Water Research 309 (2027) 127216

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Chemical log removal credit assessment for a reverse osmosis barrier using a long term operational dataset

Marzieh Niksefat [a], Kathy Northcott [b], Toyah Foote [c], Peter J Scales [d,*]

[a] Rheological Consulting Services, Department of Chemical Engineering, University of Melbourne, VIC, 3010, Australia
[b] Veolia Australia and New Zealand, Level 5, 324 St Kilda Road, Southbank, VIC, 3006, Australia
[c] Seqwater, 117 Brisbane Street, Ipswich, QLD, 4305, Australia
[d] Department of Chemical Engineering, University of Melbourne, VIC, 3010, Australia

HIGHLIGHTS

- A LRV chemical removal credit framework for reverse osmosis (RO) has been comprehensively tested.
- Analysis of 4.5 years of full-scale plant data shows the proposed indicative LRV credit system for RO is robust.
- The chemical classification approach was tested in detail for 87 chemicals of the >500 chemicals in the dataset.

ARTICLE INFO

Keywords:
Reverse osmosis
Chemical groupings
Log reduction value (LRV)
Chemicals of concern
water recycling

ABSTRACT

Broadscale assignment of removal credits (as log reduction values (LRVs)) for chemical contaminants in reverse osmosis (RO) systems is not as well established as equivalent LRVs for pathogens in water reuse operations. Existing studies on contaminant removal are predominantly based on laboratory and pilot-scale experiments with limited replication, and where monitoring has been undertaken at full scale, the availability of datasets is limited. This study presents an analysis of a comprehensive dataset.

A further issue is that it is preferable to assign removal credits (typically used for regulatory purposes) to classes of chemicals rather than to individual chemical species, whereby a single chemical is chosen as a conservative representative of a sub-class of chemicals that share similar mechanistic removal traits. The combination of assigning removal credits using a lower-tail LRV estimate calculated from the LRV distribution measured under the most challenging operational conditions and using a chemical that is conservative in removal performance for a given sub-class of chemicals (in terms of removal mechanism) offers a highly conservative but defensible approach that makes it consistent with the allocation of removal credits for pathogens.

A long-term operational dataset from a full-scale advanced water treatment plant producing purified recycled water was used to establish the criteria that allowed statistical evaluation of representative operational performance indicators of an RO water treatment barrier. Parameters including sulfate, sodium, and electrical conductivity were used in this instance, inclusive of establishing the minimum replicate data requirements for LRV estimation. The dataset (>4 years of data, >500,000 readings, >100 chemical species) was then evaluated across a range of chemical sub-classes using indicative LRV credits to validate the approach and assess the robustness of the conservative credit assignment. Application of conservative sub-class-level assumptions resulted in no cases where the presence of a molecule or ion in the RO permeate was not predicted (i.e. no false negative results). This provides confirmation that these groupings allow protective coverage of observed chemical removal behaviour through an RO barrier.

These findings demonstrate that full-scale operational monitoring data can support statistically robust, risk-based assessment of RO chemical removal. The extensive operational dataset supports the derived chemical sub-classes and the process for identification of performance surrogates. There were no observed gaps in the RO chemical sub-class groupings, demonstrating that chemical sub-classes (using performance surrogates) can be used to assess individual chemical removal LRVs without undertaking challenge testing for individual chemicals.

* Corresponding author.
E-mail address: peterjs@unimelb.edu.au (P.J. Scales).

https://doi.org/10.1016/j.watres.2026.127216
Received 19 June 2026; Received in revised form 16 October 2026; Accepted 24 October 2026
Available online 26 October 2026
0043-1354/Crown Copyright © 2026 Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Increasing pressure on global water resources has accelerated the implementation of advanced water recycling and potable reuse systems as critical components of long-term water security strategies (Florides et al., 2024; WHO, 2017). These systems rely on multiple treatment barriers to ensure reliable removal of both pathogens and chemical contaminants. Reverse osmosis (RO) is widely applied as an important barrier due to its high rejection of dissolved salts and trace organic compounds (Alhussaini et al., 2024; Drewes and Khan, 2015; Jeffrey et al., 2022). As the number of reuse schemes expand, regulatory frameworks increasingly require demonstration of chemical as well as pathogen reduction in the treated water. Although the assignment of log removal values (LRVs) or credits to individual treatment barriers is well established for pathogens, credit allocation for chemicals of concern (CoC) is not as advanced (Arden et al., 2024; Jeffrey et al., 2022; Pang et al., 2023; WHO, 2017).

For chemical contaminants, the principal challenge is not the ability of RO to achieve high removal, but the absence of a statistically defensible validation framework for assigning reduction credits under full-scale operating conditions (Pang et al., 2023; Scales et al., 2019). In the case of pathogens, there are established frameworks that rely on representative surrogates of membrane performance, conservative reduction targets, and operational monitoring (WHO, 2017). There is no equivalent widely accepted approach that exists for the allocation of chemical removal credits (Pang et al., 2023; Scales et al., 2019). This gap limits the translation of mechanistic understanding of RO performance in a consistent approach to demonstrate management of chemical risks for advanced water treatment schemes. Addressing this gap requires full-scale operational datasets that can quantify variability, define conservative lower-tail LRVs (not just mean LRVs), and test whether sub-class-level credit assumptions remain protective under real operating conditions.

Chemical rejection by polyamide RO membranes (common to most operational RO systems and the type considered herein) is governed primarily by size exclusion and electrostatic interactions, with chemical removal strongly influenced by molecular size, charge, and hydrophobicity (Alturki et al., 2010; Bellona et al., 2004; Heo et al., 2020; Khoo et al., 2022; Ma et al., 2024; Shin et al., 2022; Wu et al., 2022). These relationships provide a mechanistic basis for grouping chemicals into sub-classes with similar removal behaviour, enabling interpretation of RO performance beyond compound-specific analysis (Pang et al., 2023; Scales et al., 2019). Such an approach offers a tractable pathway for the assignment of log reduction credits through an RO barrier at the sub-class level, rather than for individual chemical compounds. Given the large number and diversity of contaminants present in recycled water systems (Drewes et al., 2018; Finnerty et al., 2023; Ong, 2018), moving to a sub-class description has the potential to provide a feasible approach to monitoring and demonstrating reduction of chemicals across RO. An approach where all chemicals, not just chemicals of concern, are considered. Additionally, the need for ever expanding analysis of individual chemicals to regularly demonstrate the removal of many known and emerging chemicals of concern is reduced. However, practical implementation remains limited by differing approaches to sub-class definition, selection of representative chemical indicators, and a standardised approach to demonstration of conservative coverage of chemical occurrence.

