Towards harmonization of water quality management : a comparison of chemical drinking water and surface water quality standards around the globe

Reference:
Van Winckel Tim, Cools Jan, Vlaeminck Siegfried, Joos Pieter, Van Meenen Els, Borregán-Ochando Elena, Van Den Steen Katleen, Geerts Robbe, Vandermoere Frédéric, Blust Ronny.- Towards harmonization of water quality management : a comparison of chemical drinking water and surface water quality standards around the globe
Journal of environmental management - ISSN 1095-8630 - 298(2021), 113447
Full text (Publisher's DOI): https://doi.org/10.1016/J.JENVMAN.2021.113447
To cite this reference: https://hdl.handle.net/10067/1807650151162165141

Towards harmonization of water quality management: A comparison of chemical drinking water and surface water quality standards around the globe

Tim Van Winckel[1,2**], Jan Cools[2**], Siegfried E. Vlaeminck[1*], Pieter Joos[1,3], Els Van Meenen[3], Elena Borregán-Ochando[2], Katleen Van Den Steen[3], Robbe Geerts[4], Frédéric Vandermoere[4], and Ronny Blust[5]

[1] *Department of Bioscience Engineering, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium*

[2] *Institute of Environment and Sustainable Development, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium*

[3] *Water-Link, Mechelsesteenweg 111, 2840 Rumst, Belgium*

[4] *Department of Sociology, University of Antwerp, Sint-Jacobstraat 2, 2000 Antwerpen, Belgium*

[5] *Department of Biology, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium*

* Corresponding author: siegfried.vlaeminck@uantwerpen.be

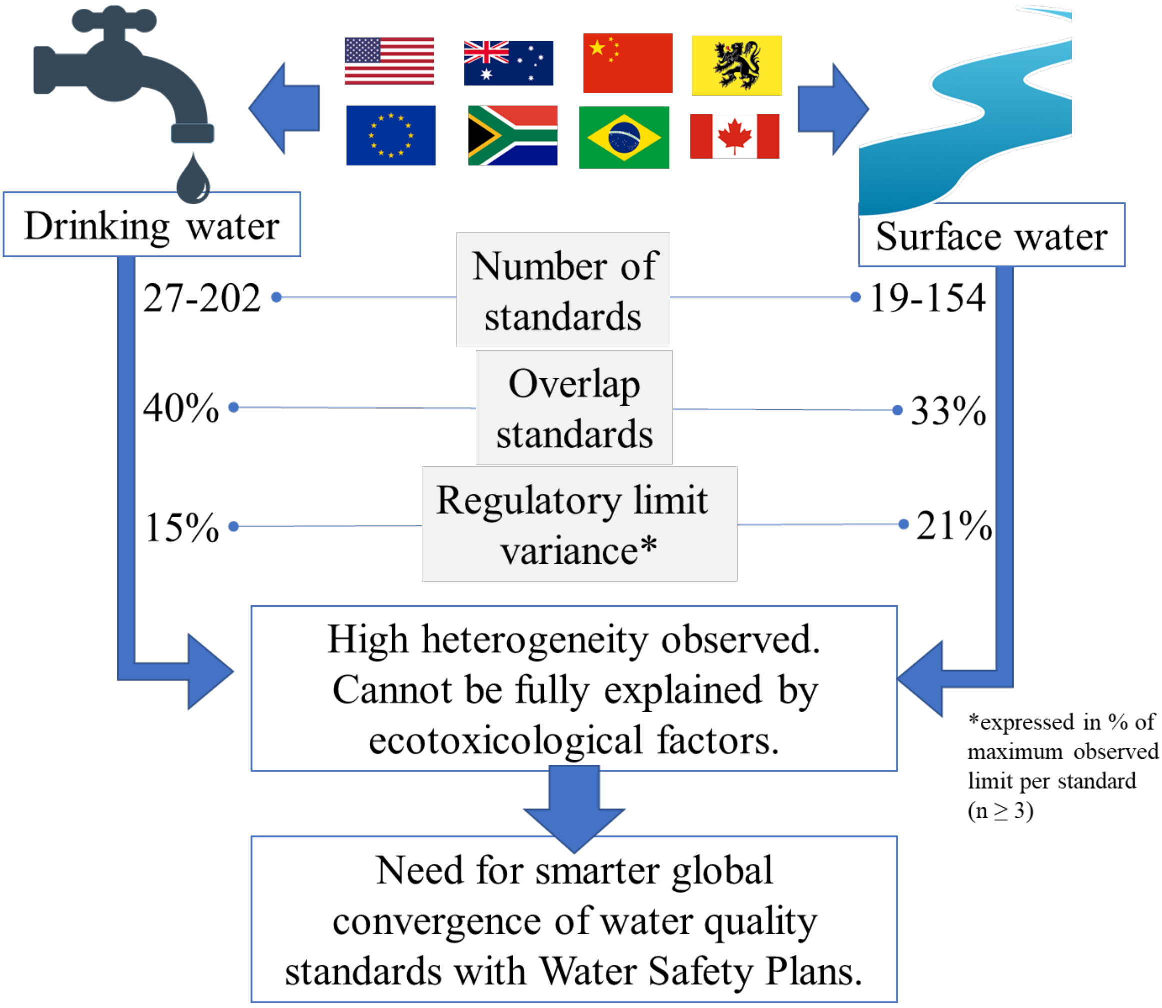
** Equally contributed to this manuscript

Abstract

Water quality standards (WQS) set the legal definition for safe and desirable water. WQS impose regulatory concentration limits to act as a jurisdiction-specific legislative risk-management tool. Despite its importance in shaping a universal definition of safe, clean water, little information exists with respect to(dis)similarity of chemical WQS worldwide. Therefore, this paper compares chemical WQS for drinking and surface water matrices in eight jurisdictions representing a global geographic distribution: Australia, Brazil, Canada, China, the European Union, the region of Flanders in Belgium, the United States of America, and South Africa. The World Health Organization's list is used as a reference for drinking water standards. Sørensen–Dice indices (SDI) showed little qualitative similarity in the compounds that are regulated in drinking water (median *SDI* = 40%) and surface water (median *SDI* = 33%), indicating that the heterogeneity within a matrix is substantial at the level of the standard. Quantitative similarly for matching standards was higher than the qualitative per Kendall correlation (median = 0.73 and 0.58 for drinking water and surface water respectively), yet variance observed within standards remained inexplicably high and pronounced for organic compounds. Variations in WQS were more pronounced for organic compounds. Most differences cannot be easily explained from a toxicological or risk-based point-of-view. Historical development, ease of measurement, and (toxicological) knowledge gaps on the risk of a vast number of organic compounds are theorized to be the drivers. Therefore, this study argues for more tailored, risk-based approach in which standards incorporated into water safety plans and dynamically set for compounds that are persistent and could pose a risk for human health and/or aquatic ecosystems. Global variations in WQS should in the end not necessarily be avoided but be globally harmonized, yet flexible to ensure a global up-to-date definition of safe and desirable water everywhere.

Keywords: World Health Organization, Risk assessment, Water safety plan, Environmental policy, Toxicology, Potable water

Graphical abstract



1. Introduction

The natural and anthropogenic water cycles have been subjected to increased stress throughout the last few decades due to rapid urbanization, intensification and global change (Reid et al., 2019; Schwarzenbach et al., 2010). In order to safeguard public and ecosystem health, legislative jurisdictions worldwide have developed water quality standards (WQS) as part of their regulatory framework. WQS describe the desired condition of a water matrix and how this condition can be achieved. This is frequently done by setting a regulatory limit on the allowable concentration of a specific parameter or chemical compound. These limits can differ depending on the applicable water matrix, can be acute or chronic, and could be summations of groups (e.g. pesticides). WQS therefore act as a jurisdiction-specific legislative risk-management tool. However, no harmonized approach to water quality risk-management exists globally. Water Safety plans (WSP), which require water quality monitoring along the drinking water production chain, are being rolled out in multiple jurisdictions, yet do not replace the fixed list of contaminants (WHO and IWA, 2017; World Health Organization, 2009). One objective of World Health Organization's (WHO) WSP framework is to create a dynamic list of WQS based on high-risk contaminants measured throughout the water chain and which therefore acts on the current threats within the water supply chain.

In respect of drinking water, the WHO introduced the Guidelines for Drinking-Water Quality (GDWQ) to comprehensively propose what constitutes "safe and desirable" drinking water and details how jurisdictions can achieve this status (World Health Organization, 2017). The GDWQ formulates non-binding guidelines WQS for microbial, chemical, radiological, and acceptability (taste, odour) which jurisdictions can adapt. Indeed, no global framework for water quality standards defining safe drinking water exists. A World Health Organization

(2018) review on the adaptation of WHO-recommended WQS in 104 countries revealed that more than half make direct or indirect reference to the GDWQ. However, the review did not elaborate beyond listing the number of countries that adapted of the WHO-recommended standards and spread (min, median, max) of the regulatory limit. Boyd (2006) found that there are discrepancies in the measured compounds and the corresponding standards between Canada, the European Union (EU), the United States of America (USA), and Australia. However, the analysis was predominantly descriptive, nor not include emergent powers such as Brazil and China or discussed surface water regulations.

For surface water, no generally accepted global guidelines for WQS exist, creating potential disparities among jurisdictions in terms of which contaminants are to be measured and what regulatory limits are to be set. Furthermore, 60-80% of the worldwide fresh water usage (domestic, industrial, or agricultural) originates from surface water, a matrix that is most at risk of potential contamination supplies (FAO, 2016; Wada et al., 2014). A UN Water (2016) review of the surface water quality regulatory instruments in the EU, South Africa, Canada, the USA, and China found a wide diversity of regulatory frameworks between countries. Specific quantitative information on the standards used in the reviewed countries, however, was not provided.

