Fixed-Flowrate Total Water Network Synthesis under Uncertainty with Risk Management

Cheng Seong Khor¹,², Benoit Chachuat¹, Nilay Shah¹

¹Centre for Process Systems Engineering (CPSE), Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom.

²Chemical Engineering Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia.

Abstract

This work addresses the problem of integrated water network synthesis under uncertainty with risk management. We consider a superstructure consisting of water sources, regenerators, and sinks that leads to a mixed-integer quadratically-constrained quadratic program (MIQCQP) for a fixed-flowrate total water network synthesis problem. Uncertainty in the problem is accounted for via a recourse-based two-stage stochastic programming formulation with discrete scenarios that gives rise to a multiscenario MIQCQP comprising network design in the first stage and its operation in the second stage acting as recourse. In addition, we extend the model to address risk management using the Conditional Value-at-Risk (CVaR) metric. Because a large number of scenarios are often required to capture the underlying uncertainty of the problem, causing the model to suffer from the curse of dimensionality, we propose a stepwise solution strategy to reduce the computational load. We illustrate this methodology on a case study inspired from the water network of a petroleum refinery in Malaysia. The presence of nonconvex bilinear terms necessitates the use of global optimization techniques for which we employ a new global
MIQCQP solver, GAMS/GloMIQO and verify the solutions with BARON. Our computational results show that total water network synthesis under uncertainty with risk management problems can be solved to global optimality in reasonable time.

**Keywords**: water network; uncertainty; fixed-flowrate; multiscenario; mixed-integer nonlinear program (MINLP); Conditional Value-at-Risk (CVaR)

1 **Introduction**

The optimal synthesis of water network presents a significant challenge for the design of process systems particularly in the face of scarcity of freshwater resources and increasingly stringent environmental regulations on effluent discharge. A specific class of the problem termed as total water network synthesis involves a simultaneous consideration of both water-using units and water (or wastewater) treatment operations (Foo, 2009). Water-using units represent water sources or sinks, including freshwater sources, with their corresponding contaminant concentrations. Water treatment or regeneration operations act as intermediate processes to reduce contaminant levels as necessary before the sources can be subject to reuse/recycle in the sinks. The goal is to synthesize a network that integrates these water-using and water-regeneration operations by optimizing a certain objective, which is typically “environomic” in nature, i.e., as based on economics in maximizing profit or minimizing cost as well as meeting certain environmental sustainability criteria, while complying with constraints on the water users and/or final discharge limits to the environment.
In our earlier work (Khor et al., 2012b), a deterministic fixed-flowrate formulation of the water network synthesis problem is presented that assumes fixed values of all model parameters. However, in actual operating conditions, there are often significant variations or stochastics in the parameter values. Indeed, literature data on effluent quality in process plants typically indicates significant variability in the regenerator efficiency for contaminant removals (Tchobanoglous et al., 2004). Figure 1 displays a representative trend of such removal efficiencies for six contaminants over a duration of one month as sampled to compare the influent and effluent of a reverse osmosis skid at a petroleum refinery in Malaysia (Khor et al., 2009). As evidenced from the plot, there are substantial variations in the removal ratio parameter of such a membrane regenerator unit, thus it is imperative to account for uncertainty in this parameter in a model formulation. In general, physical reasons contributing to uncertain contaminant removals are mainly due to fluctuations in operating condition as a result of fouling and leaks (resulting from ageing) in pipelines. For a membrane regenerator unit, such physical phenomena may lead to frictional pressure loss in the membrane channel and pressure drop in the module manifolds with consequential varying operating conditions (Maskan et al., 2000). Note that these sources of uncertainty are externally imposed, i.e., they are exogenous (as opposed to endogenous) uncertainty.
Nonetheless, only a small fraction of work in the literature to date considers the arguably more practical problem of water network synthesis under uncertainty. The work by Koppol and Bagajewicz (2003) represents a novel attempt at addressing uncertainty in water network synthesis. A mixed-integer linear program (MILP) is proposed that adopts a discrete scenario generation approach using bounded uniform distribution to represent uncertainty in contaminant mass load of water-using units and coupled with a scenario reduction technique. Additionally, the model considers financial risk management on the total network cost. The authors advocate that it is not possible to mitigate risk when the operating cost is much larger than the capital cost because a design with minimum expected operating cost usually poses minimum risk, but when capital cost is comparable to operating cost, reuse of wastewater is amenable to reducing risk.
Al-Redhwan et al. (2005) employs two-stage stochastic programming to formulate a fixed-load water network model with uncertain mass load parameter. Karuppiah and Grossmann (2008) also utilizes a similar framework with consideration for an additional uncertain parameter in the contaminant removal ratio. A major contribution of this work is to globally optimize the nonconvex mixed-integer nonlinear program (MINLP) by using relaxation via McCormick’s (1976) convex and concave envelopes for the bilinear terms and linear overestimators constructed from secants for concave terms in the objective function. The authors propose a spatial branch-and-cut scheme with a Lagrangean decomposition approach to solve the multiscenario model.

