (2013)[S9]	Byers et al. (2014)[S7]
8	Fthenakis and Kim[37] [S12]	Gheewala et al. (2011)[S13]
9	Sovacool and Sovacool (2009)[S14]	Dodder (2014)[S15]
10	Scott et al. (2011)[S16]	Gjorgiev and Sansavini (2018)[S17]
11	Siddiqi and Anadon (2011)[S18]; DOE[13] [S19]	Zhu et al. (2015)[S20]
12	DeNooyer et al. (2016)[S5]	Scanlon et al. (2013)[S21]
13	Averyt et al.[24] [S22]	Fernández-Blanco et al. (2017)[S23]
14	van Vliet et al. (2016)[S24]	Feng et al. (2014)[S25]
15	Feeley et al. (2008)[S26]; van Vliet et al. (2012)[S27]	Qin et al. (2015)[S28]
16	Feng et al. (2014)[S25]; Gerbens-Leenes et al.[4] [S29]	Wu and Chen (2017)[S30]
17	Mielke et al. (2010)[S31]; Bazilian et al. (2011)[S32]	Srinivasan et al.[42] [S33]
18	Stillwell et al. (2010)[S34]; Hoekstra et al. (2011)[S35]	Fingerman et al. (2011)[S36]
19	Spang et al.[33] [S37]; Averyt et al. (2013)[S38]; Rio Carrillo and Frei[36] [S39]	Larsen and Drews[28] [S40]
20	Qin et al. (2015)[S28]; Kyle et al.[43] [S41]; Hussey and Pittock (2012)[S42]; Davies et al.[38] [S43]	Zhou et al. (2016)[S44]

The superscripted number next to the authors' names indicates the reference number as it is listed in the Supplementary Note 1. Some positions include more than one paper because they have the same metric. The centrality metrics address the connectivity in the network and not the content, so the relevance is not defined by the work, but by the relation of that paper to others in the network.

most citations (in-degrees) and papers connecting papers of different clusters, or disciplines (betweenness). These terms are described in the Methods section. Some positions include more than one paper because they have the same metric. The reference list of all the articles in Table 1 is provided in Supplementary Note 1.

The top 20 WEN positions result from a mathematical analysis of the WEN network, considering papers as nodes and citations as edges. The metrics do not evaluate a paper's content, but relations to other papers. Papers with the most citations have a central position in the network, showing the importance of data sources in the WEN. The top four papers in the in-degrees metric are Macknick et al.[19], Macknick et al.[20], Gleick[21] and Meldrum et al.[22]. All, but Macknick et al.[20], are review papers providing a compilation of water intensities from different sources. Grey literature is also present in the lists. The most cited reports are from the International Energy Agency (IEA)[23], DOE[13], and Averyt et al.[24], compiling different data sources, providing water intensities for electricity systems.

Publications with relatively few citations, but critical for network connectivity, rank highly in the betweenness metric. They connect publication clusters on different topics that would otherwise be disconnected from the WEN. This is one of the main ideas of the nexus approach[25]. Review papers providing data are also fundamental and connect different topics, for example, Sanders[26], Macknick et al.[19], Cooper and Sehlke[27], and Larsen and Drews[28]. The information in Table 1 reveals data-source importance. WEN publications providing water intensities are not only the most cited, but also essential to connect knowledge in the network.

Old data sources echoing throughout the WEN

Figure 1 shows the WEN citation network with the references used to derive the water intensities of electricity generation technologies and their fuels in peer-reviewed papers. It shows the relationships between data sources and WEN papers. The data source networks per technology are given in Supplementary Note 2.

Grey literature water intensities from the last century moved along the papers through time. In the 1970s and 1980s, water intensities were reported in grey literature (Fig. 1, green squares on the left), for example, Gold et al.[29] (S107). Then, review papers, for example, Gleick[21] (S4), compiled the US water intensities of these grey literature publications, providing water intensity ranges and median values per technology. Later, case studies used the data from these review papers to assess the water intensities of electricity systems in other countries. Recently, a new wave of review papers, for example, Jin et al.[30] (S163), compiled case studies. But all of them still include the original sources in various ways.

Three papers dominate: Gleick[21] (S4), Macknick et al.[19] (S1) and Meldrum et al.[22] (S6). Moreover, although a large number of peer-reviewed publications were included here (854), the original data sources can be traced back to a few US-based grey literature sources from the last century, described above, echoing throughout the WEN publications of today.

Sometimes, the same water intensity has been repeatedly copied and treated as an independent observation, for example, the water intensities of coal-fired thermal power plants (TPPs) were compiled by Gleick[31] in 1993 (S66), passed on to Gleick[21] in 1994 (S4), copied by Inhaber[32] in 2004 (S55), next by Meldrum et al.[22] in 2013 (S6), then by Spang et al.[33] in 2014 (S37), Jornada and Leon[34] in 2016 (S141) and Jin et al.[30] in 2019 (S163).

WEN review papers have changed through time in terms of content and the pool of papers that they include. Usually, these review papers are connected as time passes, probably following a path dependency[35]. Based on the chronology and succession of