From a risk-management point-of-view, the WQS list should be very either flexible with the help of a measuring campaign within a WSP framework (with a focus on human and ecosystem health), or so comprehensive that most current and future threats are covered. Available literature indicates that neither is currently the case. While a multitude of publications have explored the implementation and efficacy of WSP (Roeger and Tavares, 2018; String and Lantagne, 2016; Tsoukalas and Tsitsifli, 2018), no in-depth comparison of current chemical

WQS has, to the authors' knowledge, been conducted so far. Additionally, the meta-analysis by String and Lantagne (2016) revealed that many WSP-related publications do not highlight monitoring approaches in spite of international and cross-governmental organisations (i.e. WHO and EU) indicating interest in harmonization of standards and improved comparability of monitoring results (European Commission, 2013; 2015; World Health Organization, 2017). The WSP approach suggests that monitoring should not be a fixed checklist, but instead a more flexible instrument driven by the risks in the water supply zone. One could therefore argue that a WSP should not attempt to meet the WQS demands, but rather itself should set the WQS. A major caveat in the WSP approach as proposed by the WHO is that it is currently focused on human health as the objective rather than ecosystem health. Unifying efforts like WSP with a focus on ecosystems are not well established. The European Union (EU) is a notable exception with its Water Framework Directive (2000/60/EC), where it imposes a list of priority substances for which environmental quality standards (EQS) are set by law for substances in surface waters (Directive 2008/105/EC). Every member state must incorporate these priority substances in their surface water legislation. Other legislations have similar ecosystem-centric WSP frameworks, yet none have, to the authors' knowledge, the legal power that the EU priority substances list has.

Jurisdictions to this day work with fixed lists of WQS incorporated into a rigid and slow-to-change legislative system and may not be fully prepared for new and emerging threats. The need for a comprehensive and critical study into the origin, meaning, and impact of (dis)similarities between fixed WQS lists within and between drinking and surface water regulations is dire. In response to these knowledge gaps, this study will compare the traditional fixed-list chemical WQS used for drinking and surface water quality for eight jurisdictions (Australia, Brazil, Canada, China, Europe, Flanders region in Belgium, USA, and South

Africa). The GDWQ will be used as baseline for drinking water quality standards. The central goal of this study is to provide concrete insights in the global differences in drinking water and surface water quality standards and whether this heterogeneity can be explained and is justified. Focus is placed on chemical standards, covering heavy metals, pesticides and emerging pollutants which are categorized in inorganic and organic contaminants. While microbiological, ecological and radiological monitoring are essential components of a water quality monitoring programme, these are considered outside the scope of this paper. The importance of adequate microbial standards has been widely discussed in the literature (Cabral, 2010; Ramírez-Castillo et al., 2015).

2. Material & Methods

2.1. Water quality standards

The WQS for both drinking and surface water were obtained for eight jurisdictions around the world (Australia, Brazil, Canada, China, Europe, the Flanders region in Belgium, the USA, and South Africa), in addition to the WHO guidelines. The selection of countries included in this study was determined in order to ensure that every (populated) continent is reflected in the analysis. Moreover, attention was given to the amount of influence the jurisdiction has on the continent and internationally. The standards were sourced from legal publications and no distinction was made between enforceable standards and guidelines. As such, henceforth "regulatory limit" can denote both legally binding and recommended concentration limits. A summary of the different sources can be found in **Table 1**.

WQS are set, implemented, and enforced at different levels in the jurisdictions studied. They are legally set (or recommended) at the federal level in all countries studied but Belgium. The federal governments of Brazil and China have the power to enforce the WQS for both drinking water and surface water **(Table 1)**. Brazil, however, delegates the monitoring and actual enforcement of the standards to the individual states albeit with varying efficacy (Val et al., 2019). The USA has enforceable drinking water standards; however, the surface water standards are set on a statewide level. The USA's Environmental Protection Agency does provide a list with recommended standards, which is used in this study. Moreover, the Fifth Amendment potentially challenges the actual enforcement of these standards on private properties (Carlton, 2016), but these legal nuances were not considered. In South Africa, the federal government enforces the drinking water standards and provides guidance on surface water quality standards to the provinces. In Australia and Canada, the federal government can only issue guidelines and has delegated the competence on water quality standards to the provinces. The European Union has the power to set the drinking water and surface water quality standards but delegates the implementation and enforcement to the EU Member States. Belgium, one of the EU member states, delegates that power to the its regions. In this paper, the region of Flanders was used as an example of the incorporation of WQS legislation at EU-level.

Formulation of surface water quality standards is generally more complex than their drinking water counterparts. The lack of international guidelines and the wide scope they aim to serve are theorized to be the main drivers. Surface water standards should not only protect public health, but also species in freshwater ecosystems. Therefore, surface water quality standards typically have acute and chronic regulatory limits. The former denotes a concentration that, once breached, will lead to acute toxicity, i.e. mortality or serious toxicological effects over a short exposure period. Chronic regulatory limits on the other hand aim to manage chronic toxicity, i.e. the adverse effects after continuous exposure of a chemical compound for a prolonged period. Chronic standards are typically set as a maximum yearly average. Within this study, only chronic standards were considered. These standards are stricter than acute ones and are generally more relevant from an ecological perspective to assess long-term effects.

In some jurisdictions, surface waters are classified based on their ecological status or use type. China has five classes (PRC Enviromental Protection Bureau, 2002), ranging from the most stringent Class 1 (applicable to spring water and water in national nature reserves) to the least stringent Class 5 (applicable to surface water for agricultural or general landscaping). Brazil has a similar structure with three classes (Conselho Nacional do Meio Ambiente Brasil, 2005). In both cases, the class applicable to surface water reserved to produce drinking water was used. For the Chinese legislation, this was considered Class 2, while for Brazil Class 1 was used.

The guidelines provided by the Australian federal government take a different approach (ANZECC and ARMCANZ, 2000). Rather than assigning multiple classes to different types of fresh water, trigger values are deduced from a combination of single species toxicity tests, i.e. a species sensitivity distribution. These trigger values were subsequently extrapolated using the method described in Aldenberg and Slob (1993) to account for multiple contaminants. The trigger value indicates the total percentage of aquatic species protected, ranging from 80% to 99%. For the purpose of this study, the trigger values corresponding to 95% of the species within the ecosystems protected were chosen in order to allow a fair comparison to the other surface water standards and guidelines.

Clearly, each jurisdiction has its own nuances attributed to the proper implementation of WQS. However, these nuances are out of scope of this study given that its purpose is to look towards the diversity and heterogeneity of listed WQS and their respective regulatory limit, not the effectiveness of their implementation or enforcement. Therefore, no distinction is made between mandated standard and recommendations within this paper. Additionally, the assumption was made that the guidelines set by these countries' federal governments are adapted in a similar or less stringent variant (Australia Productivity Commission, 2000; Bakker, 2011)

2.2. Statistics

Given the large differences between sample sizes and data distributions, non-parametric tests were used throughout the study. When a parametric metric was used (e.g. variance), normality was checked. All statistics were performed in R version 4.0.1.

2.2.1. *Summary statistics*

The median and median average deviation (MAD) were used as summary statistics for a given distribution. The MAD is a robust statistic, meaning that it does not make assumptions of the underlying distribution (e.g. outliers). The following notation will be used throughout the manuscript to denote median and MAD: *median [MAD, n]*, with *n* the number of observations in the distribution. Both median and MAD were calculated in R using the median() and mad() function present in the stats package (v 3.6.2).

2.2.2. *Heterogeneity indices: Sørensen–Dice index & Kendall rank correlation*

The Sørensen–Dice index is a measure used to quantify the degree of qualitative similarity (i.e. presence or absence of an element) between two groups (Dice, 1945; Sørensen, 1948). The SDI was used here to compare the similarity of the WQS being monitored, i.e. whether jurisdictions monitor the same WQS or not. The SDI is given by the Equation (1):

$$SDI = \frac{2M_{11}}{2M_{11} + M_{10} + M_{01}} \tag{1}$$

Where SDI denotes the Sørensen–Dice index , $M_{11}$the total numbers of standards present in both jurisdiction A and B, and $M_{01}$and M_{10} is the total numbers of standards present only in jurisdiction A or B respectively.

Quantitative similarity is obtained by the Kendall correlation because a normal distribution cannot be assumed. Kendall was preferred over Spearman given its slight edge in robustness and better handling of small sample sizes (Croux and Dehon, 2010).

SDI was calculated using the dist.binary() function, method 5 of the ade4 package (version 1.7-16). Kendall correlation coefficients were obtained using the “kendall” method of cor() function in the stats package.

2.2.3. *Hypothesis testing*

To determine if two samples significantly differed from each other, the Mann-Whitney-U test was used as non-parametric counterpart of the Student's t-test. The Kruskal-Wallis rank sum test was used to test if all samples were from the same distribution and can be seen as the non-parametric version of the analysis of variance test. Both Mann-Whitney-U and Kruskal-Wallis were calculated using the kruskal.test() and wilcox.test() respectively from the stats package.

2.2.4. *Levene's test*

Levene's test was used to test for homo- or heterogeneity of variances between the regulatory limits among the different jurisdictions. The test was performed with the leveneTest() function present in the car package (version 3.0-8) (Fox and Weisberg, 2018).

2.2.5. *Multidimensional scaling*

Multidimensional scaling (MDS) was used to visualize the high-dimensional relationships between WQS lists in a 2D plane. Regulatory limits were standardized using the Wisconsin double standardization technique (Cottam et al., 1978). Commonly used in ecological datasets, in this study, the regulatory limits were standardized based on the maximum observed concentration of a specific standard across jurisdictions and then divided by the number of standards present in the list of jurisdictions. This ensures equal emphasis among standards and their respective regulatory limits. Bray-Curtis distances were thereafter calculated on the standardized data to highlight potential dissimilarities. The calculated distance matrix was then scaled to its principal coordinates using principal coordinates analysis (PCoA) (Borg and Groenen, 2005). Wisconsin double standardization, Bray-Curtis distances and coordinate calculations were executed using the wisconsin(), vegdist() functions in the vegan package (version 2.5-7), whereas PCoA coordinates were calculated using the pcoa() function in the ape (version 5.5) package (Oksanen et al., 2013).