Tan et al. (2007) utilizes a Monte Carlo simulation approach to investigate the sensitivity of solutions obtained from water pinch analysis by accounting for uncertainty in the mass load parameter. It is found that fluctuations in processing conditions can lead to process disruptions and affect product quality and process stability. In another work considering mass load uncertainty, Zhang et al. (2009) attempts to develop a resilient water network by aiming for a low value of a metric introduced as the tolerance amount of a water-using unit, which measures the difference between the limiting and actual mass load of a network. Feng et al. (2011) addresses the design of a multicontaminant water-using network with consideration for mass load uncertainty. The proposed NLP-based approach adjusts stream flows through a combination of optimization and heuristics while preserving the optimal water network structure obtained under nominal condition for minimum freshwater use.
In a series of papers, Chang et al. (2009) employ a flexibility analysis approach to cater for a water network operation under uncertainty in freshwater quality, mass load, removal ratios, and maximum inlet and outlet concentrations of both water-using and treatment units. A MINLP model is used to assess the feasibility of a nominal design and to improve its flexibility index by relaxing the maximum freshwater capacity and installing new and/or removing existing pipelines. An initialization procedure is developed to help convergence to a global optimum.

This line of work is continued by Riyanto and Chang (2010) that adopts a MINLP with heuristics based on active constraints to improve the operational flexibility of a water network design. Later, Li and Chang (2011) propose a simplified NLP for an efficient computation of the flexibility index that obviates a need for elaborate initialization and the high computational expense entailed with use of a MINLP.

Hung and Kim (2011) consider uncertain inlet flowrate and mass load of water-using units. The uncertain parameters are represented as average values via a multiperiod MINLP formulation. The model addresses uncertainty by incorporating buffer tanks to handle sudden contaminant level rise and supplementary pipelines to supply freshwater to ensure feasibility. An insights-based decomposition strategy that involves iteratively solving a sequence of MILP and LP relaxations by fixing the concentration and flowrate of water-using units, respectively is proposed to initialize the solution.

Synthesis of water regeneration systems particularly within the context of a total water network system investigates options for regeneration–reuse besides direct reuse or recycle. Tan et al. (2009) consider the use of partitioning regenerators specifically membrane technologies to
optimize water allocation involving multiple water sources and sinks in an MILP model. Khor et al. (2011) incorporates a detailed nonlinear membrane regenerator model as exemplified for a reverse osmosis network in a fixed-flowrate total water network synthesis MINLP formulation. Sotelo-Pichardo et al. (2011) address the retrofit of a total water network by applying reconfiguration strategies that include repiping, modifying the capacity and performance of existing water-using units, and reuse of existing regenerators as well as installing new regenerators. A recent work by Tokos et al. (2013) employs a biobjective MINLP optimization model that accounts for both economic and environmental impacts of batch and semi-continuous total water network systems. A Pareto curve of optimal solution is obtained by implementing variants of the weighted-sum method. On the other hand, Lim et al. (2013) propose a reformulation approach for biobjective to single objective total water network synthesis problems that are incorporated with eco-design principles.