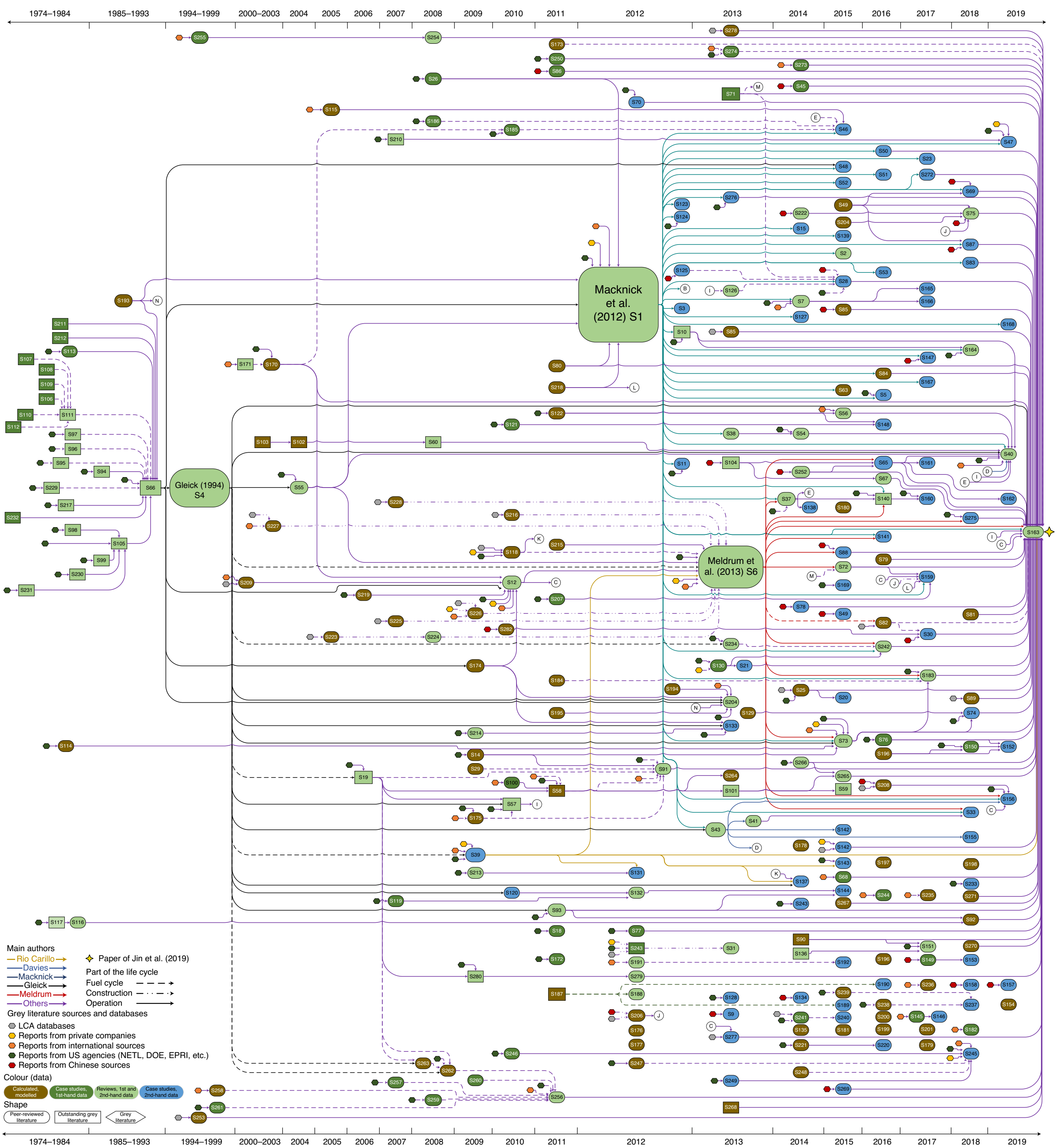

Fig. 1 | Network of literature data sources. WEN citation network with the references used to derive the water intensities of electricity generation technologies and their fuels. The bubble size reflects the number of times data sources were used in other studies. The paper position reflects the publication date. The full citation list of the references in this figure is given in Supplementary Note 1. LCA, life-cycle assessment; NETL, US National Energy Technology Laboratory.

publications, we have identified four generations of review papers that compile data sources for fuels or electricity production. In most studies, researchers compile data sources to inform their own research intention. For example, Larsen and Drews[28] performed an extensive review to estimate the water intensities of the European electricity system.

First-generation papers, for example, Gleick[21] (S4), Rio Carillo and Frei[36] (S39), and Fthenakis and Kim[37] (S12), used water intensities from the grey literature. Second-generation data compilation papers, for example, Macknick et al.[19] (S1), Davies et al.[38] (S43) and Meldrum et al.[22] (S6), used first-generation reviews and incorporated new reports of first-hand data sources from additional

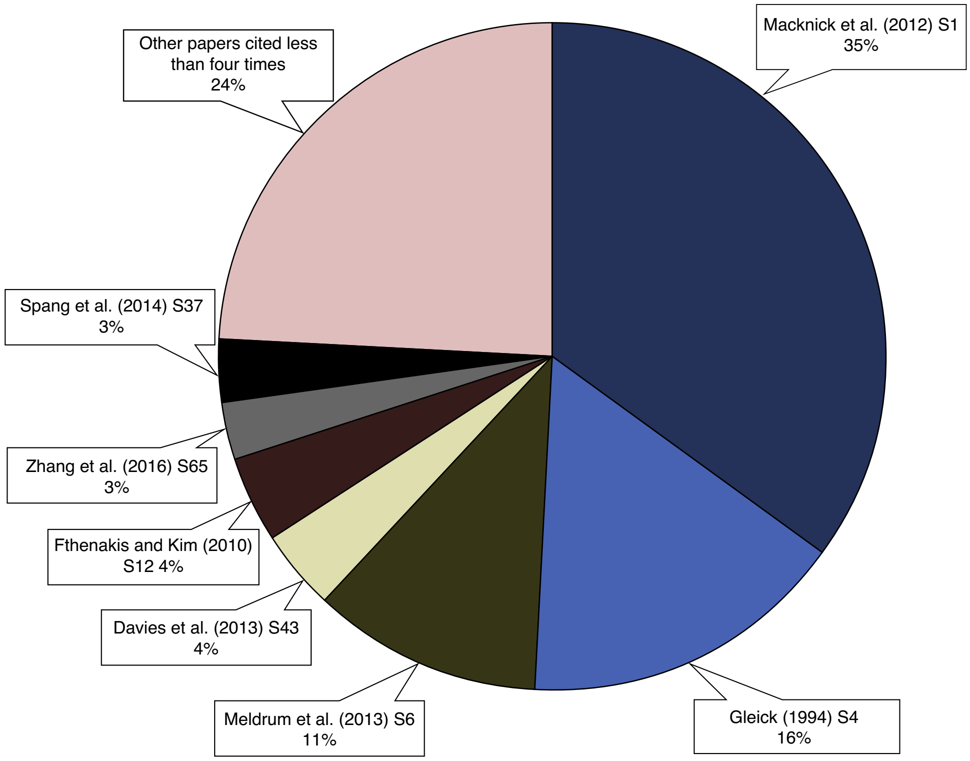

Fig. 2 | Peer-reviewed papers serving as data sources for case studies in the WEN. Chart showing the seven principal data sources to the WEN literature. The full list of references cited in this figure is given in Supplementary Note 1.

grey literature sources. Third-generation reviews, for example, Spang et al.[33] (S37) and Ali and Kumar[39] (S73), compiled data sources from first- and second-generation data compilations. New information from first-hand data sources and other grey literature is rarely added here. Recently, a fourth generation of review papers has appeared, for example, Larsen and Drews[28] (S40) and Jin et al.[30] (S163), who compiled data sources from first-, second- and third-generation reviews, excluding new data.

Echo chambers have been present since the third generation. However, gradually, different studies and review papers have grown in number. Therefore, new review papers summarizing water intensities are mostly based on published journal papers, rarely providing any new first-hand data or relevant new grey literature. It is, therefore, likely that they include repeated information.

Few original water intensities for hundreds of WEN papers

Figure 2 shows the peer-reviewed papers used as data sources for WEN case studies.

Only seven papers provided water intensity data to 76% of the papers in the sample: Macknick et al.[19], Gleick[21], Meldrum et al.[22], Davies et al.[38], Fthenakis and Kim[37], and Zhang et al.[40]. The first three provided water intensities to more than 62% of the WEN papers. The five most influential papers, their data sources and relationships are reported in Supplementary Note 3. Five papers, excluding Meldrum et al.[22], only used US data. The Americanization of data sources relates to the long US history of assessing water and electricity systems together, for example, Gold et al.[29]. In particular, although now fairly dated, US first- and second-hand data sources and grey literature contributed to the current knowledge regarding water use by electricity systems globally.

Macknick et al.[19] is the most used data source paper in the WEN literature. Its influence is so important that papers use its water intensities even for technologies that were not reported in it, as for instance in the case of the oil-fired TPPs reported by Davies et al.[38] and Spang et al.[33]. Instead of using oil-related water intensities, they assumed that oil-fired power plants have similar water intensities to gas-fired power plants, and used the natural gas water intensities provided by Macknick et al.[19].