3. Results and discussion

Across the eight jurisdictions investigated, 360 and 298 unique standards were identified for drinking water and surface water respectively. Out of the 360 drinking water standards identified, 39 (11%) were inorganic and 321 (89%) organic. In respect to surface water, 42 (14%) out of the 298 standards were considered inorganic whereas 256 (86%) organic standards were shared among the jurisdictions. Across all jurisdiction, 132 standards were shared between drinking and surface water regulations, which is 24% of the 526 unique standards found across all jurisdictions and matrices. A complete list of the standards and respective regulatory limits for both matrices can be consulted in the **Supplemental A Table SA1 and SA2.**

3.1. Heterogeneity within regulatory standards for drinking water

3.1.1. *Heterogeneity in number of standards measured*

Figure 1 shows the total number of standards included in the jurisdiction's respective drinking water quality **(A)** and surface water quality **(B)** regulations. For drinking water, the EU has the lowest number of mandated chemical compounds (29). However, the EU does require that all relevant pesticides and their metabolites must be measured and cannot individually exceed 0.1 μg/L, making its true count a lot higher. Flanders, with 193 listed standards, the jurisdiction with the second highest number of mandated compounds, is a practical application of this directive. A total of 140 compounds that Flanders mandates as a result of the European Drinking Water Directive (98/83/EC) are pesticides and their relevant metabolites. However, this number can change depending on what is put on the Flemish "watchlist" (see **Section 3.5**). The Australian recommendations were the most comprehensive (202 standards), whereas the

South African legislation provided the least amount of coverage (33 standards) for a single jurisdiction.

Most WQS (215; 59.7%) were unique to a single jurisdiction (**Figure 2A**), predominantly Flanders and Australia, which are also the jurisdictions with the largest monitoring programmes. A full breakdown of SDIs per country can be consulted in **Supplemental B Figure SB1A**. Only 13% (47) of the standards were measured by five or more legislations and could therefore be considered widespread. Overall, the median [MAD, n] Sørensen–Dice index (SDI) for drinking water was 0.40 [0.16, 36]. The SDI can be interpreted as a percentage of overlap. Therefore, half of the combinations shared more than 40% of their combined compounds. Note that the SDI only considered the presence or absence of a standard, not the regulatory limit. Brazil and the USA shared the highest similarity (62%) between their collective standards, followed by Brazil and both China and Canada (60%). Flanders had the lowest amount of overlap (29% [4%, 8]). This was predominantly caused by the significantly larger number of standards within Flemish legislation compared to most other jurisdictions. South Africa was a close runner-up with a median SDI of 32% [11%, 8] and moreover considerably more heterogeneous as indicated by the larger median absolute deviation. Whereas South Africa had a 50% overlap with Europe, the African country shared only 17% of the collective standards with Australia. The full matrix of SDI can be consulted in **supplementary A Figure S1A**.

Interestingly, 100% of drinking water standards shared by all jurisdictions are inorganic compounds. Excluding standards attributed to only a single jurisdiction (n = 145), half of the inorganic compounds (n = 31), i.e. the median, were measured by six (or more) out of eight jurisdictions. This was only three out of eight jurisdictions for organics (n = 114). Five organic compounds were listed by all expect South Africa and could therefore be considered universal. These are 1,2-dichloroethane, benzene, benzo(a)pyrene, vinyl chloride and total trihalomethanes. Benzene, 1,2-dichloroethane and vinyl chloride are important precursors for industrial more complex molecules but are also considered carcinogenic (Kielhorn et al., 2000; Rana and Verma, 2005). Benzo(a)pyrene is a byproduct of incomplete combustion. Benzo(a)pyrene can be found in exhaust fumes from diesel vehicles, wood burning, and coal tar (Srogi, 2007). Trihalomethanes are important disinfection byproducts potentially produced in drinking water production (Liang and Singer, 2003).

3.1.2. Heterogeneity in regulatory limits

Kendall's rank correlations were performed between the different jurisdictions to elucidate the relationship between their regulatory limits of matching compounds. Note that only compounds present in both jurisdictions were considered and thus sample sizes were unequal and considerably smaller than the jurisdiction's total standards. These ranged from 15 between South Africa and both Europe and Canada, to 65 between Flanders and Australia. The median sample size was 38 [21, 36]. A matrix of the number of overlapping standards as well as the correlation coefficients can be consulted in **Supplemental B Figure SB1B/C**. Overall, a strong correlation was found between jurisdictions. The median overall correlation was 0.73 [0.19, 36]. Regulatory limits of Flanders had significantly lower correlation with Australia, Canada, and the WHO. This is predominantly because of the stringent regulatory limit for pesticides imposed by the European Union compared to Australia and Canada, both of which also have numerous pesticides within their list but determined the regulatory limit per individual pesticide.

Both the EU and South Africa had the best median correlation with the other jurisdictions (0.90 [0.04, 8] and 0.90 [0.02, 8] respectively), but also the lowest matching compounds (median 21 and 20 for the EU and South Africa respectively). The high correlation with other jurisdictions therefore is a consequence of low similarity, though not because of lower statistical confidence. Their matching standards are more universally accepted. Indeed, standards listed by South Africa are frequently measured by other jurisdiction: 50% of its standards are also measured by five [4.4, 33] or more other jurisdictions. The average normalized variance in regulatory limits of standards measured by South Africa was 8.2%, significantly smaller ($p = 3.5 \times 10^{-5}$) than the average variance of all compounds measured by three or more jurisdictions 15.1% (**Figure 3**).

Indeed, **Figure 3** elucidates that overall a large spread in variances between regulatory limits can be observed, ranging from 0 (all equal) to 32% of the maximum observed concentration of the standard (oxamyl). Only three compounds - aluminum, arsenic and sodium - have equal regulatory limits across all probed jurisdictions. Di-(2-ethylhexyl)-phthalate, a common plasticizer which acts as endocrine disruptor, is the organic WQS with the lowest variance – 2.6% of maximum identified concentration (10 µg/L) – though is only mandated by four out of the nine jurisdictions. With respect to drinking water, the variance in limits between countries was generally higher **(Figure 3)**.

Figure 4A describes the spread of the regulatory limits within a jurisdiction. Levene's test was found to be significant (p-value = 0.008), indicating that the spread of the regulatory limits was not consistent between jurisdictions. However, this was expected given the large heterogeneity in the amount and types of standards measured between jurisdictions. All jurisdiction lists apart from Flanders and South Africa have a median regulatory limit between 10-50 µg/L, indicating some high-level similarities in terms of regulatory limits between jurisdictions. South Africa had a median concentration of 200 μg/L, whereas Flanders' median was 0.1 μg/L. South Africa's discrepancy could mainly be attributed to its disproportionately large ratio of organic to inorganic standards. Whereas the median organic to inorganic standards measured by a jurisdiction was 3.1:1, South Africa's ratio was 0.38:1. Regulatory limits of inorganics are generally higher, explaining South Africa's distribution shift to the right in **Figure 4A**. Flanders' low median limit is due to the EU's rule that all pesticides have a regulatory limit of 0.1 μg/L.

3.1.3. WHO guidelines: how widespread is their implementation in jurisdictions?

The guidelines for drinking water formulated by the WHO contain 91 recommended standards, which include heavy metals, various pesticides, and persistent pollutants that pose a threat to human health. The list contains 20 inorganic compounds and 71 organic compounds. All investigated jurisdictions incorporated some of the recommendations provided by the WHO, ranging from 21 (South Africa) to 61 (Australia) of the 91 recommended standards (**Figure 1A**). In total, 82 WHO-recommended standards were identified. The nine compounds that none of the jurisdictions covered include herbicides (mecroprop, molinate, chlorotoluron), halogenated acetonitriles and other organics. The lack of implementation of these compounds is widespread across the globe. They only had a median 12% [7%, 11] implementation rate in the WHO study that investigated 104 jurisdictions (World Health Organization, 2018). It should be noted that the three herbicides are covered under the EU's pesticide limit per Directive 98/83/EC and therefore apply to its member states as well. However, the overall (human) toxicity of these compounds is low (Younes and Galal-Gorchev, 2000), and typically not very well understood (i.e. acetonitriles) (Villanueva et al., 2014). This presumably induces a low sense of urgency for both policy makers and toxicologists. Water safety plans, as discussed later on, could potentially better identify whether these compounds are frequently measured and as such catalyze the study of frequently measured pollutants which are not well understood. The overall median implementation rate for the eight jurisdictions discussed in this study was 87.5% for inorganics and 37.5% for organic standards. The 104 countries in the WHO study had a median implementation rate of 85.6% and 20.7% for inorganics and organics respectively. A two-sided Mann-Whitney *U* test revealed that the sample of jurisdictions used here is representative for inorganic standards (p-value = 0.32, n = 20), yet not fully representative for organics (p-value = 0.02, n = 71). However, a significant Spearman correlation (r = 0.61; p-value = 3×10^{-8}; n = 71) was found between this study's sample and the WHO study.

Interestingly, the large majority (82%) of the drinking water standards that appear on the list of five or more jurisdictions were the ones recommended by the WHO. This indicates that the WHO recommendation list potentially influenced jurisdictions in shaping their lists, subsequently promoting a level of homogeneity. However, despite this influence on the lists of individual jurisdictions, **Figure 5** shows that a lot of variation remains with respect to the regulatory limit of WHO-recommended standards between jurisdictions. In **Figure 5**, the ratio between the jurisdiction's regulatory limit and its corresponding WHO's recommended limit was calculated for all standards. A ratio smaller than one points at a lower, more stringent limit than the WHO recommends, and a ratio higher than one indicated a higher, less stringent limit.

Visually, **Figure 5A** hinted at a higher degree of agreement between regulatory limits for inorganic than organic compounds. However, a Mann–Whitney *U* test revealed that this was not significant (p-value = 0.17, n = 334). There were also significant differences between countries per Kruskal-Wallis rank sum test (χ^2 = 39.0, p-value = 1.9x10-6, df = 7, n = 334), even when only organics were taken into account (χ^2 = 44.6, p-value = 1.6x10-7, df = 7, n = 208). No significant effect of jurisdiction was found when looking only at inorganic parameters (χ^2 = 10.4, p-value = 0.17, df = 7, n = 126), meaning all jurisdictions made similar modifications to the WHO guidelines for inorganic standards.