Table 1 summarizes the salient features of some representative work on optimal water network synthesis under uncertainty.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Model Formulation</th>
<th>Uncertain Parameter and Representation</th>
<th>Solution Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koppol and Bagajewicz (2003)</td>
<td>MILP with financial risk management</td>
<td>Contaminant mass load using discrete scenarios</td>
<td>Branch-and-bound with CPLEX</td>
</tr>
<tr>
<td>Al-Redhwan et al. (2005)</td>
<td>Two-stage stochastic NLP</td>
<td>Contaminant mass load using discrete scenarios</td>
<td>Local optimization using generalized reduced gradient (with CONOPT2)</td>
</tr>
<tr>
<td>Reference</td>
<td>Model Formulation</td>
<td>Uncertain Parameter and Representation</td>
<td>Solution Strategy</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Chang et al. (2009),</td>
<td>MINLP via flexibility</td>
<td>Freshwater quality, mass load, removal ratios, and maximum inlet and outlet concentrations of both water-using and treatment units using</td>
<td>Initialization based on heuristics to help convergence to global optimum</td>
</tr>
<tr>
<td>Riyanto and Chang (2010),</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li and Chang (2011)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. (2009)</td>
<td>Water pinch analysis</td>
<td>Mass load (details on representation are not provided)</td>
<td>Introduces a resilience metric that aims for a low difference between the limiting and actual mass load</td>
</tr>
<tr>
<td>Feng et al. (2011)</td>
<td>NLP and MINLP</td>
<td>Mass load using discrete deviation from nominal values</td>
<td>Adjusts stream flows while maintaining nominal network structure with minimum freshwater use</td>
</tr>
<tr>
<td>Hung and Kim (2011)</td>
<td>Multiperiod MINLP</td>
<td>Inlet flowrate and mass load of water-using units using average values</td>
<td>Decomposition-based iterative solution of sequence of MILP and LP relaxations</td>
</tr>
</tbody>
</table>

It is the aim of this paper to investigate the problem of fixed-flowrate water network synthesis under uncertain conditions by extending the formulation presented in Khor et al. (2012b). The formulation employs the classical recourse-based two-stage stochastic program and uses discrete scenarios to approximate the underlying probability distribution of the uncertain removal ratio parameters. The two-stage framework comprises network design in the first stage and its operation in the second stage acting as recourse. The resulting formulation gives rise to a multiscenario MINLP, or more specifically, a mixed-integer quadratically-constrained quadratic program (MIQCQP).

The paper also addresses risk management in view of the possibility of high costs of freshwater and wastewater treatment in view of the deregulation and privatization taking place in water resources management (Bakker, 2010). The situation is compounded with increasingly stressed
water resources globally due to growing population, pollution, inefficient irrigation, etc. (International Water Management Institute (IWMI), 2007). Consequently, a water network is exposed to higher costs if more freshwater is demanded and/or more wastewater needs treatment to meet operating and regulation requirements. According to a 2008 report by the United Nations Environment Programme (UNEP, 2008), industry is the second largest user of freshwater worldwide after agriculture (Zimmerman et al., 2008). Hence, there are practical reasons to enhance our water network synthesis formulation to consider risk management. We strive to meet this purpose by appending the Conditional Value-at-Risk (CVaR) metric to the objective function. However, a large number of scenarios are often required to capture the underlying uncertainty of the problem, causing the model to suffer from the curse of dimensionality. To circumvent this issue, we adopt a stepwise strategy that tractably solves a risk-neutral version of the formulation using a small number of scenarios to reduce the computational load in a subsequent scenario-ordering scheme for simulating the Value-at-Risk (VaR) parameter. The strategy also serves to provide initial values for solving the nonlinear risk-averse mean–CVaR model.