Existing literature on RO chemical removal is dominated by laboratory and pilot-scale studies, typically based on a single or a small number of observations (Albergamo et al., 2019; Fujioka et al., 2013; Ma et al., 2024; Radjenović et al., 2008). While these studies provide important mechanistic insights, they lack the benefit of demonstration at scale, allowing for characterisation of multiple chemical sub-classes across an extended period of operation. Replicate measurements and systematic evaluation of uncertainty are rarely reported, and sub-class level analyses are uncommon. As a result, current literature is insufficient to support statistically defensible chemical credit assignment (at the sub-class level) to support validation requirements as part of regulatory approval processes.

In parallel, the role of surrogate or indicator parameters in chemical validation remains poorly defined (Pang et al., 2022). Parameters such as conductivity, sulfate, selected trace organics or additives such as WT-30 have been proposed as operational indicators of RO performance, but their relationship to specific mechanistic chemical classes remains poorly established (Jeffrey et al., 2022). They do not universally provide a conservative representation of contaminant removal (Drewes et al., 2018; Pang et al., 2023), rather an indication that the barrier is operating above a minimum threshold. This contrasts with the allocation of pathogen reduction credits, where surrogate frameworks are well established and underpin operationally simple and robust performance verification strategies. These strategies are not typically published in the peer-reviewed literature but sit within state or national regulatory frameworks or industry-wide representative bodies (Sarkis et al., 2013; WaterVal, 2017). Unlike the case for pathogens, no universally accepted surrogate framework exists for the selection of chemical parameters chosen for validation in RO systems such as target parameters that are meant to be removed and/or inactivated by each barrier including real hazards, such as specific chemicals with public health standards, as well as indicators or surrogates (Drewes et al., 2018).

A risk-based approach provides a pathway to address these limitations by linking source water composition, mechanistic classification, and log reduction credit allocation alongside performance verification requirements. One such approach is the WaterVal framework (WaterVal, 2017). This framework has been used to standardise the process and approach to validating water treatment process technologies incorporating identification of influencing operational factors that affect the efficacy of the process to reduce the target contaminant. The efficacy of the treatment process is demonstrated by the allocation of pathogen log reduction credits and has recently been adapted to chemical removal, specifically for the RO barrier in previous work by Scales et al. (2019) (Scales et al., 2019) and Pang et al. (2022, 2023) (Pang et al., 2023, 2022). In such an approach, chemicals are grouped into sub-classes based on removal-relevant properties, and validation effort is prioritised according to potential contribution to treated water risk (Drewes et al., 2018; Pang et al., 2022). However, implementation of such approaches requires statistically robust full-scale data or sufficiently replicated pilot/laboratory data that can be used to support defensible estimation of log reduction performance and to demonstrate that sub-class level assumptions provide conservative coverage of observed chemical removal performance behaviour. The present study evaluated this classification framework using a long-term, full-scale operational dataset.

A framework for allocation of removal (in terms of LRV) is well established for pathogens, where health-based targets derived from quantitative microbial risk assessment (QMRA) are implemented through Water Safety Plan frameworks combining an initial validation with critical control point operational monitoring (Jeffrey et al., 2022; NRMMC et al., 2009; Sarkis et al., 2013; Victorian Department of Health, 2013; WHO, 2017). Within this structure, the validation provides evidence of conservative removal performance under defined conditions, while operational monitoring of integrity and performance surrogates confirms that treatment processes continue to operate within the established limits, defined at the time of validation (Sarkis et al., 2013; WHO, 2017).

The diversity and variability of chemical contaminants in recycled water systems preclude validation at the level of individual compounds (Finnerty et al., 2023; Ong, 2018; WHO, 2017). A tractable approach, therefore, requires abstraction of chemical behaviour into representative groups that reflect treatment-relevant removal mechanisms. End-use classifications (e.g. pharmaceuticals or pesticides) support prioritisation but do not predict treatment performance. Mechanistic sub-classes directly reflect the governing processes of RO rejection and

allow extrapolation of removal behaviour within defined chemical domains (Dickenson et al., 2009; Drewes et al., 2018). This enables validation of removal credits to be undertaken using representative compounds that provide conservative estimates of sub-class-level removal, rather than requiring compound-specific assessment.

As already noted, rejection of dissolved chemicals by polyamide RO membranes is governed by size exclusion and electrostatic interactions, resulting in systematic variation in removal as a function of molecular size, hydrophobicity and charge (Alturki et al., 2010; Bellona et al., 2004; Heo et al., 2020; Khoo et al., 2022; Ma et al., 2024). These relationships provide the basis for grouping chemicals into mechanistic sub-classes with comparable removal behaviour, rather than interpreting performance at the level of individual compounds. The biggest issue for RO membranes is primarily associated with small neutral molecules. These represent a limiting case for defining a representative conservative surrogate for that sub-class and show a high sensitivity to operating conditions (Antony et al., 2011; Ebrahimzadeh et al., 2021; Wu et al., 2022).

In terms of descriptors for sub-class designation, molecular width and molar volume (MV (cm^3/mol) are observed to provide a more reliable predictor of rejection than conventional metrics such as molecular weight cut-off (MWCO), particularly in distinguishing consistently rejected compounds from those with variable removal (Albergamo et al., 2019; Doederer et al., 2014; Pang et al., 2023; Scales et al., 2019). Therefore, based on mechanistic relationships, grouping chemicals into sub-classes defined by size (as defined by molar volume) and charge, as previously proposed by Scales et al. (2019) and further developed by Pang et al. (2023), is expected to result in removal performance that is broadly consistent within any one sub-class (Pang et al., 2023, 2022; Scales et al., 2019).

In full-scale systems, high rejection combined with pre-treatment often results in permeate concentrations below detection limits, limiting reliable LRV calculation (Dickenson et al., 2011; Ishida et al., 2021; Pang et al., 2023, 2022). Spiking experiments are therefore required (Albergamo et al., 2019; Fujioka et al., 2013; Pype et al., 2015). Choice of spiking or challenge test chemicals does not always include compounds that are conservative to a chemical sub-class, limiting the applicability of validation results to the allocation of removal credits (Allinson et al., 2018; Doederer et al., 2014; Heo et al., 2020; Radjenović et al., 2008).

This study evaluates the previously developed chemical sub-class and LRV allocation approach using long-term operational data from a full-scale advanced water treatment plant producing purified recycled water with RO as one of the treatment barriers. The integrated aims were to:

(i) quantify the statistical behaviour and convergence of representative RO performance indicators
(ii) evaluate minimum data requirements for stable and defensible LRV estimation of chemical removal
(iii) evaluate sub-class-based chemical occurrence to identify chemical domains most likely to challenge RO removal performance
(iv) use the dataset as a test case to assess the efficacy of the previously developed risk-based chemical removal credit allocation framework at the sub-class level using representative RO performance indicators.

2. Materials and methods

This study utilised a full-scale operational dataset from the Luggage Point Advanced Water Treatment Plant (LPAWTP), part of the Western Corridor Recycled Water Scheme (WCRWS), in Queensland, Australia. The scheme is managed by Seqwater and operated by Veolia Australia and New Zealand and has a catchment of greater than 0.75 million persons. In this multi-barrier system, RO is implemented as one of a series of barriers for the removal of particulates, dissolved inorganic species and organic micropollutants.