The fact that only six papers echo through the existing literature, and water intensities of electricity generation are copied over and over again, also implies that relevant recent publications receive little attention. For instance, Jiang and Ramaswami[41] have provided first-hand data of the water intensities of 19 Chinese coal power plants, showing, for the first time, seasonal variation of the water intensities of Chinese power plants. Despite its importance, so far, this paper has only nine citations, with only one citation using its reported water intensities, the rest cited for background information.

Existing knowledge dominated by US data

Figure 3 shows the countries of origin of data sources in relation to the countries for which assessments were made.

US sources (64% of existing data sources) and US studies (38% of total studies) dominate the WEN. Almost all US studies used US data sources, except studies that used data from Meldrum et al.[22]. Only a quarter of studies used data sources from the country of the study itself, 40% used US data instead. For instance, 59% of studies providing a global perspective used US studies as their main data source. Even 18% of Chinese studies used US data, especially for cases considering fuels other than coal. The problem is that the application and context of the underlying source are potentially incompatible with the study to which the water intensity has been transferred. For example, Srinivasan et al.[42] assessed the water consumption of electricity production in India using the water intensities of Macknick et al.[19], Meldrum et al.[22] and Kyle et al.[43], who adopted data from Davies et al.[38]. Hence, most of the data originate

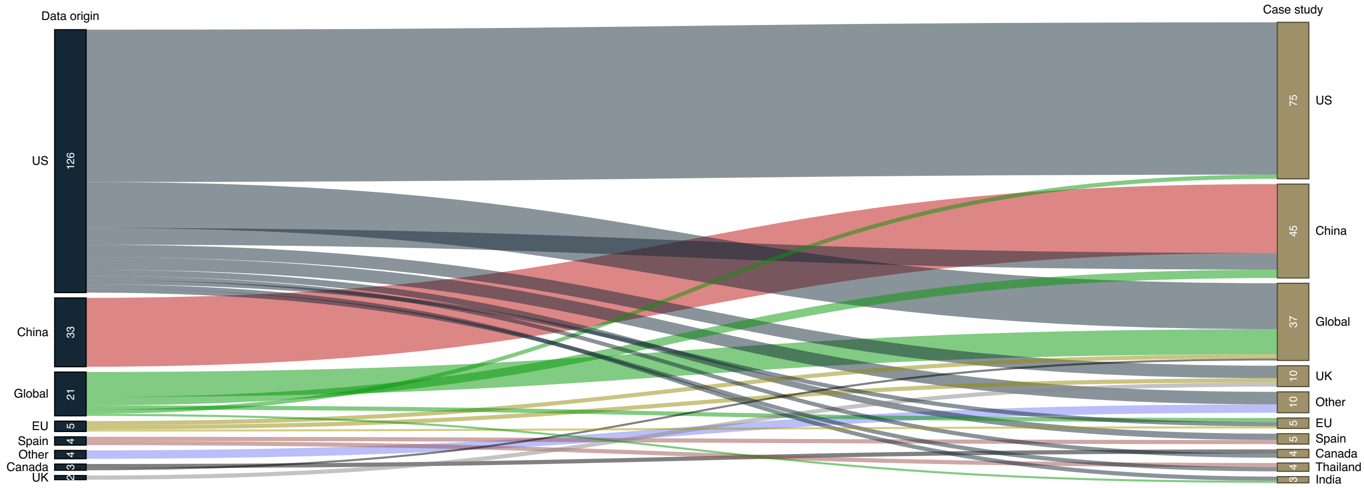

Fig. 3 | Countries of origin of data sources in relation to the countries for which assessments were made. Countries where the data sources were generated (left) linked to the case study countries to which the data sources relate (right). The category Other includes Egypt, the United Arab Emirates, the Middle East and North Africa (MENA) region and Greece.

in the United States. However, besides a few sentences to address data origin, the paper does not discuss the uncertainties and consequences of using US data sources, so that the results are probably irrelevant for the Indian electricity sector.

Original data sources double-counted in review papers

Figure 1 highlights two recent review papers, Larsen and Drews[28] and Jin et al.[30], as their work includes many WEN publications of different periods. This might suggest that both are comprehensively the most up-to-date review articles, providing essential information on the water intensities reported in the WEN literature. However, they are also more prone to act as echo chambers, so we took a closer look at the water intensities they assessed. Figure 4a,b shows the papers used as data sources of the water intensities of coal-fired TPPs in the study of Jin et al.[30]. Figure 4a shows papers when links are excluded, Fig. 4b when links are included.

Figure 4a shows a diversity of data sources (40), assuming Jin et al.[30] gave independent water intensities. However, Fig. 4b shows that 60% of these sources are related, and 50% of the water intensities originate from similar sources. For example, Jin et al.[30] included water intensities of Gleick[21] and Ali[44]. However, Ali[44] used data from Ali and Kumar[39], who used data from Meldrum et al.[22], who used data from Fthenakis and Kim[37], who used data from Gleick[21]. The pool of water intensities, therefore, includes averages that already contain water intensities that were previously averaged. For instance, the report of Gleick[21] provides average water intensities from a range of sources. Meldrum et al.[22] averaged these water intensities with additional information from other sources, including Fthenakis and Kim[37], who also used Gleick[21]. When Jin et al.[30] considered these papers separately, they calculated averages using data from Gleick[21] and Meldrum et al.[22], so the values reported by Gleick[21] were counted and averaged with different samples multiple times.

Larsen and Drews[28] identified the lack of data sources and the uncertainties of water intensities in the literature. However, they failed to see the double counting in their calculations. The sources of Larsen and Drews[28] are detailed in Supplementary Note 4.

Figure 5 provides a visual example of this double counting in the WEN literature. It shows that the average water intensities of the four papers on the left differ. Paper 4 considers the most data sources and comprises almost all the available information. However, the sources consider more than twice the same water intensity (wi). For instance, wi5 is included directly in the averages presented in paper 4, but is also included twice indirectly because it is included in papers 1 and 3. This is also the case for the recent reviews described above. Figure 5 also shows why this double counting is not evident. Averaging different water intensities from the same data pool generates value differences, suggesting new values, although the underlying water intensities are the same.

Larsen and Drews[28] and Jin et al.[30] concluded that water intensity ranges per technology remain the same after including new data sources. However, due to the echo chambers in reviews, this conclusion is questionable as it could be a reflection of the continuous use of the same data pool.

Discussion

Echo chambers have appeared in the WEN literature because of a lack of relevant data and unreflective use of easily available sources. Electricity companies have first-hand data on water use by power plants and fuel life cycles. The WEN literature relies on companies' information, which frequently lacks transparency and validation. Moreover, regulations differ among countries. Hence, data are only available in countries where regulations request energy companies to provide water use data, for example, for thermoelectric generation, including nuclear and geothermal, but not for renewable technologies using hydropower or biomass. The water intensities of hydropower or biomass, however, can be estimated on the basis of data from external and independent sources, for example, climate and crop yields, and are probably more reliable. Papers failing to report dependencies have helped the echo to continue in scientific journals as peer-review processes legitimize the unreflective use of copied data sources. This echo chamber in the WEN literature has translated into the double counting of water intensities, producing false ranges of water intensities.