The presence of nonconvex bilinear terms necessitates the implementation of a global optimization approach to ensure reliable solutions. In this regard, we solve both the risk-neutral and risk-averse formulations using a new global MIQCQP solver, GAMS 24.0.2/GloMIQO 2.1 (Misener and Floudas, 2012). We also verify the solutions with GAMS 24.0./BARON 11.9.1 (Tawarmalani and Sahinidis, 2005), a state-of-the-art commercial global solver.
The rest of the article is organized as follows. Section 2 formally describes the problem under uncertainty which we wish to address, while Section 3 proceeds to outline in general, our proposed optimization-based framework with a computationally-tractable solution strategy. Then, we present formulations for the risk-neutral and risk-averse problems of water network synthesis as two-stage stochastic programs in Section 4. Section 5 reports computational results using the two commercial solvers before concluding the paper in Section 6.

2 Problem Statement

This work aims to optimally synthesize a total water network, which integrates both water-using and water treatment operations, as given the following elements of a water network:

- a set of fixed-flowrate water sources \( i, i \in I \), with known flowrates \( F_{\text{SO}}(i) \) and concentrations \( C_{\text{SO}}(i,q) \) of the contaminants \( q \in Q \);
- a set of fixed-flowrate water sinks \( j, j \in J \), with known flowrate requirements \( F_{\text{SI}}(j) \) and maximum allowable inlet concentration limit \( C_{\text{max}}(j,q) \) of the contaminants \( q \);
- a set of water regenerators \( k, k \in K \), with uncertain removal ratios of contaminants; and
- a freshwater source \( i = \text{FW} \) with known contaminant concentrations that can be purchased to supplement available water sources.

2.1 Uncertainty Description
As elucidated, uncertainty in water network synthesis primarily concerns contaminant removal ratios in the regeneration subnetwork. Removal ratio or removal efficiency \( R(k, q, s) \) of a targeted contaminant \( q \) for a regenerator \( k \) is defined as the amount \( q \) (expressed in ratio or percentage) removed by \( k \) from an incoming water stream under an operating scenario \( s \). The uncertain removal ratio parameters can be assumed to be represented by a bounded uniform probability distribution within a certain known interval defined by its lower and upper bounds. This assumption is reasonable because it is possible to specify the range within which removal ratios vary. In our model, we treat the removal ratios as random variables that may take on multiple possible discrete values under different network operation scenarios. For this purpose, we sample random values of the removal ratios from a known discrete probability distribution using a Monte Carlo-simulation based technique.

### 2.2 Two-Stage Stochastic Programming Framework

Fluctuations in the removal amounts, as exemplified by Figure 1, may consequently necessitate a recourse action of adjusting freshwater supply for diluting flows to the regenerators and the sinks. Such a recourse action ensures that the network may still be operated without violating the maximum allowable inlet concentration (MAIC) limits of the sinks, particularly for discharge limits that are bounded by strict environmental regulations. In this regard, another major recourse action involves regulating wastewater flows to the effluent treatment system (ETS) to comply with discharge limits.
The decision variables are divided into first and second stages corresponding to before and after information on the uncertainty have been revealed, respectively. Similar to the approach of Karuppiah and Grossmann (2008), the first-stage decisions pertain to determining the optimal network topology in terms of interconnections of the entities comprising sources, regenerators, and sinks. These decisions are represented by two types of design variables, which have to be taken prior to revelation of uncertainty in the system. They are the 0–1 variables \( y_{A,B}(\alpha, \beta) \) describing existence of the interconnections between an origin entity A (as indexed by \( \alpha \)) and a destination entity B (as indexed by \( \beta \)), as well as the continuous variables on the maximum allowable total stream flows \( F_{A,B}^{\text{max}}(\alpha, \beta) \).