The RO system utilised Toray TML20D membranes and operates at approximately 85% recovery, with a permeate flux of 17–20 L m^{-2} h^{-1}, consistent with full-scale operating conditions reported for other advanced reuse applications (Pang et al., 2023). On this basis, the dataset and RO membranes utilised provide a representative basis for evaluating removal performance under realistic operational conditions.

Indicators such as electrical conductivity (EC), sulfate, and sodium were routinely monitored and provided consistent measures of membrane performance. The dataset spanned approximately 4.5 years of continuous plant operation and included routine monitoring of paired feed and permeate streams. A total of 516 parameters (individual chemicals and parameters such as total nitrate) were monitored, with the database comprising over 500,000 measurements.

Parameters with <10% detection frequency in the feed were excluded from further data analysis, and RO feed concentrations below the method detection limit (MDL) were removed to ensure that LRV calculations were based on measurable concentrations.

No observations were excluded as statistical outliers, as the dataset was intended to represent the range of full-scale operating conditions.

For RO permeate concentrations below the MDL, a value of 0.5 × MDL was applied to enable a minimum LRV estimation. This substitution provides a conservative estimate of low concentrations without systematically overestimating removal performance.

LRVs were calculated from paired feed and permeate concentrations. Statistical measures, including mean LRV, standard deviation (SD), and standard error of the mean (SEM), were evaluated as a function of cumulative sample size. Observations were added in chronological order in the cumulative analysis. These cumulative sequences were used as descriptive assessments of the operational record, as their apparent convergence may be influenced by the order in which observations are included.

SEM was calculated as:

$$SEM = SD / \sqrt{n}$$ where n is the number of samples.

Relative SEM was calculated as:

$$\text{Relative SEM (\%)} = (SEM / \text{Mean LRV}) \times 100$$

The Detection Ratio (DR), defined as the ratio of feed concentration to MDL (Dickenson et al., 2011), was used to assess the validity of LRV calculations. For DR ≥ 5, LRV was calculated using the 0.5 × MDL substitution, while for DR < 5, only a minimum LRV was reported due to insufficient analytical sensitivity. The DR = 5 threshold was used to ensure sufficient analytical range between the feed concentration and detection limit for calculation of a meaningful LRV, consistent with the approach of Dickenson et al. (2011).

Data analysis including calculation of the SEM was performed, and a relative SEM criterion of <5% was adopted as a pragmatic stability threshold to identify:

(i) when additional observations had diminishing influence on the estimated mean LRV and
(ii) the minimum number of observations required for a reliable estimate of the mean LRV.

Using relative SEM<5% is not proposed as a universal threshold, but as an operational benchmark for assessing data sufficiency in long-term monitoring datasets. This criterion was not used independently to establish stability of the conservative lower-tail credit.

Following screening, the dataset used for analysis comprised over 100,000 observations. Parameters with low detection frequency were excluded to ensure sufficient data density for statistically reliable analysis and to remove compounds that did not contribute meaningfully to the risk assessment. Sulfate and sodium were selected as representative indicators of both the performance of the membrane and their respective chemical sub-classes, while EC was evaluated as an operational integrity indicator. It is not proposed that these parameters are used as quantitative surrogates for the removal of the uncharged organic chemical sub-

classes or of molecules with a lower molecular (hydrated) size.

Conservative LRVs were calculated as the mean observed LRV minus 1.96 standard deviations:

$$\text{Conservative LRV} = \text{mean LRV} - 1.96 \times \text{SD}$$

This metric is representative of a parametric lower-tail estimate of the observed LRV distribution and is not an empirical percentile or a confidence interval for the mean. It represents a conservative estimate of lower-tail operational performance.

This framework separates the precision of the mean LRV from stability of the assigned conservative credit. Mean-LRV precision was evaluated using relative SEM, while an indication of validation-relevant credits was derived from the parametric lower-tail estimate of observed operational performance. Removal credits that could be used for regulatory purposes would require separate testing under the most challenging operational conditions.

To compare chronological analysis of the dataset as distinct from random assessment, repeated random subsampling was conducted for sulfate, sodium and EC. Whilst neither procedure is appropriate for estimation of a credit that could be used for regulatory purposes (since the data is produced under a range of conditions as distinct from the most challenging operational conditions), it is useful to obtain an indicative credit for framework evaluation purposes. For each indicator, subsets containing 10 to 250 observations, in increments of 10, were randomly selected without replacement from the complete operational dataset. A total of 5000 subsets were generated at each sample size. For each subset, the conservative LRV was calculated as the mean LRV minus 1.96 standard deviations. The corresponding indicative credit was then obtained by rounding the conservative LRV down to the nearest 0.1 LRV.

Each subset-derived credit was compared with the credit calculated from the complete dataset and classified as higher, the same or lower. The proportion of subsets that produced a higher credit was used to assess the risk of over-crediting (i.e. if conservative test data was not available and the entire dataset was used) when fewer observations were available. The sample size at which the proportion of subsets producing a higher credit fell to 5% or less was reported as a dataset-specific sensitivity measure and was not interpreted as a universal minimum sampling requirement. This 5% threshold was used as a conservative criterion for this study and is not intended to represent an established regulatory requirement. The analysis was performed in Python (version 3.14.6) using NumPy (version 2.5.1) with a fixed random seed for reproducibility.

Chemicals were subsequently classified into RO removal sub-classes based on molecular size and charge, consistent with established mechanistic predictors of RO rejection behaviour and previous work from the same group. The classification framework and sub-class definitions are provided in Table 1.

Table 1
Chemical sub-classes used in this study to classify removal performance through an RO membrane based on charge and molecular size, expressed as molar volume (MV; cm^3/mol) (Pang et al., 2023).

Class	Sub-Class	
	Charge (+, ++, o)	**MV (cm^3/mol)**
Inorganic	+	-
	++	-
	o	-
Organic	+/ ++	MV≤120
	+/ ++	120<MV≤200
	+/++	200<MV≤400
	+/++	MV>400
	o	MV≤120
	o	120<MV≤200
	o	200<MV≤400
	o	MV>400

+: monovalent; ++: multivalent; o: neutral; MV = molar volume.

3. Results and discussion

3.1. Conventional RO performance and integrity indicators

The analysis focuses on three elements: (i) the statistical analysis of representative RO integrity and performance indicators (sulfate, sodium, and EC), (ii) assessment of whether the available data are sufficient for precise estimation of the mean LRV and an indicative conservative credit, and (iii) sub-class-based profiling of source-water chemicals and comparison of actual RO performance with that predicted using indicative conservative LRVs. The purpose here is not simply to quantify RO removal, but to test whether routine full-scale operational data can support statistically defensible chemical removal credits that can later be applied at the sub-class level. Together, these results evaluate the practical application of the previously developed risk-based RO chemical classification and credit allocation framework using a long-term, full-scale operational dataset.