Echo chambers dominate the WEN literature. A consequence of the lack of data is that scholars have transferred case study findings to different contexts, ignoring original spatial and temporal contexts, assumptions and system boundaries, and neglecting temporal and spatial differences due to climate, geography and operation conditions. This may have happened unintentionally and unknowingly. For example, Hardy et al.[45] reviewed the WEN in Spain, using the water intensities reported by Spanish authors, Rio Carillo and Frei[36] and Linares et al.[46]. The latter used US data from

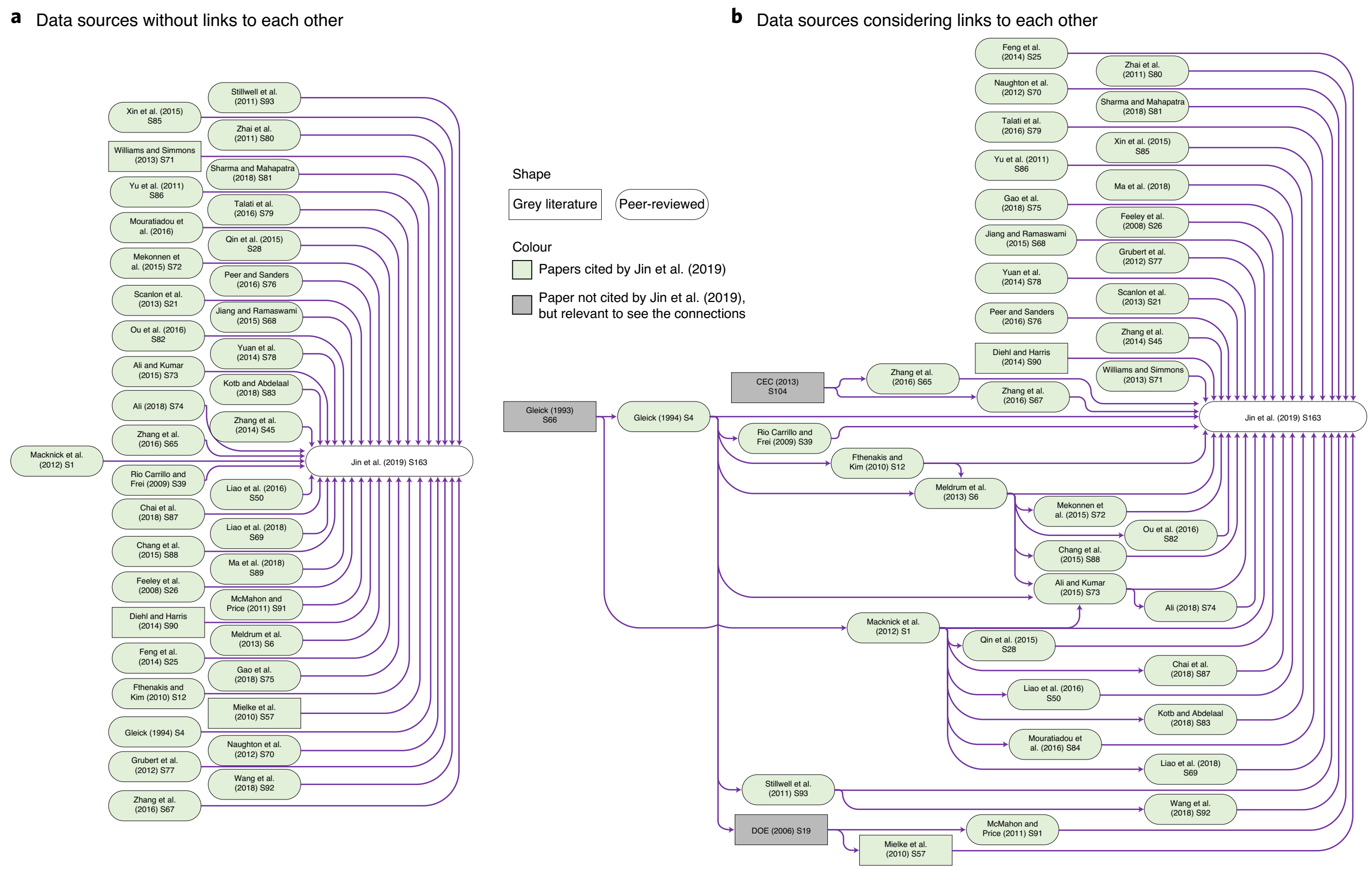

Fig. 4 | Papers used as data sources of the water intensities of coal-fired power plants in the review of Jin et al.[30]. **a**, The apparent diversity of the data sources without links to each other, showing the large number of sources considered. **b**, The chart shows how the data sources are related to each other. The full citation list of the references in this figure is given in Supplementary Note 1.

the Electric Power Research Institute (EPRI)[46]. Despite Hardy et al.[45] looking at Spanish sources, some of the water intensities used came from EPRI[47], data relevant for US electricity systems and probably incompatible with Spanish systems as they have different geography, climate and electricity mix, among other things. These factors are important for assessing whether water intensities are compatible or not[48]. Similarly, authors may use software packages such as REMIND[49] assuming that estimations result from software calculations, but being unaware that these are again based on data provided by Macknick et al.[19], so that the derived data on the water use of electricity systems may be based on incompatible systems.

Echo chambers need to be identified and addressed in science to (1) avoid misuse and misinterpretation of data, (2) address uncertainty, (3) avoid biases and (4) improve the visibility of relevant information in papers that is hidden by echo chambers. Echo chambers may perpetuate misconceptions, for example, on water-intensive technologies, affecting policymaking and deterring their future deployment. For instance, some papers based on earlier data have shown that hydropower is one of the most water-intensive electricity generating technologies, for example, Gleick[50] (1992) and Mekonnen and Hoekstra[51] (2012). However, the water intensities of hydropower technologies show large variation due to geography, climate and energy planning. Recent publications, such as Vaca-Jiménez et al.[15], have shown that some hydropower technologies have relatively small water intensities, comparable to the intensities of thermoelectric power plants with once-through cooling. Policy based on old data would limit hydropower deployment, ignoring recent first-hand data questioning this 'perceived wisdom'.

Echo chambers hide uncertainty by excluding temporal and spatial limitations of existing data sources that are only relevant for studies within a specific system boundary. New papers providing water intensities based on new data sources risk being considered unreliable if their results differ from values reported in pre-eminent publications, and risk rejection in the peer-review process so that popular data sources keep echoing. Finally, echo chambers prevent new information sources from gaining visibility. Papers providing new and insightful information regarding the water intensities of electricity-generating technologies have received little attention and few citations, so new papers are drawn to the echo chambers. If echo chambers are identified, these insightful papers attract interest, and WEN sources improve.