Overall, the spread in ratios is quite large, ranging from 1000 times more stringent than the WHO, to 30 times less stringent (**Figure 5A**). This spread is peculiar, given that most WHO regulatory limits are based on toxicology studies, with methodologies and relevant background information well documented. Notable here is that the USA, Canada and Australia had a considerable larger fraction of ratios above one (**Figure 5B**), indicating more relaxed regulatory limits than the WHO prescribes, though this trend disappears when looking at all measured standards (**Figure 4A**). Flanders, and by extension the EU, has a considerable number of standards with lower regulatory limits than the WHO's recommendations. South Africa adapted the least number of standards from the WHO, though all but three of those were set at the regulatory norm that the WHO recommends.

To conclude, the multidimensional scaling (MDS) analysis shown in **Figure 6B** still resulted in four distinct regions of similarity. The MDS was based on the Bray-Curtis distance after double normalization of the data (maximum and total, see Material and Methods) and thus takes both the presence/absence as well as the regulatory limit into account. Based on the MDS, we can conclude that the USA, Canada, and Brazil measure similar WHO parameters with limits, whereas the same is true for China, Australia, and Flanders/EU. South Africa does not strongly correspond with any other jurisdiction yet is more associated with the USA/Canada/Brazil cluster than with China/Australia.

3.2. Heterogeneity in surface water regulations: How does it compare to drinking water?

The number of standards in surface water quality lists in the studied jurisdictions ranged from 19 (South Africa) to 154 (Flanders) parameters as shown in **Figure 1B**. On top of the 49 standards mandated by the EU, Flanders added an additional 105 organic compounds as WQS. Four jurisdictions (Australia, Brazil, Canada, and China) have surface water quality standards for 79-102 compounds. Similar to drinking water, the majority of standards were only present in a single jurisdiction (Figure 2B). The length of the surface water WQS list was for many jurisdictions considerably different compared to their corresponding drinking WQS lists. For Australia, USA, and South Africa, the number of standards for surface water is substantially smaller (40-60%) than for drinking water. For China, Brazil, and Flanders, approximately the same number of standards exist for surface water compared to drinking water. However, the MDS analysis in **Figure 6A** revealed that they are not necessarily the same standards nor do they have similar regulatory limits. The exception is China, where the distance between the coordinates was small. Moreover, China's surface water parameters were more like other jurisdiction's lists of drinking water standards than surface water standards. Within this study, China's 'Class II' regulations were used which appertain bodies of water used for drinking water production. This may have biased the aforementioned similarity. Canada and the EU were the only jurisdictions where more parameters are to be monitored in surface water compared to drinking water (**Figure 1B**). Overall, there is also more disagreement between jurisdictions in terms of what parameters should be measured, as indicated by the smaller percentage of standards measured by five or more jurisdictions (8.5% vs 12.3%). Every jurisdiction incorporates some of their drinking water standards within their surface water regulations (**Figure 1B**). Both China and Brazil incorporated the largest number of drinking water standards in their surface water standards (36 each). South Africa, the USA and the EU incorporated less than 10 standards each. Interestingly, all but one shared standard are WHO recommended parameters, magnesium, a major ion and essential nutrient, being the exception.

The overall median SDI index for surface water standards was 0.33 [0.09, 28], significantly lower than for drinking water standards per Mann-Whitney *U* test (p-value = 0.001, n = 28), indicating a higher amount of disagreement between surface water standards despite the overall smaller sample size (Figure SB1). The maximum SDI (0.52, between China and Brazil) was also about 16% lower than the maximum for drinking water (0.62, between USA and Brazil). The median SDI per country ranged from 22% to 41% for the European Union and Brazil respectively, which resulted in a significant difference between the jurisdictions' SDI distribution as per the Kruskall-Wallis test (χ^2 = 14.6, p-value = 0.04, df = 7, n = 56).

The median Kendall correlation of the regulatory limits of the set of matching compounds between jurisdictions overall was high (0.58 [0.13, 28]), yet significantly lower than drinking water per Mann-Whitney *U* test (p-value = 5×10^{-5}, n = 28). The USA shared the highest overall median correlation with the EU (0.66), though the latter had a higher spread (MAD = 0.09 versus 0.036). Like South Africa's trend in the context of drinking water regulations, the USA's surface water standards match poorly with other jurisdictions (median = 16 standards) and these standards are generally better accepted. A similar explanation is true for the European Union (median = 16 standards). The EU's list is comprised of 'priority substances' and thus is generally more agreed upon worldwide. China's list had the lowest median correlation because it more resembles a typical drink water list as shown in **Figure 6A** and discussed above.

The lower overall observed Kendall correlation translated into a generally higher variance within regulatory limits of standards measured by three or more jurisdictions (**Figure 3**). The mean was 21% of the maximum observed concentration per standard which was significantly higher (p-value = 2×10^{-5}, n = 163) than the variance observed in drinking water standards. The overall spread of the variances, however, was similar between drinking and surface water. Only one standard (dichloromethane) had equal regulatory limits across all jurisdictions (Brazil, China, EU, and Flanders) that included it.

A similar trend is visible in **Figure 4B**, where the overall distribution of regulatory limits within a jurisdiction is visually more varied than for surface water. In particular, the 5% percentile value is significantly lower (p-value = 0.003) in surface water compared to drinking water. Indeed, the majority of standards present in both the drinking and surface water lists of a jurisdiction had a lower regulatory limit for surface water as visualized in **Figure 7**. The reason that surface water standards are typically different from drinking water regulations could be attributed to the following: (i) ecosystems are toxicologically more complex and diverse than a single species as humans. (ii) aquatic species complete their entire life cycle in water, (iii) exposure is continuous while drinking is not (the recommended fluid intake for humans is two to three litres per day (Gleick, 1996)), and (iv) Application of safety factors (or uncertainty factors) which are generally more conservative for humans than for ecosystems to keep the risk in the human population as low as possible (i.e. human risk is managed at the individual level, ecosystem risk at the population or community level) (European Chemical Agency, 2008). Whereas the disconnect between drinking water and surface water makes sense on a toxicological level, it could lead to scenarios in which a limit is violated for the source (surface water) but not the final product (drinking water), which from a policy or legislative viewpoint is inefficient. However, not all drinking water is produced from surface water. Ground water is also commonly used, and not elaborated on in this study. Additionally, when a body of water is designated for drinking water production, the standards may be more aligned. This is visible in **Figure 7** with Brazil and China, whose surface water quality lists are for bodies of water used for the production of drinking water, because their median ratios are in both cases close to or equal to one. Last, jurisdictions with a suboptimal distribution network may experience deterioration of water quality throughout the supply chain. Slightly more relaxed standards at the final sink could therefore ease some regulatory pressure.

3.3. A summarizing overview of WQS using multidimensional scaling

Figure 6A presents a summarizing helicopter view of differences and similarities between and within drinking and surface water quality standard lists using multidimensional scaling. The MDS analysis gives a visual representation of the Bray-Curtis distances between WQS lists that were standardized based on maximum observed concentration per standard and number of standards in a given list. Bray-Curtis considers absence or presence of a standard, as well as its regulatorily limit. The closer lists are together, the more similar they are.

Drinking water lists (upright triangles) are clustered relatively far away from the surface water lists (downwards triangles), meaning that overall, both are dissimilar in both what is measured and the limit. This was also apparent from the more in-depth analysis performed in **Section 3.2**. On a jurisdiction level, the distance between the drinking water and surface water lists is connected. Here, the USA and the EU are very far apart, confirming the differences stated above. Similarly, one can see that the surface water lists that have an effect on drinking water catchment areas (Brazil, China) are closer to the drinking water cluster than others. The drinking water cluster is more tightly packed than the surface water one, confirming the analyses above that show more heterogeneity within surface water standards.

Within a given matrix, surface water lists are more diverse than and thus more spread out confirming the analysis performed above. Here, Flanders and EU's lists were similar, which makes sense given that Flanders' list is based on the priority substances dictated by the European Union. For drinking water, the MDS analysis showed that the jurisdictions that followed the WHO's recommendations more closely are clustered together, whereas Flanders, Australia, South Africa, and the EU are more spread out.

3.4. Why inorganic standards are more broadly accepted

Based on the previous analyses, one can conclude that the traditional inorganic standards are widely accepted and incorporated into legislations, whereas more heterogeneity exists for organic standards such as pesticides, persistent organic pollutants, and other harmful organics. This heterogeneity could be explained by a number of reasons: (i) historically limited documentation and understanding of the risk of a vast number of harmful organics, (ii) high demands on analytical sensitivity (sub-micro and nanogram per liter) and the need for multiple complex and expensive analytical instruments (Noguera-Oviedo and Aga, 2016; Schmidt, 2018), and (iii) the cost, complexity and therefore capacity to continuously operate a monitoring network (Behmel et al., 2016).

A wide range of organic compounds are harmful. The eChemPortal, the global gateway to information on the properties, hazards, and risks of chemicals, holds information on more than 800,000 substances (OECD, 2020). Prioritization of potential high-risk substances requires an understanding of their occurrences, transformation pathways, and toxicity in the environment, which in view of their number cannot be comprehensive. However, initiatives such as REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) tackle this issue by harmonizing the reporting and legislation pertaining to potential (toxicological) hazards of chemicals on a European level (Hengstler et al., 2006; Williams et al., 2009). REACH further works with a 'read-across assessment framework' (RAAF) which allows for grouping chemicals that are expected to exhibit the same toxicological properties. This would decrease the amount of substances that need regulation and require an individual toxicological limit (European Chemicals Agency, 2017).

The risk of most of the inorganics, such as heavy metals, has been known for a large part of history. The Romans understood the toxicity of lead, arsenic, and copper (Retief and Cilliers, 2000) and many (heavy) metals could be detected with reasonable accuracy in the 19th century with help of spectroscopy developed by Kirchhoff and Bunsen (Thomas, 1991). The advancement in our knowledge of the presence and risks associated with organic substances in water has only been developed over the last few decades, in parallel with the appearance of highly sensitive analytical instrumentation (Noguera-Oviedo and Aga, 2016; Schmidt, 2018). Chromatography coupled with mass spectrometry is required for a majority of the organic compounds, and only recent advancement here has allowed for the resolution and sensitivity required. The setting of drinking and surface water standards also requires information on the toxicity of the compounds under acute and chronic exposure scenarios and this is reasonable well documented for the most common heavy metals, but only for a rather limited number of organic compounds. One could therefore hypothesize that a predominant reason for the broader embrace of inorganic standards is that the study, detection and toxicology is better understood and agreed upon. Moreover, the creation and embrace of novel policies is generally a slower process than the science it depends on (Smith, 2017).