Defensible assignment of removal credits requires not only estimation of the mean LRV but also explicit quantification of variability and lower-tail performance, as operationally, credits must reflect conservative system behaviour (Helsel, 2010; NRMMC et al., 2009). However, most studies of RO chemical removal are based on limited observations and do not report variability or uncertainty, constraining their application for assignment of credits and regulatory interpretation (Albergamo et al., 2019; Fujioka et al., 2013; Kimura et al., 2004; Radjenović et al., 2008).

Mean LRV behaviour (Figs. 1a, 2a and 3a) was used to assess consistency of removal performance. For all indicators, the cumulative mean changed progressively less as further observations were added, indicating that estimates based on small sample counts were more sensitive to the particular observations included. Sulfate exhibits higher rejection (mean LRV~2.7–2.8) than sodium and EC (mean LRV~1.6–1.7), consistent with established RO removal mechanisms in which multivalent ions are more strongly rejected due to combined size exclusion and electrostatic interactions (Antony et al., 2011; Bellona et al., 2004; Pang et al., 2023; Scales et al., 2019).

Relative SEM (Figs. 1b–3b) provides a direct measure of uncertainty

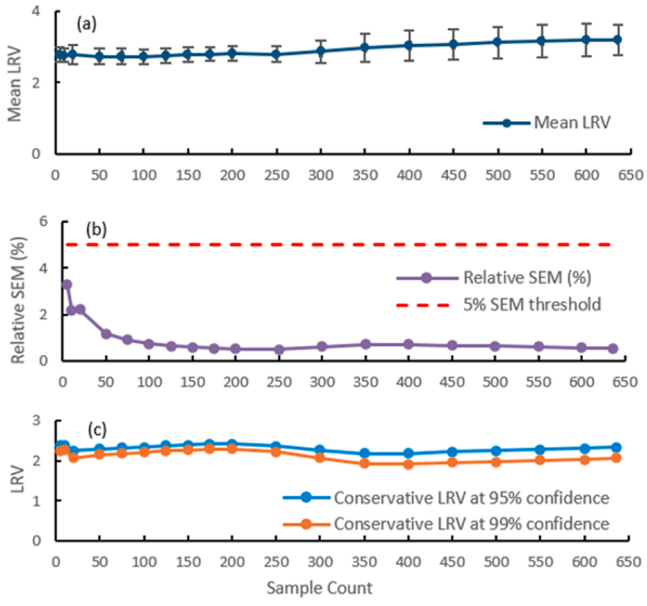


Fig. 1. Statistical evaluation of sulfate LRV estimates with increasing sample size: (a) mean LRV with ±1 SD error bars; (b) relative SEM with the 5% stability threshold indicated; (c) conservative LRV estimates based on the mean LRV − 1.96 SD and mean LRV − 2.58 SD, respectively.

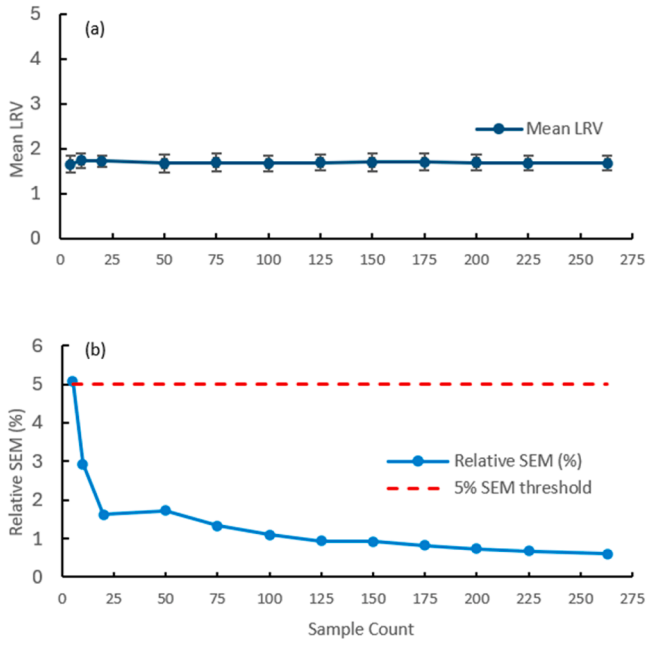


Fig. 2. Statistical evaluation of sodium LRV estimates with increasing sample size: (a) mean LRV with ±1 SD error bars; (b) relative SEM showing stabilisation of statistical uncertainty with increasing monitoring points.

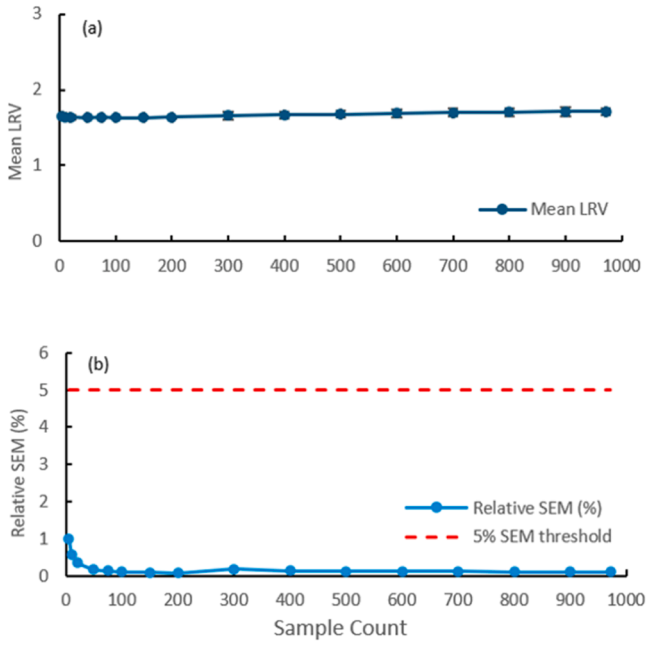


Fig. 3. Statistical behaviour of EC LRV estimates with increasing sample size: (a) mean LRV; (b) relative SEM demonstrating high statistical stability even at low sample counts.

in the estimated mean and is used herein as the primary criterion for precision of the mean LRV. SEM decreases with increasing sample size, reflecting improved confidence in the estimate of the central tendency (Helsel, 2010; Montgomery and Runger, 2019). For sulfate and sodium, relative SEM decreases rapidly and falls below 5% at approximately 15–25 observations. The sample size associated with this criterion depends on the observed LRV variability, which is influenced by feed concentration variability and the frequency of permeate measurements below the detection limit. Beyond this range, further reductions are marginal, indicating diminishing returns from additional observations for estimation of the mean LRV. EC exhibits consistently low relative SEM (<1%) due to low variability, supporting its role as a stable operational indicator of RO membrane integrity, but not as a direct surrogate for broader chemical sub-class removal behaviour, as EC reflects the removal of both inorganic and charged organic species. However, in all cases EC removal needs to be monitored online in feedwater and combined permeate to allow the real-time detection of any catastrophic breach of integrity. A catastrophic breach of integrity impacts all chemical sub-class removal.