Conclusion and implications of echo chambers in science. Just three papers dominate the WEN literature, forming the main data sources for a large share of the publications in this field, creating echo chambers: Gleick[21], Macknick et al.[19] and Meldrum et al.[22] compiled grey literature from the 1970s and 1980s, focusing on the US electricity system. Most available water intensity data sources are based on compilations of decade-old data from grey literature presented as new information. Despite the validity and reliability of these sources at the time they were written, the fact that they are old implies an essential drawback for present water intensity analyses as technologies have considerably evolved. Recent studies providing first-hand data, for example, Grubert and Sanders[52] and Peer and Sanders[48], are overwhelmed by these old sources, potentially also because the new data may deviate too much from what has been repeated over time. These articles show that there

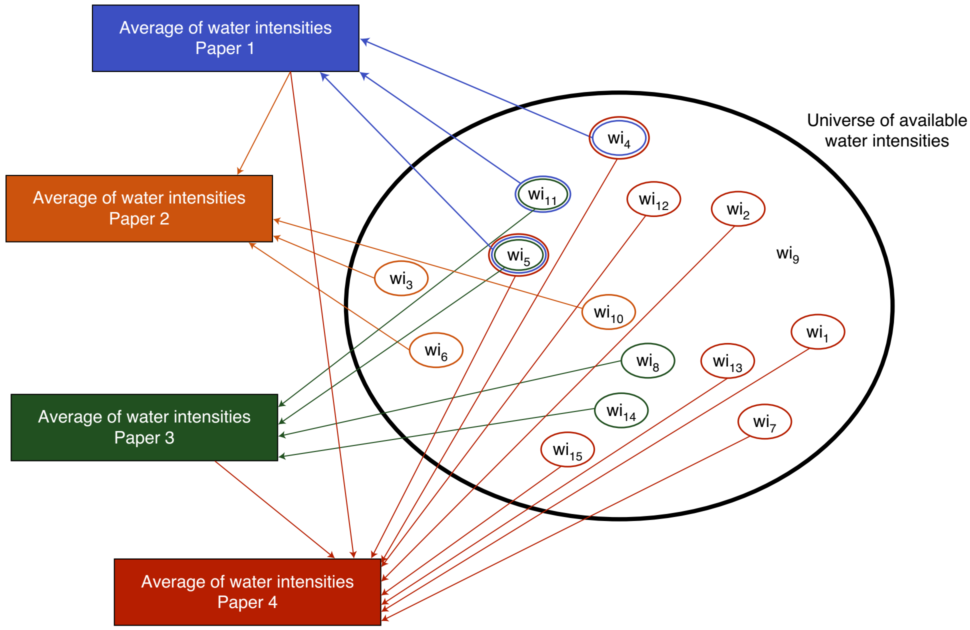

Fig. 5 | Double counting in the WEN literature. Example of the double counting of water intensities by averaging more than once the same water intensities. wi_{1-15} refer to different water intensities, and the arrows show the water intensities that were used for the hypothetical average in the different papers. For example, the water intensities wi_3, wi_6 and wi_{10} are included in the average of paper 2. Moreover, the averages of papers may also include the average of other papers. For instance, the average of paper 2 also includes the average of paper 1, as paper 2 considers that paper 1 is unique and separate from the rest of the pool of water intensities.

is a notable deviation between the water intensities reported in old sources and those reported in first-hand data, particularly in the United States.

There are two approaches to generating water intensity data sources. The first, a bottom-up approach, in which the water intensities are estimated and compiled on the basis of power plant operation data. This has the problem of a lack of data, and transferability of context and technology-specific data. The second, a top-down approach, in which statistical agencies estimate water use data per economic sector. This has the problem of being highly aggregated in terms of sector and spatial units. Most data sources that we found used data from US power plants, the first approach, although more Chinese data sources using regional or national government statistics have become available in recent years (Supplementary Note 8), the second approach. Either way, the use of American or Chinese data should be applied with great care to other technologies or countries that differ from the United States or China. Data sources echoing throughout the literature hide uncertainties, give a false sense of reliability and diversity of information, bias results, hide new information and, therefore, should be used with precaution.

Currently, reports based on these decade-old, US-based data sources have been used in a variety of applications and scenario formulations for many electricity systems around the world. Unreliable and unrelated data are present in reports that are used as inputs for national energy policy plans, for example, the DOE[13] (when they are used outside the United States) or IEA[23], which could translate echo chambers into important societal impacts.

Until recently, most science has been monodisciplinary. However, new fast-growing scientific trends connect disciplines, for example, the nexus concept in sustainability science. Scientific disciplines will probably become more interconnected in the future. For the WEN, this interconnection has caused scholars to use data sources unreflectively, probably because they were unfamiliar with a discipline outside their own. As original data become fuzzy and blurred, trends towards multi- and cross-disciplinary science accentuate echo chambers. The scientific community needs to recognize, prevent or resolve these echo chambers.

WEN echo chambers are located in data sources compiled from decade-old information and are present in recent literature as water intensities presented as new information. It is hard to identify echo chambers in conventional literature reviews, as repeated water intensities are often hidden in averages or median values. Detailed analysis can trace original sources and map them to find relationships. Nonetheless, researchers may avoid echo chambers and double accounting by following due diligence in assessing the origin, quality and constraints of the data sources they use, including differences in terminology related to water use[53]. The repetition of a data source is not problematic in itself, especially if there is an apparent lack of diversity of data sources. The problem resides on using it unreflectively and applying results to incompatible contexts. The Supplementary Information includes an additional discussion regarding the ways to avoid echo chambers.

Finally, the scientific echo chamber effect is likely to intensify in the future unless it is identified and addressed. If scientists continue to use preferred data sources and keep citing unreflectively, studies will create false perceptions of legitimacy and validity. Based on the WEN literature example, this study warns that echo chambers may be present in other scientific disciplines too, especially those that repeatedly use decade-old data sources. Future studies should address and solve echo chambers and provide reliable, spatially relevant data. In general, all scientific communities should address their echo chambers by following a full-disclosure approach, in which all sources are used reflectively.

Methods

In this study we used a hybrid literature review approach that combines the common keyword research-intensive literature review (for example, Albrecht et al.[54]) with bibliometric and network analyses (citation, co-citation and co-occurrence analyses). This approach to extensive literature reviews was adopted from Newell et al.[55]. These hybrid approaches, which include network analysis tools (for example, Liu and Mei[56]) are useful to identify research fronts, collaboration networks and influential authors.

We included as many papers as possible and identified the connections between them, understanding their contents and scopes. The papers we included on water use for electricity generation, and their fuels, were published between 1974 and 2019. The most recent paper is from November 2019. Thus, our analysis excludes papers published after this date.

The study answers five research questions:

- What consitutes the main water–electricity literature on electricity generation systems, their fuels and applied water sources?
- What are the main data sources of the WEN literature for different fuels, for example, coal, gas, nuclear, hydropower, solar, wind and biomass energy?
- What are the characteristics of available publications, that is, the origin, quality and use of first- or second-hand data?
- What are the most important echo chambers?
- What are the implications of echo chambers in science?

The method included two phases. In the first phase, steps 1–11, we gathered and consolidated WEN paper databases, clustering papers based on inputs and outputs, scopes and paper types. We focused on all the connections between papers (in terms of citations), creating an overview of existing WEN literature. In the second phase, steps 12–18, we assessed the water use data sources of papers, focusing on citations revealing connections between sources. The flow and number of papers studied between the steps are shown in Supplementary Fig. 1 in the Supplementary Methods.