Historical knowledge gained on the risk of organic compounds in water quality was typically related to an understanding of the chemical drivers of public health or environmental crisis. Examples are the pesticide DDT affecting (predatory) seabird eggs (Cox, 1991; Risebrough et al., 1967), the carcinogenicity of benzo[a]pyrene and other polycyclic aromatic carbons because of chimney sweeps' carcinomas (Cook et al., 1933), the bio-accumulative toxicity and consequent global termination of polychlorinated biphenyls (PCB) (de Boer, 2005), and the modern example of the concerns surrounding perfluorooctanoate (PFOA) spread and exposure (Steenland et al., 2010; Trudel et al., 2008). Toxicological studies are typically lengthy, costly, complicated, and raise ethical concern due to the testing on animals (Rand, 2020; Scholz et al., 2013). Precautionary measures are often taken for high-risk organics such as pesticides. The EU mandates a blanket-wide regulatory limit of 0.1 μg/L for pesticides and their (relevant) metabolites regardless of their actual toxicity. The total concentration of pesticides cannot surpass 0.3 μg/L. While this can be an effective risk-mitigation approach, it could potentially put unnecessary strain on municipalities that need to adhere to these strict limits. Rigorous toxicological testing of compounds that occur in the matrix is preferred to blanket-wide limits as set for pesticides. The REACH framework could be extended to regulatory limits for WQS.

3.5. A Risk-based approach: A smarter way to protect human health and the environment?

Water safety plans are the practical outcome of this paradigm shift to risk-based approaches. The WHO formulated the basis of a water safety plan, which now acts as foundation for many jurisdictions' own water safety plans (World Health Organization, 2009). The WHO defines a water safety plan as "The most effective means of securing the safety of a drinking water supply (…) through the use of a comprehensive risk assessment and risk management approach that encompasses all steps in the water supply from catchment to consumer." Australia, Canada, China, Brazil and the European Union have, amongst many other jurisdictions not discussed in this study, rolled out some form of water safety plan (WHO and IWA, 2017).

Whereas there is no formal legal inclusion of the term "water safety plan" in the EU's Drinking Water Directive (98/83/EC), Article 7 & 8 in combination with Annex II and III (Commission Directive 2015/1787) do enforce increased implementation of risk-based approaches, such as water safety plans. In Belgium, drinking water regulations are left to the regions (Flanders, Brussels, Wallonia). The Flanders Environment Agency (*Vlaamse milieumaatschappij*, VMM) is responsible for the enforcement of the Flemish implementation of the Directive 98/83/EC.

To comply with the EU's risk-based approach, the VMM created a framework which utilizes a "watchlist" of chemicals that are not standardized in the Drinking Water Directive but could potentially end up in the drinking water. These do not have regulatory limits but drinking water municipalities are obliged to qualitatively determine their potential presence. The watchlist is based on three pillars: (i) presence in raw water catchment areas during measurement campaigns carried out by all drinking water municipalities, (ii) degree of national sales of individual pesticides, (iii) the octanol/water partition coefficient (K_{ow}) of organics determined in (i). A low K_{ow} is correlated with a decreased removal efficiency in drinking water production technologies. Organics on the established watch list need to be routinely screened. The produced watchlist is updated every 1-3 years. Currently, 255 compounds are on the watch list, ranging from pesticides, metabolites, personal care products, and pharmaceuticals. A total of 135 pesticides/metabolites are currently in the list and were also incorporated in the Flemish WQS list for the analyses performed above (e.g. pesticides and metabolites).

If a compound without standardization is repeatedly detected in the drinking water matrix, VMM will issue a precautionary limit. This is done based on a hybrid approach of the Dutch threshold of toxicological concern (TTC) (Kroes et al., 2005) and German health-related indication value (*Gesundheitlichen Orientierungswert*, GOW) (des Umweltbundesamtes, 2003). Precautionary limits can range from 0.01 to 27 µg/L. So far, 22 compounds have been given a precautionary limit, whereas 97 compounds have not.

Flanders provides an example of a practical application of a dynamic WQS list through the concept of a watchlist. Many approaches to water safety plans exist across legislations and the authors do not imply the Flemish model is superior to others. It does show, however, the general trend that jurisdictions are moving away from regulating drinking water purely from a rigid legal perspective striving for homogeneity across jurisdictions, to a more proactive, dynamic, and flexible system where the actual risks for a catchment area or aquatic ecosystem are systemically mapped and mitigated, albeit with the creation of heterogeneity between standards. It is therefore clear that a global paradigm shift is required in order to migrate to a risk-based approach were modern techniques such as non-target screening of micropollutants are utilized to scan the water matrices for reoccurring (organic) compounds and regulatory limits are set based on an international library of toxicological data such as the REACH framework. In this way, a smart and transparent unification of water quality standards can be achieved while only measuring the compounds relevant to the area.

4. Conclusion

Water quality standards objectify the definition of safe water. Therefore, one might expect a high degree of agreement between jurisdictions. This paper, however, concludes the contrary. A large variation exists in the number of standards incorporated in a jurisdiction's legislation as well as in their respective regulatory limits. This holds true for both drinking water and surface water regulations.

Jurisdictions generally agree more on what compounds to measure for drinking water and their respective regulatory limit, most likely influenced by the recommendation list published by the WHO. However, even some WHO-recommended compounds differed up to a thousand-fold. Surface water regulations were generally more complex and diverse than those formulated for drinking water. This was evident in the lower Sørensen–Dice index and correlation between legislations, and the generally bigger scatter observed in the multidimensional scaling. Whereas some degree of heterogeneity can be expected due to the more complex ecosystem it's designed to protect, regulatory limits should converge, which was not observed in this study.

Standards with the lowest variation in both matrices were predominantly inorganic, reflecting our longer standing and knowledge of the adverse effects of this relatively limited set of compounds. The high number of organic parameters that could be present in either matrix is therefore bigger and thus toxicological diversity can be expected. Therefore, water safety plans, such as the one rolled out in Flanders and by other legislations, in combination with integration in the REACH-like platform, could be a useful tool to convert a majority of the unexplained and seemingly arbitrary heterogeneity into functional variation based on local risk. In the end, while the nature of the compounds measured - especially in surface water matrices - can differ from place to place, regulatory limits should not. They should be derived from internationally accepted standards unless future evidence is presented that indicates that sensitivities are different across water types, climate zones, and species diversity.

Acknowledgements

This work was funded with a gift from Water-Link to the University of Antwerp (Chair Water-link). The authors would like to thank Gustavo Gomes De Sousa and Weiqiang Zhu for their help sourcing and translating the Brazilian and Chinese legislations, Janne Spanoghe for the constructive criticism, and Helen Lyons for the diligent proofreading of the manuscript.

References

Aldenberg, T. and Slob, W. 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. Ecotoxicology and environmental safety 25(1), 48-63.

ANZECC and ARMCANZ 2000. Australian and New Zealand guidelines for fresh and marine water quality. Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Canberra, 1-103.

Australia Productivity Commission (2000) Arrangements for Setting Drinking Water Standards: International Benchmarking, Productivity Commission.

Bakker, K. (2011) Eau Canada: The future of Canada's water, ubc Press.

Behmel, S., Damour, M., Ludwig, R. and Rodriguez, M.J. 2016. Water quality monitoring strategies - A review and future perspectives. Sci Total Environ 571, 1312-1329.

Borg, I. and Groenen, P.J. (2005) Modern multidimensional scaling: Theory and applications, Springer Science & Business Media.

Boyd, D.R. (2006) The Water We Drink: An International Comparison of Drinking Water Quality Standards and Guidelines: a Report Prepared for the David Suzuki Foundation, David Suzuki Foundation.

Cabral, J.P. 2010. Water microbiology. Bacterial pathogens and water. International journal of environmental research and public health 7(10), 3657-3703.

Canadian Council of Ministers of the Environment (2002) Canadian environmental quality guidelines, Canadian Council of Ministers of the Environment.

Carlton, J. 2016. Drought by Fifth Amendment: Debunking Water Rights as Real Property. BYU J. Pub. L. 31, 409.

Conselho Nacional do Meio Ambiente Brasil 2005. Resolução CONAMA nº 357/2005. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento. Diário Oficial [da] União (53).

Cook, J.W., Hewett, C. and Hieger, I. 1933. 106. The isolation of a cancer-producing hydrocarbon from coal tar. Parts I, II, and III. Journal of the Chemical Society (Resumed), 395-405.

Cottam, G., Goff, F.G. and Whittaker, R. (1978) Ordination of plant communities, pp. 185-213, Springer.

Cox, C. 1991. Pesticides and birds: from DDT to today's poisons. Wildl. Dis 24(1), 51-61.

Croux, C. and Dehon, C. 2010. Influence functions of the Spearman and Kendall correlation measures. Statistical methods & applications 19(4), 497-515.

de Boer, J. (2005) Encyclopedia of Analytical Science, pp. 214-225, Elsevier Science Ltd.

des Umweltbundesamtes, E. 2003. Bewertung der Anwesenheit teil-oder nicht bewertbarer Stoffe im Trinkwasser aus gesundheitlicher Sicht. Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz 46, 245-248.

Dice, L.R. 1945. Measures of the amount of ecologic association between species. Ecology 26(3), 297-302.

European Chemical Agency 2008. Characterisation of dose [concentration]-response for environment.

European Chemicals Agency 2017 Read-Across Assessment Framework (RAAF).

European Commission 2013. Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy. Off. J. Eur. Union 226, 1-17.

European Commission 2015. European Council directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption Off. J. Eur. Union 260.

FAO 2016 AQUASTAT Main Database (FAO), F.a.A.O.o.t.U.N. (ed).