Parametric lower-tail LRV estimates (Fig. 1c) illustrate how the calculated conservative operational performance changes as observations are added, which is important for credit allocation. At low sample counts, the lower-tail estimates calculated as mean LRV − 1.96 SD and mean LRV − 2.58 SD vary substantially, reflecting sensitivity to the limited set of observations included. As the sample size increases, these estimates become less variable in the chronologically ordered cumulative analysis. However, visual convergence of a single cumulative sequence does not independently demonstrate that a proposed sample size will consistently reproduce the same conservative credit, as the apparent result may depend on the ordering of observations. Operational testing for validated credit assignment would typically use challenge test measurements or samples taken over a short space of time under tight operational conditions and in many cases be restricted in number. Credit stability was assessed separately using repeated random subsampling to reduce dependence on the original observation order (Haan, 2002; Helsel, 2010).

The relative SEM of the mean fell below 5% at approximately 20 observations for sulfate and sodium, indicating reasonable precision of the estimated mean LRV (Haan, 2002; Helsel, 2010; Montgomery and Runger, 2019). However, this result does not establish that a different set of 20 observations would produce the same conservative credit.

To estimate the upper limit of the number of measurements/samples required, repeated subsampling (Fig. 4) showed that 20 observations were not sufficient to ensure stable conservative credit assignment for sulfate or sodium across a subset with operational data from a wide range of operational conditions. At this sample size, 34.5% of sulfate subsets and 59.9% of sodium subsets produced credits higher than those derived from the complete datasets, whereas no higher credit was observed for EC. When a 5% upper limit was used for the proportion of subsets producing a higher credit, the corresponding sample sizes were approximately 190 for sulfate, 210 for sodium and 10 for EC. Among the indicators evaluated, sodium required the largest number of observations to reach this level in the present dataset. These findings show that the number of observations required to estimate the mean LRV with acceptable precision is not necessarily sufficient to support stable conservative credit assignment, although it is likely that the number would be closer to 20 if only conservative operational data was utilised. There were no data showing values below the MDL in this analysis. It is further assumed that an approach that resulted in non-detects in the permeate would not be used for a validation exercise for credit assignment. This analysis assessed sensitivity to subset composition but did not account for temporal autocorrelation. The reported sample counts should therefore be interpreted as dataset-specific indicators of credit stability rather than as numbers of independent observations or universal minimum requirements. This contrasts with most published studies of RO chemical removal, which rely on single or limited measurements and do not capture uncertainty (Albergamo et al., 2019; Fujioka et al., 2013; Kimura et al., 2004; Radjenović et al., 2008).

The results are summarised in Table 2, which presents the sample sizes associated with the relative SEM criterion for the mean LRV and the corresponding indicative conservative LRV credits. The difference between the mean and conservative LRV values reflects the influence of operational variability on the credit that can be assigned. In allocation of an actual removal credit, the data would be obtained under conservative operational conditions where the variability would be lower. Under these circumstances, the sample size indicated from the analysis in Fig. 4

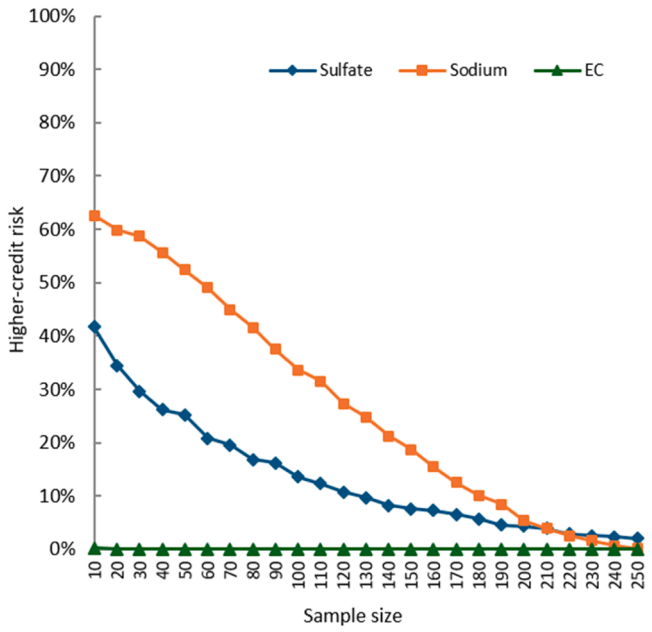


Fig. 4. Probability of assigning a higher conservative LRV credit as a function of sample size. (Credits were calculated as mean LRV – 1.96 SD and rounded down to the nearest 0.1 LRV based on 5000 repeated random subsets for each indicator).

is representative of an upper limit and the sample size to achieve precision in the mean-LRV, a lower limit.

3.2. Evaluation of the chemical sub-class framework

The critical test of the previously developed sub-class framework is whether chemical sub-class indicative LRVs derived from literature can predict all observed permeate detections under full-scale operation, without any unpredicted observation of chemicals in the permeate (i.e. a false negative). Chemical occurrence in RO feedwater was evaluated within the previously proposed mechanistic sub-class framework to determine whether validation can be reduced to a limited number of representative chemical groups. A total of 516 chemicals and parameters were monitored over 4.5 years of operation, of which 87 compounds had a detection frequency of >10%, a threshold applied for further analysis of the compound. These compounds were classified into sub-classes consistent with established predictors of RO removal behaviour (Pang et al., 2023; Scales et al., 2019) as outlined in Table 3.

For each sub-class, the number of chemicals present in the RO feed (Chemicals), those predicted to be detected in the RO permeate (Predicted), and those detected in the RO permeate (Detected) are shown in Table 3. Predicted detections (No. Likely Detected) were estimated using the maximum observed RO feed concentrations and indicative LRV values as given in Table 3. Using the maximum observed RO feed concentration for each chemical across multiple sampling events (often >100) and an indicative LRV values provides a conservative predication of potential permeate occurrence. Analysis of full-scale system occurrence data provides an evaluation of the previously proposed sub-class-level indicative LRV values and their suitability as potential conservative default values for chemicals within these sub-classes (Drewes et al., 2018; Pang et al., 2023). The indicative LRV values were taken from the work of Pang et al., 2023, which was conducted on a full-scale plant based on six sampling events and using the removal performance of a compound that was at the lower end of the MV range for each sub-class.