Consolidating the paper's database. We first designed and constructed a database of WEN literature. In step 1, we collected peer-reviewed papers using keyword searches in Scopus and Web of Science with five keyword combinations among (1) water, (2) energy, (3) electricity, (4) fuels and (5) nexus. The keywords captured a large range of papers that included most electricity generation technologies, their fuels (including bioenergy sources) and their relation to water resources.

In step 2 we consolidated the database by removing non-accessible papers, discarding papers that were (1) not written in English, (2) not accessible due to incorrect information and (3) not peer-reviewed. The final database contained a mix of WEN and water–energy–food (WEF) papers, as the keyword search also included papers with topics related to food, agriculture or irrigation. Based on the paper title, abstract and keywords, in step 3 we eliminated papers considering food-related topics in WEF papers.

In step 4 we classified the papers based on the input and output of the studied system. Water, the input, was classified as (1) freshwater, (2) saline water and (3) wastewater. Energy, the output, was classified as (1) fuels, (2) electricity and (3) heat. Papers outside this classification, for example, papers considering only general WEN concepts, were excluded. For papers that considered water and energy as both input and output, we only considered water input for electricity systems. The classification was based on a paper's title, abstract and keywords, and if more information was required, then it was refined by a second step that involved the reading of the whole paper. An overview of the inputs and outputs of the WEN literature is provided in Supplementary Note 5.

With the database completed and classified, step 5 involved inventorying papers according to category to define the extent of coverage per topic and electricity system.

The compiled database contained papers from the keyword search. However, papers might have been left out due to limitations in the keywords. Authors may not have framed their work inside the WEN or WEF, or possibly defined them in very specific terms, which may have prevented them from turning up in the keyword search. The papers that were left out by the keyword search needed to be included in the database. Papers address a specific topic, and their authors include a variety of relevant references with a unifying theme. Therefore, instead of searching for individual references, we searched the reference lists of the papers from the keyword search to include missing papers.

In step 6 we formed a general reference list of all the papers in the database. We obtained the references of the papers in the database by extracting their metadata from the Scopus and Web of Science catalogues. Many papers in the reference list, however, did not refer to the WEN and were excluded.

In step 7 we discarded papers through a citation co-occurrence analysis, in which all referenced papers were compared and counted. By adopting Schiebel's hypothesis[57], we assumed that in databases with a similar theme, papers with many citations include important information for a specific subject. Therefore, we included only papers that appeared more than three times in the reference list. The aim of this step was to identify possible papers that were left out in the keyword search.

Any reference is counted as part of another only if both papers' metadata entirely coincide. However, during the publication process, critical reference metadata might be lost due to flawed citation. To avoid this, we compared references to homogenize the citation list in two steps. In step 8 we sorted publications in the list based on three metadata parameters (Digital Object Identifier (DOI), title and authors/year) to compare similar publications that may have different metadata. We identified several cases in which one of these parameters was wrongly written or missing. Therefore, the network analysis would have missed the citation, hindering the correct accounting of the article's citation number. We corrected the flawed or missing metadata of the publications.

Sometimes authors publish reports, or pre-prints, disclosing water intensities. This grey literature is then upgraded to full papers that are published in a scientific journal after peer review. In terms of the water intensities reported, these papers are fundamentally the same publication, for instance, the case of Macknick et al.[19] and Macknick et al.[58]. The latter is the grey-literature version of the former. We acknowledge that combined publications differ from a metadata perspective, but combining them is justified as there is no fundamental difference between the publications' content and the reported water intensities are the same. In step 9 we compared the water intensities reported in similar publications by the same authors. For this, we separated the papers by author name, and identified those publications that have the same authors, similar title and were published around the same date. Then, we compared their reported water intensities to identify whether they were similar. A list of papers that were considered to be the same publication despite metadata differences is given in Supplementary Note 6.

We repeated steps 6 and 7 once more to capture paper references added in the first round.

Identifying influential publications in the WEN. In step 10 we constructed a co-citation network with the papers in the database to find connections between papers. Co-citation is a network analysis tool used for literature connectivity assessments based on citations in which papers are considered as nodes and citations as edges between nodes[59]. Citation connections between papers were identified one by one, preferably by using the DOI number, to avoid errors arising from comparison of other metadata details, for example, title or author names.

In step 11 we listed the influential publications of the co-citation network based on two network centrality metrics: in-degrees and betweenness. These metrics are often used in network analysis to estimate network connectivity[60], showing influential publications not only in terms of higher co-occurrence, but also in terms of network connectivity. In-degrees quantifies the number of incoming links to a node, counting how many times a paper has been referenced. Betweenness indicates which papers serve as knowledge bridges by considering which nodes connect to other nodes. The co-citation network and centrality metrics were assessed using Matlab 2019b[61] and its digraph functionality (for directed networks). Additional information on the network analysis and metrics used is provided in Supplementary Note 7.

Identifying water intensity sources. In step 12 we identified the papers that refer uniquely to case studies of water intensities (water consumption and water withdrawal rates) for electricity generation and the fuels used for this purpose. Water intensities for electricity generation form only one part of papers that assess whole energy systems. For example, Okadera et al.[62] assessed the water intensity of Thailand's energy production and supply, including transportation fuels, energy exports and electricity generation. In these cases, we studied the part related to electricity generation and the fuels used for this purpose, excluding the rest.

In step 13 we assessed the data sources of the papers from step 12 by inventorying citations in the method sections and supporting information. We then followed the data sources until we arrived at the initial publication. The original data sources were traced, including peer-reviewed papers, grey literature and reported first-hand data, for example, the reports of energy companies. We acknowledged the difference between consumptive and withdrawal rates, and made sure that data sources coincided with the type of water use (consumptive or withdrawal).

In step 14 we mapped the data sources co-citation network, including connections between papers, distinguishing between water intensity data sources of power plant operation, construction, decommission and fuel life cycles. We also included peer-reviewed papers. Grey literature was included in the network analysis. However, when drawing the network map, they were only included if the information came from relevant international sources, for example, the IEA[23], or if their use was recurrent, improving the network map figures. Otherwise, they were defined in terms of their nature, whether they came from an international or national agency report on the subject, a book or a private company report. Moreover, these sources were also differentiated by country of origin.

Data origins are paramount to understanding WEN data sources. In step 15 we assessed data sources of the United States and China, the most engaged countries in the WEN[30]. Chinese data sources are often documented in Chinese. To overcome the language barrier, six Chinese scientists supported us and performed a survey of the characteristics of the Chinese sources of water intensities. An overview of the survey and the results are provided in Supplementary Note 8.

Spotting echo chambers. Social networks identify echo chambers by quantifying repetition in a network[63]. However, the repetitive use of data sources does not necessarily define an echo chamber, especially not for topics with limited data sources. The problem resides in the way these sources are used. In the WEN, echo chambers appear when water intensities are repeated through publications that

(1) considered them new, disconnected and global, even when they came from the same original data source, which was described decades ago for a specific case study, and (2) applied them to new studies without considering the operational characteristics of the original studies.