The second-stage decisions are recourse or corrective actions with respect to the first-stage decisions after the uncertainty is revealed. Here, the optimal operating policy of the network is established by the continuous recourse variables of flows and concentrations in each network interconnection, particularly freshwater supply and wastewater inlet flows through the regenerators, as stated earlier. These second-stage operational flows can be different during network operation in accordance with the discrete-valued scenarios assumed by the uncertain removal ratios. Thus, flow in an interconnection is sent at a rate \( F_{A,B}(\alpha, \beta, s) \) from an origin entity A to a destination entity B for a particular realization of the operating scenario \( s \), while the associated contaminant concentrations leaving and entering a regenerator \( k \) for a set of contaminants \( q \in Q \) is given by \( C_A(k, q, s) \). Figure 2 provides a diagrammatic representation of the proposed framework for a piping interconnection between two network elements A and B.
Conventional two-stage stochastic programming, which exploits the underlying structure of the problem presented here, provides a suitable framework because uncertainty in removal ratios is exogenous in nature as aforementioned. Moreover, a stochastic program with complete recourse formulation (Sen and Higle, 1999) can be considered as opposed to a decision-dependent approach (Tarhan et al., 2009).

The objective of the proposed two-stage stochastic program is to synthesize an integrated water system that minimizes the total annualized cost of designing the network in the first stage and the expected cost of operating the network in the second stage. A similar framework has been adopted by Karuppiah and Grossmann (2008) for a water network synthesis problem but nonetheless for a fixed-load water-using unit representation. This modeling framework has also found recent applications in Li et al. (2011) for a pooling problem of natural gas production and Yunt et al. (2008) for a portable power generation system with varying demand levels. We
emphasize again that it is at the second stage that the flexibility of the design is checked by considering variations through the recourse operating variables in accommodating the uncertain parameter realizations. This part of the model is also the most computationally demanding. It is noteworthy that the integer decisions do not appear in the second stage and hence, our formulation do not fall under the class of problems of stochastic mixed-integer recourse programs, which is treated, among others, in the work of Schultz (2003).

3 Water Network Model Formulation

3.1 Superstructure Representation

Based on the work of Khor et al. (2012a, b), we adopt a source–regenerator–sink superstructure that allows all feasible interconnections of the network entities to cater for alternative configurations involving direct water reuse/recycle, regeneration–reuse, and regeneration–recycle. Figure 3 shows a simplified superstructure around both the membrane and non-membrane regenerators. The representation leads to a MINLP formulation by adopting the notations explained earlier as based on Figure 2.
3.2 Regenerator Models

Conventional models for water regenerators typically involve a single pooled inlet stream from a source and a single pooled outlet stream to a sink (see, e.g., Karuppiah and Grossmann (2008)), which we directly employ for the non-membrane regenerators (NM). On the other hand, we adapt such a representation for the membrane regenerators (M) that typically consist of two outlets, namely a permeator stream (MP) which is of lower concentration than a rejector stream (MR). The outlet flow of a regenerator in a conventional model would have been separated into a permeator and a rejector by using a splitter, but both resultant streams would then share the same concentration. To avoid such a possible concentration discrepancy, we represent the permeator and the rejector as units instead of streams in our proposed linear membrane regenerator model. Our proposed model also accounts for full interconnectivity among the regenerators by allowing an interconnection between a non-membrane regenerator $k$ with another non-membrane regenerator $k'$ or a membrane regenerator $k''$—similar cases apply for a permeator and a rejector.
3.3 Optimization Model Formulation

3.3.1 Material Balances

3.3.1.1 Water Sources

Mass balance around a source except freshwater for each scenario is represented as:

\[ \sum_{k \in K_{NM}} F_{SO,NM}(i,k,s) + \sum_{k' \in K_M} (F_{SO,MP}(i,k',s) + F_{SO,MR}(i,k',s)) + \sum_{j \in J} F_{SO,SI}(i,j,s) = F_S(i), \quad \forall i \in I \setminus \{\text{Freshwater}\}, \forall s \in S. \]  (1)

For a freshwater source, its supply flowrate is a recourse variable that varies with the scenarios:

\[ \sum_{k \in K_{NM}} F_{SO,NM}(i,k,s) + \sum_{k' \in K_M} (F_{SO,MP}(i,k',s) + F_{SO,MR}(i,k',s)) + \sum_{j \in J} F_{SO,SI}(i,j,s) = F_{FW}(s), \quad i = \text{Freshwater}, \forall s \in S. \]  (2)