Gleick, P.H. 1996. Basic water requirements for human activities: meeting basic needs. Water international 21(2), 83-92.

Health Canada 2017. Guidelines for Canadian drinking water quality—Summary table. Water and Air Quality Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa.

Hengstler, J., Foth, H., Kahl, R., Kramer, P., Lilienblum, W., Schulz, T. and Schweinfurth, H. 2006. The REACH concept and its impact on toxicological sciences. Toxicology 220(2-3), 232-239.

Kielhorn, J., Melber, C., Wahnschaffe, U., Aitio, A. and Mangelsdorf, I. 2000. Vinyl chloride: still a cause for concern. Environmental health perspectives 108(7), 579-588.

Kroes, R., Kleiner, J. and Renwick, A. 2005. The threshold of toxicological concern concept in risk assessment. Toxicological sciences 86(2), 226-230.

Liang, L. and Singer, P.C. 2003. Factors influencing the formation and relative distribution of haloacetic acids and trihalomethanes in drinking water. Environmental science & technology 37(13), 2920-2928.

Ministério da Saúde Brasil 2017. Portaria de consolidação nº 5, de 28 de setembro de 2017. Consolidação das normas sobre as ações e os serviços de saúde do Sistema Único de Saúde. Diário Oficial [da] República Federativa do Brasil.

NHMRC and NRMMC 2011. Australian Drinking Water. Guidelines Paper 6 National Water Quality Management Strategy. Commonwealth of Australia.

Noguera-Oviedo, K. and Aga, D.S. 2016. Lessons learned from more than two decades of research on emerging contaminants in the environment. J Hazard Mater 316, 242-251.

OECD 2020 eChemPortal.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P., O'hara, R., Simpson, G., Solymos, P., Stevens, M. and Wagner, H. 2013. Community ecology package. R package version, 2.5-7.

PRC Enviromental Protection Bureau 2002. 地表水环境质量标准 GB 3838–2002 (Environmental Quality Standards for Surface Water). Environmental Protection Bureau, Beijing.

PRC Minestry of Health 2006. 生活饮用水卫生标准 GB 5749-2006 (Standards for Drinking Water Quality)

Ramírez-Castillo, F.Y., Loera-Muro, A., Jacques, M., Garneau, P., Avelar-González, F.J., Harel, J. and Guerrero-Barrera, A.L. 2015. Waterborne pathogens: detection methods and challenges. Pathogens 4(2), 307-334.

Rana, S. and Verma, Y. 2005. Biochemical toxicity of benzene. Journal of Environmental Biology 26(2), 157-168.

Rand, G.M. (2020) Fundamentals of aquatic toxicology: effects, environmental fate and risk assessment, CRC press.

Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T., Kidd, K.A., MacCormack, T.J., Olden, J.D. and Ormerod, S.J. 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. Biological Reviews 94(3), 849-873.

Retief, F. and Cilliers, L. 2000. Poisons, poisoning and the drug trade in ancient Rome. Akroterion 45(1), 88-100.

Risebrough, R.W., Menzel, D.B., Martin, D.J. and Olcott, H.S. 1967. DDT residues in Pacific sea birds: a persistent insecticide in marine food chains. Nature 216(5115), 589-591.

Roeger, A. and Tavares, A.F. 2018. Water safety plans by utilities: A review of research on implementation. Utilities Policy 53, 15-24.

Schmidt, T.C. 2018. Recent trends in water analysis triggering future monitoring of organic micropollutants. Analytical and bioanalytical chemistry 410(17), 3933-3941.

Scholz, S., Sela, E., Blaha, L., Braunbeck, T., Galay-Burgos, M., Garcia-Franco, M., Guinea, J., Kluver, N., Schirmer, K., Tanneberger, K., Tobor-Kaplon, M., Witters, H., Belanger, S., Benfenati, E., Creton, S., Cronin, M.T., Eggen, R.I., Embry, M., Ekman, D., Gourmelon, A., Halder, M., Hardy, B., Hartung, T., Hubesch, B., Jungmann, D., Lampi, M.A., Lee, L., Leonard, M., Kuster, E., Lillicrap, A., Luckenbach, T., Murk, A.J., Navas, J.M., Peijnenburg, W., Repetto, G., Salinas, E., Schuurmann, G., Spielmann, H., Tollefsen, K.E., Walter-Rohde, S., Whale, G., Wheeler, J.R. and Winter, M.J. 2013. A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. Regul Toxicol Pharmacol 67(3), 506-530.

Schwarzenbach, R.P., Egli, T., Hofstetter, T.B., Von Gunten, U. and Wehrli, B. 2010. Global water pollution and human health. Annual review of environment and resources 35, 109-136.

Smith, Z.A. (2017) The environmental policy paradox, Routledge.

Sørensen, T.J. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons, I kommission hos E. Munksgaard.

South Africa Department of Water Affairs and Forestry 1996. South African Water Quality Guidelines Volume 7 Aquatic Ecosystems.

South Africa Department of Water Affairs and Forestry 2001. Compulsory National Standards for the Quality of Potable Water. SA Goverment Gazette 22355.

Srogi, K. 2007. Monitoring of environmental exposure to polycyclic aromatic hydrocarbons: a review. Environmental Chemistry Letters 5(4), 169-195.

Steenland, K., Fletcher, T. and Savitz, D.A. 2010. Epidemiologic evidence on the health effects of perfluorooctanoic acid (PFOA). Environmental health perspectives 118(8), 1100-1108.

String, G. and Lantagne, D. 2016. A systematic review of outcomes and lessons learned from general, rural, and country-specific Water Safety Plan implementations. Water Science and Technology: Water Supply 16(6), 1580-1594.

Thomas, N.C. 1991. The early history of spectroscopy. Journal of chemical education 68(8), 631.

Trudel, D., Horowitz, L., Wormuth, M., Scheringer, M., Cousins, I.T. and Hungerbühler, K. 2008. Estimating consumer exposure to PFOS and PFOA. Risk Analysis: An International Journal 28(2), 251-269.

Tsoukalas, D.S. and Tsitsifli, S. 2018 A critical evaluation of Water Safety Plans (WSPs) and HACCP implementation in water utilities, p. 600.

UN Water 2016 Towards a Worldwide Assessment of Freshwater Quality: A UN-Water Analytical Brief.

USA Environmental Protection Agency 2002. National primary drinking water regulations: Long Term 1 Enhanced Surface Water Treatment Rule. Final rule. Federal register 67(9), 1811.

USA Environmental Protection Agency 2009. National Recommended Water Quality Criteria - Aquatic Life Criteria

Val, A.L., Bicudo, C.E.d.M., Bicudo, D.d.C., Pujoni, D.G.F., Rosado, F., Spilki, I.d.S.N., Hespanhol, I., Cirilo, J.A., Tundisi, J.G. and Val, P. 2019. Water quality in Brazil. Water Quality in the Americas, 103.

Villanueva, C.M., Kogevinas, M., Cordier, S., Templeton, M.R., Vermeulen, R., Nuckols, J.R., Nieuwenhuijsen, M.J. and Levallois, P. 2014. Assessing exposure and health consequences of chemicals in drinking water: current state of knowledge and research needs. Environmental health perspectives 122(3), 213-221.

Vlaamse Overheid 2018. Decreet integraal waterbeleid (Decree integral water policy) 18 july 2003. Belgisch Staatsblad, 99909-100041.

Wada, Y., Wisser, D. and Bierkens, M.F.P. 2014. Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. Earth System Dynamics 5(1), 15-40.

WHO and IWA 2017. Global Status Report on Water Safety Plans: A review of proactive risk assessment and risk management practices to ensure the safety of drinking-water. World Health Organization, International Water Association.

Williams, E.S., Panko, J. and Paustenbach, D.J. 2009. The European Union's REACH regulation: a review of its history and requirements. Critical reviews in toxicology 39(7), 553-575.

World Health Organization (2009) Water safety plan manual: step-by-step risk management for drinking-water suppliers, World Health Organization.

World Health Organization 2017 Guidelines for drinking-water quality, WHO.

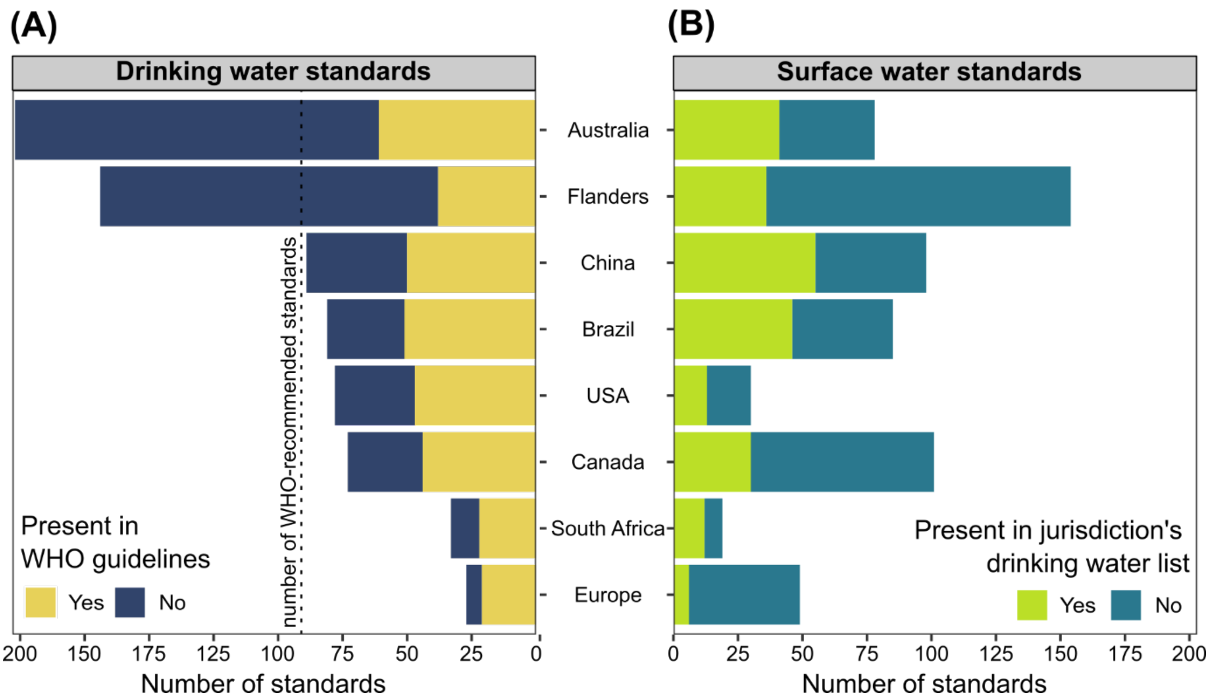
World Health Organization 2018. A global overview of national regulations and standards for drinking-water quality.