To identify echo chambers, we assessed a paper's data sources by considering (1) papers sharing common data sources, (2) the most used (popular) data sources, (3) the country of origin of the popular data sources and (4) the country of the studied electricity system.

In step 16 we identified possible echo chambers in the literature based on information gathered regarding the paper's data sources and their origins. By analysing network maps of data sources per technology (Supplementary Figs. 2–11), we focused on papers that share common data sources. Next, we highlighted papers that (1) used a data source that was connected to a previous data source, but the connection was not explicit in the paper and (2) studied a different electricity system from the one in the data source.

The approach used for the first case comprised the detailed reading of the paper. In these cases, we reviewed the method used (including information in the supporting information) to identify which water intensities were considered as input for their assessment, and how they were handled. Sometimes authors were aware of the connection and treated papers as connected. However, sometimes authors were unaware of the connection and treated them as separate, generating an echo chamber. For this group of highlighted papers, we checked whether the paper explicitly mentioned that the data sources came from a different system and discussed the limitations of the use of these sources, including an uncertainty analysis. When this was not explicit, we also considered the paper an echo chamber.

The approach used for the second case consisted of comparing the spatial and temporal boundaries of the studied system with those of the data source's system. Echo chambers appeared when authors used data of incompatible systems, whether by geography, technology, fuel source, climate or energy planning, and failed to disclose the limitations and uncertainties of using those data.

In step 17 we identified echo chambers based on the selected criteria.

Echo chambers and double counting. Echo chambers sometimes cause double counting of water intensities in review papers. In step 18 we identified the pool of water intensities used to estimate ranges, medians and averages of water intensities per power generating technology in review papers. Based on the results obtained in the previous steps, especially the network map of the data sources, we assessed whether the water intensities considered in these pools came from the same source and were duplicated.

Data availability

All datasets that support the findings of this study are publicly available, or available from the corresponding author upon reasonable request.

Received: 25 March 2020; Accepted: 20 January 2021;
Published online: 22 February 2021

References

1. IPCC: Summary for Policy Makers. In *Climate Change 2014: Mitigation of Climate Change* (eds Edenhofer, O. et al.) 1–31 (Cambridge Univ. Press, 2014); https://doi.org/10.1017/CBO9781107415324
2. IEA *Clean Energy Transitions Programme (CETP): Annual Report 2018* (OECD Publications, 2019).
3. Urban, J. J. Emerging scientific and engineering opportunities within the water-energy nexus. *Joule* **1**, 665–688 (2017).
4. Gerbens-Leenes, P. W., Hoekstra, A. Y. & van der Meer, T. H. The water footprint of bioenergy. *Proc. Natl Acad. Sci. USA* **106**, 10219–10223 (2009).
5. Mathioudakis, V., Gerbens-Leenes, P. W., Van der Meer, T. H. H. & Hoekstra, A. Y. The water footprint of second-generation bioenergy: a comparison of biomass feedstocks and conversion techniques. *J. Clean. Prod.* **148**, 571–582 (2017).
6. Murrant, D., Quinn, A., Chapman, L. & Heaton, C. Water use of the UK thermal electricity generation fleet by 2050: part 2 quantifying the problem. *Energy Policy* **108**, 859–874 (2017).
7. Smith, P. et al. Biophysical and economic limits to negative CO_2 emissions. *Nat. Clim. Change* **6**, 42–50 (2016).
8. Mekonnen, M. M., Gerbens-Leenes, P. W. & Hoekstra, A. Y. Future electricity: the challenge of reducing both carbon and water footprint. *Sci. Total Environ.* **569**, 1282–1288 (2016).
9. Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A. & Bonneau, R. Tweeting from left to right. *Psychol. Sci.* **26**, 1531–1542 (2015).
10. Jasny, L., Waggle, J. & Fisher, D. R. An empirical examination of echo chambers in US climate policy networks. *Nat. Clim. Change* **5**, 782–786 (2015).
11. Choi, D., Chun, S., Oh, H., Han, J. & Kwon, T. "Taekyoung". Rumor propagation is amplified by echo chambers in social media. *Sci. Rep.* **10**, 310 (2020).
12. Farrell, J. Politics: echo chambers and false certainty. *Nat. Clim. Change* **5**, 719–720 (2015).
13. *Energy Demands on Water Resources: Report to Congress on the Interdependency of Energy and Water* (US Department of Energy, 2006).
14. Mekonnen, M. M., Gerbens-Leenes, P. W. & Hoekstra, A. Y. The consumptive water footprint of electricity and heat: a global assessment. *Environ. Sci. Water Res. Technol.* **1**, 285–297 (2015).
15. Vaca-Jiménez, S., Gerbens-Leenes, P. W. & Nonhebel, S. The water footprint of electricity in Ecuador: technology and fuel variation indicate pathways towards water-efficient electricity mixes. *Water Resour. Ind.* **22**, 100112 (2019).
16. Vaca-Jiménez, S., Gerbens-Leenes, P. W. & Nonhebel, S. Water-electricity nexus in Ecuador: the dynamics of the electricity's blue water footprint. *Sci. Total Environ.* **696**, 133959 (2019).
17. *Water Use for Electric Power Generation* (EPRI, 2008).
18. Liu, J. et al. Nexus approaches to global sustainable development. *Nat. Sustain.* **1**, 466–476 (2018).
19. Macknick, J., Newmark, R., Heath, G. & Hallett, K. C. Operational water consumption and withdrawal factors for electricity generating technologies: a review of existing literature. *Environ. Res. Lett.* **7**, 045802 (2012).
20. Macknick, J., Sattler, S., Averyt, K., Clemmer, S. & Rogers, J. The water implications of generating electricity: water use across the United States based on different electricity pathways through 2050. *Environ. Res. Lett.* **7**, 045803 (2012).
21. Gleick, P. H. Water and energy. *Annu. Rev. Energy Environ.* **19**, 267–299 (1994).
22. Meldrum, J., Nettles-Anderson, S., Heath, G. & Macknick, J. Life cycle water use for electricity generation: a review and harmonisation of literature estimates. *Environ. Res. Lett.* **8**, 015031 (2013).
23. IEA *World Energy Outlook 2012* 33 (OECD Publications, 2012); https://doi.org/10.1787/weo-2012-en
24. Averyt, K. et al. *Freshwater Use by US Power Plants: Electricity's Thirst for a Precious Resource* (Energy and Water in a Warming World Initiative, UCS Publications, 2011).
25. Hoff, H. *Understanding the Nexus. Background Paper for the Bonn 2011 Conference: The Water, Energy and Food Security Nexus* 1–52 (Stockholm Environment Institute, 2011).
26. Sanders, K. T. Critical review: uncharted waters? The future of the electricity-water nexus. *Environ. Sci. Technol.* **49**, 51–66 (2015).
27. Cooper, D. C. & Sehlke, G. Sustainability and energy development: influences of greenhouse gas emission reduction options on water use in energy production. *Environ. Sci. Technol.* **46**, 3509–3518 (2012).
28. Larsen, M. A. D. & Drews, M. Water use in electricity generation for water-energy nexus analyses: the European case. *Sci. Total Environ.* **651**, 2044–2058 (2019).
29. Gold, H., Goldstein, D. J., Probstein, R. F., Shen, J. S. & Yung, D. *Water Requirements for Steam-Electric Power Generation and Synthetic Fuel Plants in the Western United States* (US EPA, 1977).
30. Jin, Y., Behrens, P., Tukker, A. & Scherer, L. Water use of electricity technologies: a global meta-analysis. *Renew. Sustain. Energy Rev.* **115**, 109391 (2019).
31. Gleick, P. H. *Water in Crisis: A Guide to the World's Fresh Water Resource* (Oxford Univ. Press, 1993).
32. Inhaber, H. Water use in renewable and conventional electricity production. *Energy Sources* **26**, 309–322 (2004).
33. Spang, E. S., Moomaw, W. R., Gallagher, K. S., Kirshen, P. H. & Marks, D. H. The water consumption of energy production: an international comparison. *Environ. Res. Lett.* **9**, 105002 (2014).
34. Jornada, D. & Leon, V. J. Robustness methodology to aid multiobjective decision making in the electricity generation capacity expansion problem to minimise cost and water withdrawal. *Appl. Energy* **162**, 1089–1108 (2016).
35. Geels, F. W. & Kemp, R. Dynamics in socio-technical systems: typology of change processes and contrasting case studies. *Technol. Soc.* **29**, 441–455 (2007).
36. Rio Carrillo, A. M. & Frei, C. Water: a key resource in energy production. *Energy Policy* **37**, 4303–4312 (2009).
37. Fthenakis, V. & Kim, H. C. Life-cycle uses of water in US electricity generation. *Renew. Sustain. Energy Rev.* **14**, 2039–2048 (2010).
38. Davies, E. G. R., Kyle, P. & Edmonds, J. A. An integrated assessment of global and regional water demands for electricity generation to 2095. *Adv. Water Resour.* **52**, 296–313 (2013).
39. Ali, B. & Kumar, A. Development of life cycle water-demand coefficients for coal-based power generation technologies. *Energy Convers. Manag.* **90**, 247–260 (2015).
40. Zhang, C., Zhong, L., Fu, X., Wang, J. & Wu, Z. Revealing water stress by the thermal power industry in China based on a high spatial resolution water withdrawal and consumption inventory. *Environ. Sci. Technol.* **50**, 1642–1652 (2016).
41. Jiang, D. & Ramaswami, A. The 'thirsty' water-electricity nexus: field data on the scale and seasonality of thermoelectric power generation's water intensity in China. *Environ. Res. Lett.* **10**, 024015 (2015).