3.3.1.2 Non-Membrane Regenerators

Mass and concentration balances around the inlet of a non-membrane regenerator:
\[
\sum_{i \in I} F_{i}^{i} (k, j, s) + \sum_{k' \in K_{NM}} F_{k'}^{k'} (k, j, s) + \sum_{k' \in K_{NM}} \left( F_{k}^{k} (k', j, s) + F_{k'}^{k'} (k', j, s) \right)
\]

= \( F_{i}^{i} (k, j, s), \quad \forall k \in K_{NM}, \forall s \in S. \tag{3} \)

\[
\left[ \sum_{i \in I} F_{i}^{i} (k, j, s) C_{i}^{i} (i, q) + \sum_{k' \in K_{NM}} F_{k'}^{k'} (k, j, s) C_{k'}^{k'} (k', q, s) \right] + \sum_{k' \in K_{NM}} \left( F_{k'}^{k'} (k, j, s) C_{k'}^{k'} (k', q, s) + F_{k'}^{k'} (k, j, s) C_{k'}^{k'} (k', q, s) \right)
\]

= \( F_{i}^{i} (k, j, s) C_{k}^{k} (k, q, s), \quad \forall k \in K_{NM}, \forall q \in Q, \forall s \in S. \tag{4} \)

where \( F_{i}^{i} (k, j, s) \) is the mass flowrate of the feed stream to a non-membrane regenerator.

Mass and concentration balances around the outlet of a non-membrane regenerator:

\[
\sum_{j \in J} F_{i}^{i} (k, j, s) + \sum_{k' \in K_{NM}} F_{k'}^{k'} (k, j, s) + \sum_{k' \in K_{NM}} \left( F_{k}^{k} (k', j, s) + F_{k'}^{k'} (k', j, s) \right)
\]

= \( F_{i}^{i} (k, j, s), \quad \forall k \in K_{NM}, \forall s \in S. \tag{5} \)

\[
(1 - R(k, q, s)) C_{k}^{k} (k, q) = C_{i}^{i} (i, q), \quad \forall k \in K_{NM}, \forall q \in Q, \forall s \in S. \tag{6} \)

Recall that \( R(k, q, s) \) is a random variable representing the uncertain parameter of removal ratio. It is noteworthy that the uncertainty appears as a coefficient on the left-hand side of the constraints.
3.3.1.3 Membrane Regenerators

Mass and concentration balances around entry point to the permeator of a membrane regenerator:

\[
\sum_{i \in I} F_{SO,MP} (i, k, s) + \sum_{k' \in K_{NM}} F_{NM,MP} (k', k, s) + \sum_{k'' \in K_{MP}} \left( F_{MP,MP} (k'', k, s) + F_{MR,MP} (k'', k, s) \right)
\]

\[
= F_M^F (k, s), \quad \forall k \in K_M, \forall s \in S.
\]  

(7)

\[
\left\{ \begin{align*}
\sum_{i \in I} (F_{SO,MP} (i, k, s)) C_{SO} (i, q) + \sum_{k' \in K_{NM}} F_{NM,MP} (k', k, s) C_{NM} (k', q, s) \\
+ \sum_{k'' \in K_{MP}} F_{MP,MP} (k'', k, s) C_{MP} (k'', q, s) + \sum_{k'' \in K_{MR}} F_{MR,MP} (k'', k, s) C_{MR} (k'', q, s)
\end{align*} \right.
\]

\[
= F_M^F (k, s) C_M^F (k, q, s), \quad \forall k \in K_M, \forall q \in Q, \forall s \in S.
\]  

(8)

Mass and concentration balances around entry point to the rejector of a membrane regenerator:

\[
\left\{ \begin{align*}
\sum_{i \in I} F_{SO,MR} (i, k, s) + \sum_{k' \in K_{NM}} F_{NM,MR} (k', k, s) + \sum_{k'' \in K_{MP}} \left( F_{MP,MR} (k'', k, s) + F_{MR,MR} (k'', k, s) \right)
\end{align*} \right.
\]

\[
= F_M^F (k, s), \quad \forall k \in K_M, \forall s \in S.
\]  