The occurrence-based comparison was used to assess whether the indicative sub-class LRVs conservatively bounded observed permeate occurrence, rather than to derive new compound-specific LRVs from the operational dataset. Permeate occurrence was strongly non-uniform and concentrated within a restricted subset of sub-classes. Of the 23 observed detections, 13 (57%) were associated with inorganic charged species and 7 (30%) with uncharged organics (MV <200 cm^3/mol), with the remaining detections distributed across the other sub-classes. No detections were observed for compounds with MV ≥200 cm^3/mol, despite conservative prediction assumptions. These patterns are consistent with established RO removal behaviour governed by molecular size and charge (Heo et al., 2020; Pang et al., 2023; Tang et al., 2017), and demonstrate that full-scale permeate occurrence is

Table 2
Indicative conservative LRV credits and sample sizes associated with mean-LRV precision for representative RO performance indicators.

Parameter	Sample size for mean-LRV precision (Relative SEM < 5%)	Indicative conservative LRV (mean LRV – 1.96 SD)
Sulfate	20 samples	2.2
Sodium	20 samples	1.3
EC	≤10 samples	1.6

Table 3
Distribution of detected chemicals and permeate occurrence by RO chemical sub-class. Predicted detections are based on indicative sub-class LRV values and maximum observed RO feed concentrations.

Chemical Sub-class	Chemicals	Predicted	Detected	% Detected	Indicative LRV
Inorganic – Charged – Monovalent	11	7	7	100	1.4
Inorganic – Charged – Multivalent	13	7	6	86	2.2
Inorganic – Uncharged	2	2	2	100	0–0.6
Organic – Charged – MV < 120	2	2	0	0	0.5
Organic – Charged – 120 ≤ MV < 200	6	4	1	25	1.0
Organic – Charged – 200 ≤ MV < 400	1	1	0	0	1.5
Organic – Charged – MV ≥ 400	0	0	0	0	2.0
Organic – Uncharged – MV < 120	7	7	4	57	0
Organic – Uncharged – 120 ≤ MV < 200	13	12	3	25	0.5
Organic – Uncharged – 200 ≤ MV < 400	28	11	0	0	1.0
Organic – Uncharged – MV ≥ 400	4	2	0	0	1.5

Note:
Predicted: Number of chemicals with at least one feed concentration exceeding the threshold defined by the indicative LRV and analytical detection limit (calculated as 10^(Indicative LRV + $\log_{10}$(LOD))).
Detected: Number of chemicals with at least one permeate concentration above the detection limit.
% Detected: Percentage of "at-risk" chemicals observed in the permeate.
Indicative LRVs are sub-class-level values derived from literature.

controlled by a limited number of mechanistic domains.

Comparison of predicted and observed permeate occurrences showed high sensitivity but lower specificity, consistent with the intentionally conservative prediction approach. Inorganic charged species exhibited near-complete agreement, with 7/7 monovalent and 6/7 multivalent predicted detections observed. No false-negative outcomes were observed across the reduced dataset (>100,000 observations), in that there was no instance where a chemical was detected in the RO permeate and not predicted to be there. This is the central finding of the sub-class analysis: conservative sub-class-level LRV assignments derived from previous work, bounded observed permeate occurrence under full-scale operation and satisfied a key requirement for potential conservative default LRV value application (Pang et al., 2023; Scales et al., 2019).

The mechanistic distribution of detections supports a reduced validation structure. No detections were observed for compounds with MV $\geq$200 cm^3/mol, supporting representation of higher molecular weight sub-classes by more conservative lower MV domains. Small neutral compounds (MV <120 cm^3/mol) remain a limiting case because size exclusion and electrostatic repulsion provide limited separation benefit (Bellona et al., 2004; Fujioka et al., 2013; Shin et al., 2022; Wu et al., 2022). This is consistent with previous evidence that rejection of small uncharged compounds can be governed not only by size exclusion, but also by dipolar interactions and hydrogen bonding with the polyamide active layer, leading to poor rejection for specific functional groups such as secondary amides (Fujioka et al., 2020).

Taken together, the combined evidence of permeate occurrence from the dataset, the indicative LRV removal credits at the sub-class level, and the mechanistic constraints, supports the previously identified sub-classes that define the defensible envelope for LRV removal credits. The results provide long-term full-scale evidence supporting the proposition that compound-specific allocation of LRV chemical removal credits is not required as the primary approach. Instead, a reduced set of mechanistically defined sub-classes may provide a defensible basis for assigning LRV removal credits for chemical validation across RO, covering a very large number (millions) of chemicals, consistent with previous arguments that property-based classification provides more relevant information for RO removal than provenance-based chemical groupings (Osorio et al., 2022).

It should be noted that sub-class-level LRV chemical removal credits used herein for prediction are derived from literature and represent indicative values only; formal credit assignment would require validation using targeted challenge testing under defined operating conditions fulfilling the challenge testing requirement of a validation process. The data herein suggests that the minimum number of sampling events should be 20.

The results provide full-scale operational support for implementing RO chemical LRV credit allocation using the previously developed sub-class-based framework. The permeate detections within a limited number of sub-classes, together with the absence of false-negative outcomes, supports the use of a chemical sub-class-level framework as a screening and chemical LRV credit allocation framework, provided that indicative LRVs are confirmed through targeted challenge testing (Pang et al., 2023; Scales et al., 2019).

This provides a clear basis for prioritising validation effort. Sub-classes associated with permeate occurrence or lower rejection potential define the limiting conditions, whereas sub-classes characterised by consistently high rejection may not require independent validation where their behaviour is conservatively represented by lower-MV or lower-rejection domains. This supports application of the previously developed risk-based framework in which effort is prioritised based on the likelihood of permeate occurrence rather than applied uniformly across all detected compounds (Sarkis et al., 2013; WHO, 2017).

Within this framework, validation to allocate removal credits, operational monitoring, and integrity monitoring perform distinct roles. This distinction is important because chemical removal credits cannot be transferred across membrane processes with different controlling mechanisms; membrane distillation is governed mainly by volatility and hydrophobicity, whereas RO is governed mainly by size exclusion, electrostatic effects and solute–membrane interactions (Costa et al., 2023). Validation and credit allocation defines conservative performance through targeted challenge testing, while operational monitoring (e.g. conductivity and key ionic indicators) provides continuous verification that system performance remains within the validated envelope (Pype et al., 2015; Sarkis et al., 2013; WaterVal, 2017). This distinction is consistent with established RO integrity monitoring used to indicate the presence of an integrity breach, such as conductivity, sulfate and TOC, while challenge or direct integrity tests are required to establish or confirm higher removal claims (Kumar et al., 2007). The literature is sometimes confusing in this area where compounds such as sucralose and uranine have been suggested as early warning indicators of membrane deterioration despite that fact that they are not routinely observed in the RO permeate, an important caveat for integrity monitoring (Ray et al., 2024). In short, operational indicators do not directly quantify contaminant removal and therefore cannot replace formal challenge testing to allocate LRV removal credits (WaterVal, 2017).

The analysis of RO chemical monitoring data against the previously developed chemical sub-classification framework is based on a single full-scale system. It reflects system-specific conditions including feedwater composition, membrane characteristics, and analytical detection limits. Therefore, application to other systems requires consideration of site-specific factors and, where necessary, targeted testing.