42. Srinivasan, S. et al. Water for electricity in India: a multi-model study of future challenges and linkages to climate change mitigation. *Appl. Energy* **210**, 673–684 (2018).
43. Kyle, P. et al. Influence of climate change mitigation technology on global demands of water for electricity generation. *Int. J. Greenh. Gas. Control* **13**, 112–123 (2013).
44. Ali, B. The cost of conserved water for coal power generation with carbon capture and storage in Alberta, Canada. *Energy Convers. Manag.* **158**, 387–399 (2018).
45. Hardy, L., Garrido, A. & Juana, L. Evaluation of Spain's water-energy nexus. *Int. J. Water Resour. Dev.* **28**, 151–170 (2012).
46. Linares, P., Sáenz & Sáenz de Miera, G. *Implications for Water of the World Energy Scenarios* (Economics for Energy, 2010).
47. *Water & Sustainability (Volume 3): US Water Consumption for Power Production - The Next Half Century* (EPRI, 2002).
48. Peer, R. A. M. & Sanders, K. T. Characterising cooling water source and usage patterns across US thermoelectric power plants: a comprehensive assessment of self-reported cooling water data. *Environ. Res. Lett.* **11**, 124030 (2016).
49. Luderer, G. et al. Description of the REMIND Model (Version 1.6). *SSRN Electron. J.* https://doi.org/10.2139/ssrn.2697070 (2015).
50. Gleick, P. H. Environmental consequences of hydroelectric development: the role of facility size and type. *Energy* **17**, 735–747 (1992).
51. Mekonnen, M. M. & Hoekstra, A. Y. The blue water footprint of electricity from hydropower. *Hydrol. Earth Syst. Sci.* **16**, 179–187 (2012).
52. Grubert, E. & Sanders, K. T. Water use in the United States energy system: a national assessment and unit process inventory of water consumption and withdrawals. *Environ. Sci. Technol.* **52**, 6695–6703 (2018).
53. Grubert, E., Rogers, E. & Sanders, K. T. Consistent terminology and reporting are needed to describe water quantity use. *J. Water Resour. Plan. Manag.* **146**, 04020064 (2020).
54. Albrecht, T. R., Crootof, A. & Scott, C. A.The water-energy-food nexus: a systematic review of methods for nexus assessment. *Environ. Res. Lett.* **13**, 043002 (2018).
55. Newell, J. P., Goldstein, B. & Foster, A. A 40-year review of food-energy-water nexus literature and its application to the urban scale. *Environ. Res. Lett.* **14**, 073003 (2019).
56. Liu, L. & Mei, S. Visualising the GVC research: a co-occurrence network based bibliometric analysis. *Scientometrics* **109**, 953–977 (2016).
57. Schiebel, E. Visualization of research fronts and knowledge bases by three-dimensional areal densities of bibliographically coupled publications and co-citations. *Scientometrics* **91**, 557–566 (2012).
58. Macknick, J., Newmark, R., Heath, G. & Hallett, K. C. *A Review of Operational Water Consumption and Withdrawal Factors for Electricity Generating Technologies* (NREL, 2011).
59. Zhong, S., Geng, Y., Liu, W., Gao, C. & Chen, W. A bibliometric review on natural resource accounting during 1995–2014. *J. Clean. Prod.* **139**, 122–132 (2016).
60. Fornito, A., Zalesky, A. & Bullmore, E. T. in *Fundamentals of Brain Network Analysis* (eds Fornito, A. et al.) 115–136 (Elsevier, 2016); https://doi.org/10.1016/B978-0-12-407908-3.00004-2
61. MATLAB. version 9.7.0 (R2019b). (The MathWorks Inc., 2019).
62. Okadera, T., Chontanawat, J. & Gheewala, S. H. Water footprint for energy production and supply in Thailand. *Energy* **77**, 49–56 (2014).
63. Del Vicario, M. et al. The spreading of misinformation online. *Proc. Natl Acad. Sci. USA* **113**, 554–559 (2016).

Acknowledgements
We thank Y. Shan, M. Mengdie and D. Zhao for helping us assess the characteristics and quality of the Chinese data sources. This work was supported by SENESCYT (National Secretariat of Higher Education, Science, Technology and Innovation of Ecuador).

Author contributions
S.V.-J. performed data analysis and wrote the paper. P.W.G.-L. conceptually designed the study and edited the paper. S.N. and K.H. supervised the project and provided an outline for the paper and edited it.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41893-021-00686-7.

Correspondence and requests for materials should be addressed to S.V.-J.

Peer review information *Nature Sustainability* thanks Lu Liu, Kelly Sanders and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.