Younes, M. and Galal-Gorchev, H. 2000. Pesticides in drinking water—a case study. Food and chemical toxicology 38, S87-S90.

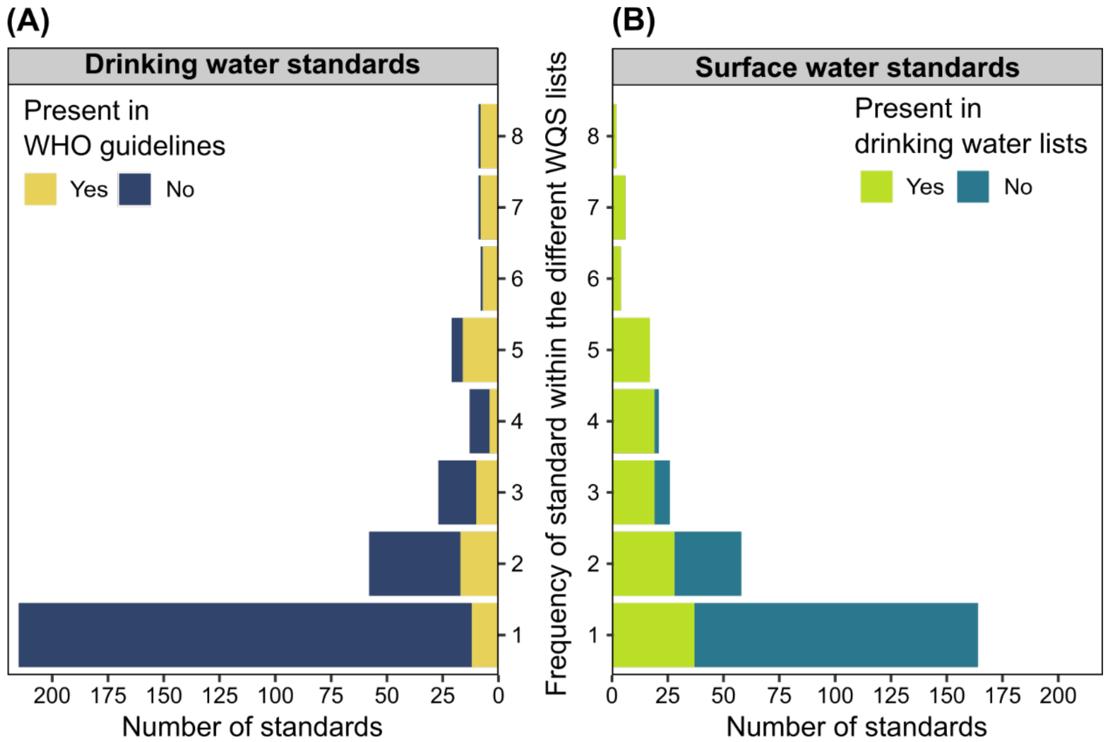
Table 1. List of drinking and surface water quality standards legislations and recommendations used in this study. Legislation names are translated into English. Original titles can be consulted in the reference list.

Jurisdiction	Drinking water quality standards regulations			Surface water quality standards regulations		
	Name	Enforc-able?	Reference		Enforc-able?	Reference
Australia	Australian Drinking Water. Guidelines Paper 6: National Water Quality Management Strategy	NO	NHMRC and NRMMC (2011)	Australian and New Zealand guidelines for fresh and marine water quality	NO	ANZECC and ARM-CANZ (2000)
Brazil	Consolidation Ordinance No. 5, of September 28, 2017. Consolidation of norms on health actions and services of the Unified Health System.	YES	Ministério da Saúde Brasil (2017)	Resolution CONAMA nº 357/2005. It disposes on the classification of the water bodies and environmental guidelines for its framing	YES	Conselho Nacional do Meio Ambiente Brasil (2005)
Canada	Guidelines for Canadian drinking water quality—Summary table.	NO	Health Canada (2017)	Canadian environmental quality guidelines	NO	Canadian Council of Ministers of the Environment (2002)
China	Standards for Drinking Water Quality (GB 5749-2006)	YES	PRC Minestry of Health (2006)	Environmental Quality Standards for Surface Water (GB 3838–2002)	YES	PRC Enviromental Protection Bureau (2002)
European Union	Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption	YES	European Commission (2015)	Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013	YES	European Commission (2013)
Flanders (Belgium)	Integral water policy decree of 18 July 2003	YES	Vlaamse Overheid (2018)	Integral water policy decree of 18 July 2003 (VLAREM annex 2)	YES	Vlaamse Overheid (2018)
South Africa	Compulsory National Standards for the Quality of Potable Water (SANS 241)	YES	South Africa Department of Water Affairs and Forestry (2001)	South African Water Quality Guidelines Volume 7: Aquatic Ecosystems	NO	South Africa Department of Water Affairs and Forestry (1996)
United States of America	National primary drinking water regulations: Long Term 1 Enhanced Surface Water Treatment Rule	YES	USA Environmental Protection Agency (2002)	National Recommended Water Quality Criteria	NO	USA Environmental Protection Agency (2009)
WHO	Guidelines for drinking-water quality	NO[1]	World Health Organization (2017)	-		

Figures



<u>Figure 1.</u> **(A)** Total number of standards in each jurisdiction's drinking water regulations. Differentiation is made between standards that are also present in the WHO guidelines (yellow bar) and that are absent (blue bar). **(B)** Total number of standards in each jurisdiction's surface water regulations. Surface water standards that are also present in the jurisdiction's drinking water regulations are indicated in green.



<u>Figure 2.</u> Number of standards that co-occur a certain amount of time between different WQS lists for both **(A)** drinking water lists and **(B)** Surface water lists.

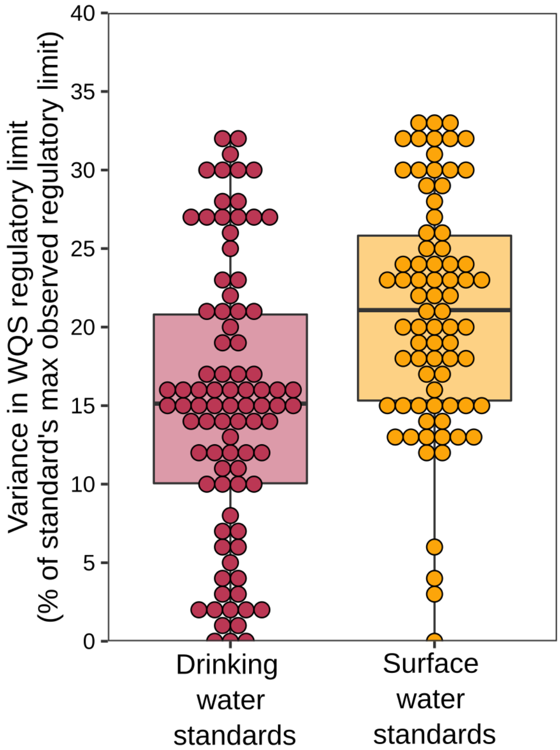


Figure 3. Variance observed in regulatory limits for a certain standard expressed in the percentage of the maximum observed regulatory limit of the respective standard. Only standards occurring in three or more jurisdictions are included.