4. Conclusions

This study demonstrates that full-scale operational RO data can be used to derive statistically defensible performance metrics and to evaluate the previously developed chemical sub-class framework to demonstrate it applicability to deriving chemical LRV credits for assessing RO performance and informing risk. Representative membrane performance (sulfate and sodium) and integrity (EC) indicators were used to demonstrate that (under normal operational conditions) the number of samples that may be required to achieve stability of the conservative credit was significantly greater than 20. Under the 5% upper limit applied in this study, sodium required the largest nominal sample count to reach this level. Because temporal autocorrelation was not explicitly accounted for, these counts should be interpreted as dataset-specific indicators of credit stability rather than numbers of independent observations or universal minimum requirements.

Chemical sub-class-based analysis showed that permeate occurrence is consistent with a mechanistic understanding of the performance of RO membranes and with conservative sub-class-level LRV allocation reported in earlier work. The framework and indicative LRV credit assignment provided high sensitivity, with no false-negative outcomes within the dataset assessed (>4 years of data and >500,000 data points). Accordingly, the findings provide long-term full-scale support for the proposition that validation and removal credit allocation does not require compound-specific assessment as the primary basis but can be implemented using a reduced set of representative sub-classes.

These findings provide long-term full-scale support for application of the previously developed risk-based framework in which effort is prioritised toward sub-classes associated with permeate occurrence or lower rejection potential, while sub-classes exhibiting consistently high rejection are unlikely to require independent testing. Removal credits can therefore be assigned at the chemical sub-class level, supported by targeted validation rather than compound-by-compound assessment. Sub-classes predicted to show poor RO rejection, particularly small neutral compounds, remain limiting cases and may require additional treatment barriers based on system risk, targeted monitoring, or more conservative credit assumptions. Operational monitoring supports ongoing verification that system performance remains within the validated envelope. The extensive operational dataset provides long-term full-scale evaluation of the proposed chemical sub-classes and the

process for identification of performance surrogates. There were no observed gaps in the RO chemical sub-class groupings, providing support for the proposition that chemical sub-classes (using performance surrogates) can be used to validate individual chemical removal LRVs without undertaking challenge testing for individual chemicals.

CRediT authorship contribution statement

Marzieh Niksefat: Writing – original draft, Validation, Formal analysis, Data curation. **Kathy Northcott:** Writing – review & editing, Data curation, Conceptualization. **Toyah Foote:** Writing – review & editing, Validation, Formal analysis, Data curation. **Peter J Scales:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge funding for this project through Project 3505 of Water Research Australia and the advice of Dr Marty Hancock. The sponsor group of companies is also acknowledged, namely: Seqwater, Melbourne Water, SA Water, Hunter Water, and Veolia Australia and New Zealand. The project was made possible through the supply of and training on a large operational dataset on a reverse osmosis treatment barrier from Seqwater.

Data availability

The data that has been used is confidential.

References

Albergamo, V., Blankert, B., Cornelissen, E.R., Hofs, B., Knibbe, W.-J., Van Der Meer, W., De Voogt, P., 2019. Removal of polar organic micropollutants by pilot-scale reverse osmosis drinking water treatment. Water. Res. 148, 535–545.

Alhussaini, M.A., Souza-Chaves, B.M., Felix, V., Achilli, A., 2024. Comparative analysis of reverse osmosis and nanofiltration for the removal of dissolved contaminants in water reuse applications. Desalination. 586, 117822.

Allinson, M., Kadokami, K., Shiraishi, S., Nakajima, D., Zhang, J., Knight, A., Gray, S., Scales, P., Allinson, G., 2018. Wastewater recycling in Antarctica: performance assessment of an advanced water treatment plant in removing trace organic chemicals. J. Environ. Manage 224, 122–129.

Alturki, A.A., Tadkaew, N., McDonald, J.A., Khan, S.J., Price, W.E., Nghiem, L.D., 2010. Combining MBR and NF/RO membrane filtration for the removal of trace organics in indirect potable water reuse applications. J. Memb. Sci. 365 (1–2), 206–215.

Antony, A., Bleach, R., Mattiello, A., Leslie, G., 2011. Organic nanoparticles for membrane integrity monitoring – towards clean water production. In: Proceedings of Chemeca 2011, Engineers Australia, ACT, Australia, pp. 1362–1370.

Arden, S., McGaughy, K., Phillips, J., Hills, L., Chiang, E., Dumler, S., Jahne, M., Garland, J., 2024. A unit process log reduction database for water reuse practitioners. Water. Res. X. 23, 100226.

Bellona, C., Drewes, J.E., Xu, P., Amy, G., 2004. Factors affecting the rejection of organic solutes during NF/RO treatment—A literature review. Water. Res. 38 (12), 2795–2809.

Costa, F.C., Dos Santos, C.R., Amaral, M.C., 2023. Trace organic contaminants removal by membrane distillation: a review on mechanisms, performance, applications, and challenges. Chem. Eng. J. 464, 142461.

Dickenson, E.R., Drewes, J.R.E., Sedlak, D.L., Wert, E.C., Snyder, S.A., 2009. Applying surrogates and indicators to assess removal efficiency of trace organic chemicals during chemical oxidation of wastewaters. Environ. Sci. Technol. 43 (16), 6242–6247.

Dickenson, E.R., Snyder, S.A., Sedlak, D.L., Drewes, J.E., 2011. Indicator compounds for assessment of wastewater effluent contributions to flow and water quality. Water. Res. 45 (3), 1199–1212.

Doederer, K., Farré, M.J., Pidou, M., Weinberg, H.S., Gernjak, W., 2014. Rejection of disinfection by-products by RO and NF membranes: influence of solute properties and operational parameters. J Memb Sci 467, 195–205.

Drewes, J.E., Anderson, P., Denslow, N., Jakubowski, W., Olivieri, A., Schlenk, D. and Snyder, S. 2018. Monitoring strategies for constituents of emerging concern (CECs) in recycled water. SCCWRP Technical Report 1032, California State Water Resources Control Board, California, USA, p. 155.

Drewes, J., Khan, S., 2015. Contemporary design, operation, and monitoring of potable reuse systems. J. Water Reuse Desalin. 5 (1), 1–7.

Ebrahimzadeh, S., Wols, B., Azzellino, A., Martijn, B.J., Van Der Hoek, J.P., 2021. Quantification and modelling of organic micropollutant removal by reverse osmosis (RO) drinking water treatment. J. Water. Process. Eng. 42, 102164.

Finnerty, C.T., Childress, A.E., Hardy, K.M., Hoek, E.M., Mauter, M.S., Plumlee, M.H., Rose, J.B., Sobsey, M.D., Westerhoff, P., Alvarez, P.J., 2023. The future of municipal wastewater reuse concentrate management: drivers, challenges, and opportunities. Environ. Sci. Technol. 58 (1), 3–16.

Florides, F., Giannakoudi, M., Ioannou, G., Lazaridou, D., Lamprinidou, E., Loukoutos, N., Spyridou, M., Tosounidis, E., Xanthopoulou, M., Katsoyiannis, I.A., 2024. Water reuse: a comprehensive review. Environments. 11 (4), 81.

Fujioka, T., Kodamatani, H., Yujue, W., Yu, K.D., Wanjaya, E.R., Yuan, H., Fang, M., Snyder, S.A., 2020. Assessing the passage of small pesticides through reverse osmosis membranes. J Memb Sci 595, 117577.

Fujioka, T., Oshima, N., Suzuki, R., Khan, S.J., Roux, A., Poussade, Y., Drewes, J.E., Nghiem, L.D., 2013. Rejection of small and uncharged chemicals of emerging concern by reverse osmosis membranes: the role of free volume space within the active skin layer. Sep. Purif. Technol. 116, 426–432.

Haan, C., 2002. Statistical Methods in Hydrology, 2nd edn. Iowa State Press, Ames, IA, USA.

Helsel, D., 2010. Much ado about next to nothing: incorporating nondetects in science. Annals of Occupational Hygiene 54 (3), 257–262.

Heo, J., Kim, S., Her, N., Park, C.M., Yu, M., Yoon, Y., 2020. Removal of contaminants of emerging concern by FO, RO, and UF membranes in water and wastewater. Contam. Emerg. Concern Water Wastewater 139–176.

Ishida, K.P., Luna, R.F., Richardot, W.H., Lopez-Galvez, N., Plumlee, M.H., Dodder, N.G., Hoh, E., 2021. Nontargeted analysis of trace organic constituents in reverse osmosis and UV-AOP product waters of a potable reuse facility. ACS. ES. T. Water. 2 (1), 96–105.

Jeffrey, P., Yang, Z., Judd, S.J., 2022. The status of potable water reuse implementation. Water. Res. 214, 118198.

Khoo, Y.S., Goh, P.S., Lau, W.J., Ismail, A.F., Abdullah, M.S., Ghazali, N.H.M., Yahaya, N. K.E., Hashim, N., Othman, A.R., Mohammed, A., 2022. Removal of emerging organic micropollutants via modified-reverse osmosis/nanofiltration membranes: a review. Chemosphere 305, 135151.

Kimura, K., Toshima, S., Amy, G., Watanabe, Y., 2004. Rejection of neutral endocrine disrupting compounds (EDCs) and pharmaceutical active compounds (PhACs) by RO membranes. J. Memb. Sci. 245 (1–2), 71.

Kumar, M., Adham, S., DeCarolis, J., 2007. Reverse osmosis integrity monitoring. Desalination. 214 (1–3), 138–149.

Ma, Q., Zhang, J., Zhu, G., Ahuja, N., Khusid, B., Zhang, W., 2024. Effects of surfactants, ion valency and solution temperature on PFAS rejection in commercial reverse osmosis (RO) and nanofiltration (NF) processes. J. Water. Process. Eng. 66, 106039.

Montgomery, D.C., Runger, G.C., 2019. Applied Statistics and Probability For Engineers. John Wiley & Sons, Hoboken, NJ, USA.

NRMMC, EPHC and AHMC 2009. Australian Guidelines for Water Recycling. National Resource Management Ministerial Council, Environment Protection and Heritage Council and Australian Health Ministers Conference, Canberra, Australia.

Ong, C.N., 2018. Water reuse, emerging contaminants and public health: state-of-the-art analysis. In: Tortajada, C., Ong, C.N. (Eds.), Water Reuse Policies For Potable Use, Routledge Special Issues on Water Policy and Governance. Routledge, Oxfordshire, UK.

Osorio, S.C., Biesheuvel, P., Spruijt, E., Dykstra, J., Van Der Wal, A., 2022. Modeling micropollutant removal by nanofiltration and reverse osmosis membranes: considerations and challenges. Water. Res. 225, 119130.

Pang, H., Allinson, M., Northcott, K., Schultz, A., Scales, P.J., 2023. Demonstrating removal credits for chemicals of concern in recycled water through a reverse osmosis barrier – a predictive framework. Water. Res. 244, 120247.

Pang, H., Allinson, M., Schultz, A., Cooper, C., Northcott, K.A., Scales, P.J., 2022. A novel adaptation of WaterVal protocols to address chemical risks in water recycling. In: Proceedings of OzWater 22. Brisbane, Australia.

Pype, M.-L., Alvarez de Eulate, E., Antony, A., Arrigan, D., Busetti, F., Le-Clech, P. and Gernjak, W. 2015. National validation guidelines for water recycling: reverse osmosis membranes, Australian Water Recycling Centre of Excellence, Brisbane, Australia.

Radjenović, J., Petrović, M., Ventura, F., Barceló, D., 2008. Rejection of pharmaceuticals in nanofiltration and reverse osmosis membrane drinking water treatment. Water. Res. 42 (14), 3601–3610.

Ray, H., Hannoun, D., Assouline, J., Vickers, J., Dickenson, E., 2024. Pilot evaluation of the potential log removal credit using chemical markers for reverse osmosis and nanofiltration. Desalination. 572, 117148.

Sarkis, S., Richard, L., Mulvenna, V. and Poon, R. 2013. Guidelines for validating treatment processes for pathogen reduction, pp. 1–80, Australia.

Scales, P.J., Wijekoon, K.C., Ladwig, C., Knight, A., Allinson, M., Allinson, G., Zhang, J., Gray, S.R., Packer, M., Northcott, K., Sheehan, D., 2019. A critical control point approach to the removal of chemicals of concern from water for reuse. Water. Res. 160, 39–51.

Shin, M.G., Choi, W., Park, S.-J., Jeon, S., Hong, S., Lee, J.-H., 2022. Critical review and comprehensive analysis of trace organic compound (TOrC) removal with polyamide RO/NF membranes: mechanisms and materials. Chem. Eng. J. 427, 130957.
Tang, S.C., Birnhack, L., Nativ, P., Lahav, O., 2017. Highly-selective separation of divalent ions from seawater and seawater RO retentate. Sep. Purif. Technol. 175, 460–468.
Victorian Department of Health 2013. Guidelines for validating treatment processes for pathogen reduction: supporting class A recycled water schemes in Victoria. Victorian Department of Health, Melbourne, Australia.
WaterVal 2017. Reverse osmosis and nanofiltration: WaterVal Validation Protocol, Australian WaterSecure Innovations Ltd, Brisbane, Australia.
WHO (eds) 2017. Potable reuse: guidance for producing safe drinking-water, World Health Organization, Geneva, Switzerland.
Wu, Y., Chen, M., Lee, H.-J., Ganzoury, M.A., Zhang, N., De Lannoy, C.-F., 2022. Nanocomposite polymeric membranes for organic micropollutant removal: a critical review. ACS. ES. T. Eng. 2 (9), 1574–1598.