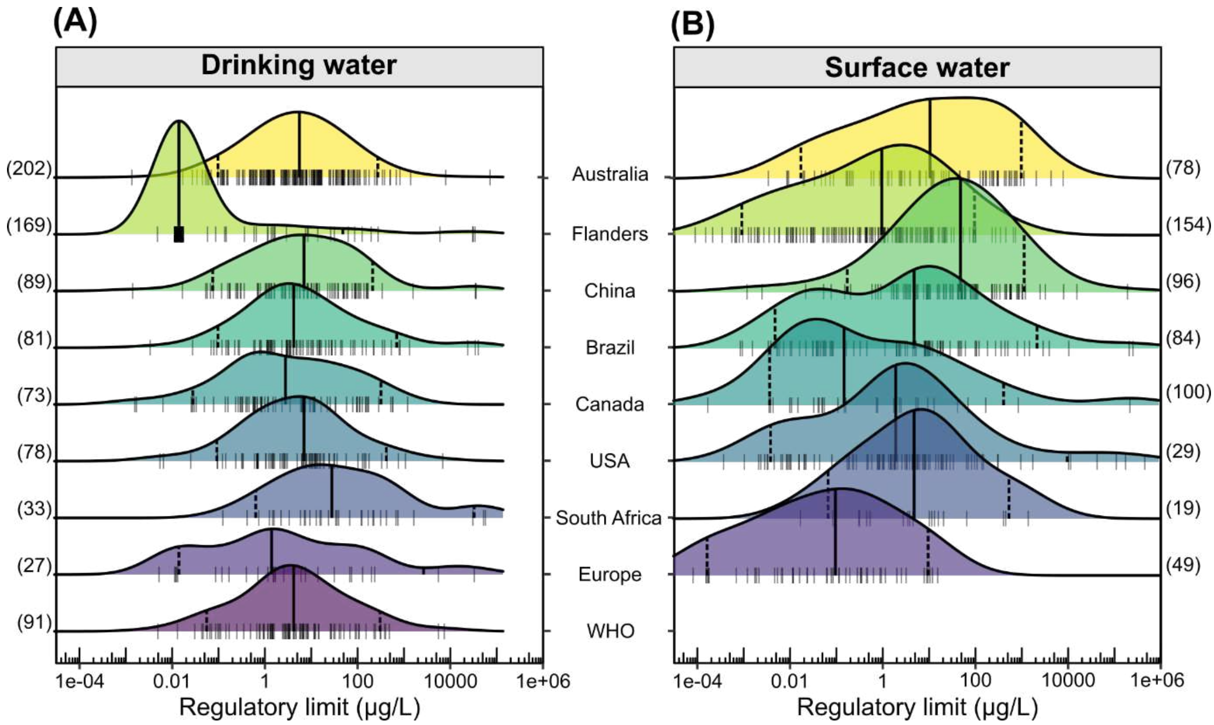
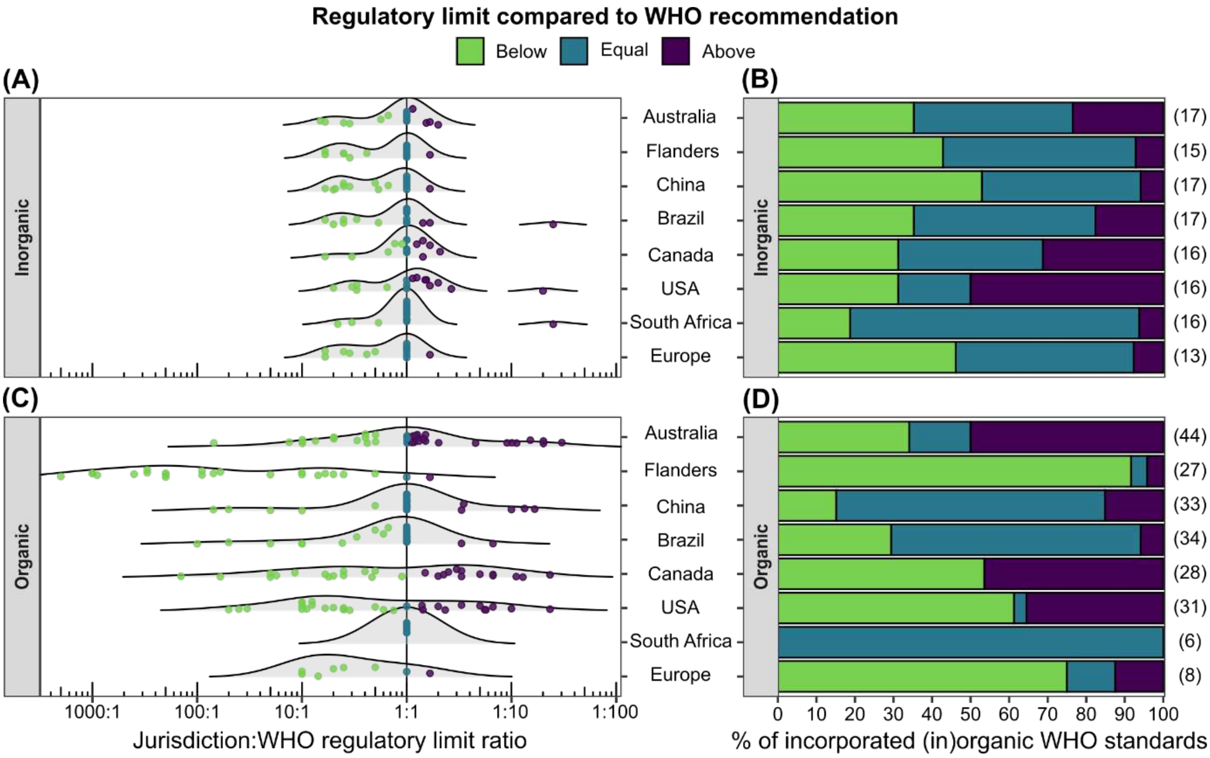
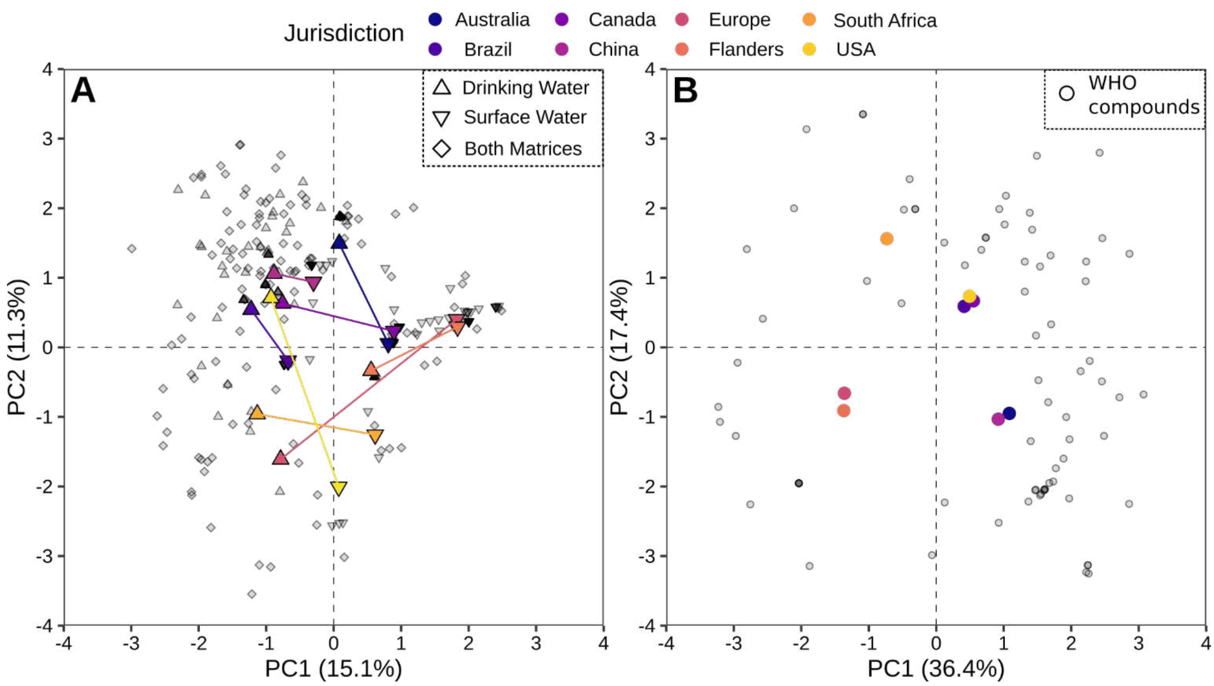


Figure 4. Distribution of regulatory limits of all considered jurisdictions for drinking water **(A)** and surface water **(B)** regulations. Each vertical tick at the base of each distribution indicates a datapoint (regulatory limit) taken up in the density curve. The vertical black line represents the median regulatory limit within a jurisdiction. The dashed lines left and right of the distribution indicate the 5% and 95% quantile respectively. The number of standards in the respective legislation's list is in between brackets.



<u>Figure 5.</u> Ratio between regulatory norms of the respective jurisdiction and the WHO's recommendations for inorganic **(A)** and organic **(C)** standards. The percentage of WHO-recommended incorporated standards below, equal or above the WHO recommended limit is given for inorganics **(B)** and organics **(D)**.



<u>Figure 6.</u> **(A)** Multidimensional scaling for the complete dataset of drinking water and surface water quality standards, with exception of WHO recommendations. The lines between the upwards and downwards triangles denote the Euclidean representation of the Bray-Curtis distance between a jurisdiction's drinking and surface water lists in the first and second principal coordinates space. **(B)** Multidimensional scaling for the WHO recommended standards included in the drinking water quality lists of the jurisdictions investigated.

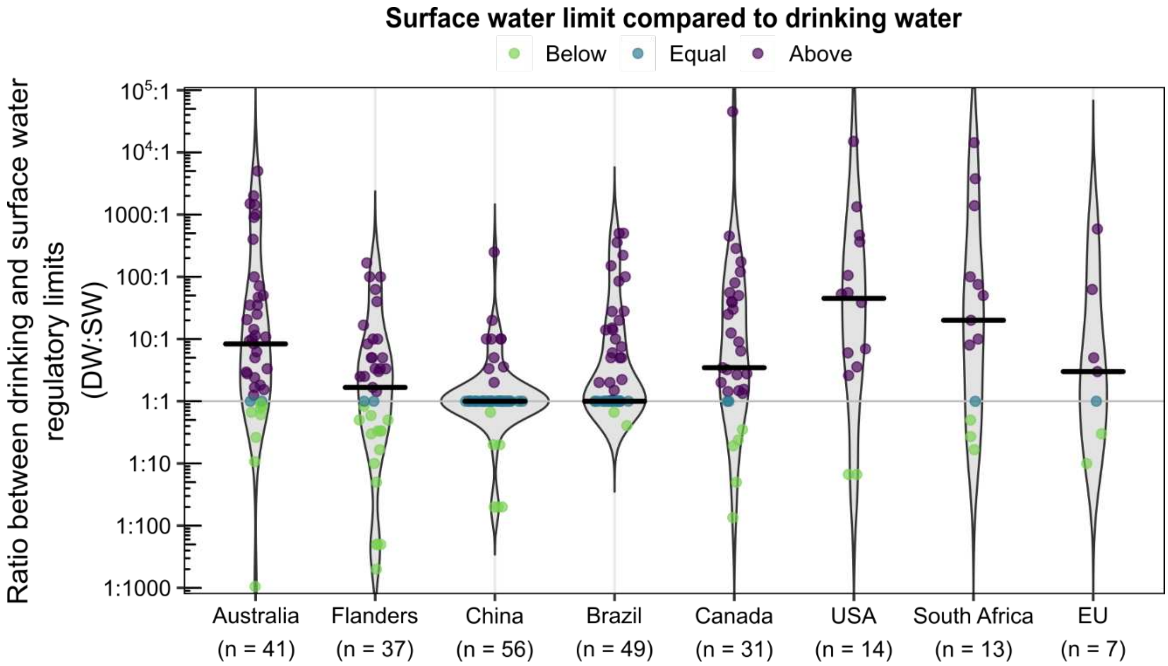


Figure 7. Ratio between drinking water regulatory limits and surface water regulatory limits for standards present in both lists in a respective jurisdiction. A ratio greater than one denotes a higher limit (i.e. laxer norm) for drinking water. A ratio smaller than one means a lower limit (i.e. more stringent norm) for drinking water.