(9)

\[
\left\{ \begin{align*}
\sum_{i \in I} (F_{SO,MR} (i, k, s)) C_{SO} (i, q) + \sum_{k' \in K_{NM}} F_{NM,MR} (k', k, s) C_{NM} (k', q, s) \\
+ \sum_{k'' \in K_{MP}} F_{MP,MR} (k'', k, s) C_{MP} (k'', q, s) + \sum_{k'' \in K_{MR}} F_{MR,MR} (k'', k, s) C_{MR} (k'', q, s)
\end{align*} \right.
\]

\[
= F_M^F (k, s) C_M^F (k, q, s), \quad \forall k \in K_M, \forall q \in Q, \forall s \in S.
\]  

(10)
It is noted that the mass flowrate of an entry point to a membrane regenerator is represented as \( F_M^F(k,s) \) without being distinguished by a subscript denoting a permeator or a rejector. This is because physically, the permeator and the rejector are entities of the same membrane regenerator unit, hence careful formulation has been employed to maintain consistency with the actual physical configuration, which also lends a more natural formulation to the underlying problem.

Mass and concentration balances around the permeator of a membrane regenerator:

\[
\sum_{j \in J} F_{MP,SI}(k,j,s) + \sum_{k' \in K_{NM}} F_{MP,NM}(k,k',s) + \sum_{k' \in K_{NM} \atop k' \neq k} \left( F_{MP,MP}(k,k',s) + F_{MP,MR}(k,k',s) \right) = \lambda(k)F_M^F(k,s), \quad \forall k \in K_M, \forall s \in S
\]  

\[
\left( 1 - R(k,q,s) \right) C_M^F(k,q,s) = \lambda(k) C_{MP}(k,q,s), \quad \forall k \in K_M, \forall q \in Q, \forall s \in S.
\]  

Mass and concentration balances around the rejector of a membrane regenerator:

\[
\sum_{j \in J} F_{MR,SI}(k,j,s) + \sum_{k' \in K_{NM}} F_{MR,NM}(k,k',s) + \sum_{k' \in K_{NM} \atop k' \neq k} \left( F_{MR,MP}(k,k',s) + F_{MR,MR}(k,k',s) \right) = \left( 1 - \lambda(k) \right) F_M^F(k,s), \quad \forall k \in K_M, \forall s \in S.
\]

\[
R(k,q,s) C_M^F(k,q,s) = \left( 1 - \lambda(k) \right) C_{MR}(k,q,s), \quad \forall k \in K_M, \forall q \in Q, \forall s \in S.
\]
where $\lambda(k)$ is the split ratio on flow as based on the liquid phase recovery factor for the permeator.

### 3.3.1.4 Sinks

Mass balance (linear) around a sink except the discharge:

$$\sum_{i \in I} F_{SO,SI}(i, j, s) + \sum_{k \in K_{NM}} F_{NM,SI}(k, j, s) + \sum_{k' \in K_{M}} (F_{MP,SI}(k', j, s) + F_{MR,SI}(k', j, s)) = F_S(j), \quad \forall j \in J \setminus \{\text{Discharge}\}, \forall s \in S. \quad (15)$$

The discharge flowrate is a recourse variable that is dependent on the scenarios:

$$\sum_{i \in I} F_{SO,SI}(i, j, s) + \sum_{k \in K_{NM}} F_{NM,SI}(k, j, s) + \sum_{k' \in K_{M}} (F_{MP,SI}(k', j, s) + F_{MR,SI}(k', j, s)) = F_W(s), \quad j = \text{Discharge}, \forall s \in S. \quad (16)$$

### 3.3.2 Quality Constraints

Constraint on the quality requirement for a sink as stipulated by its maximum allowable inlet concentration (MAIC) limit for a contaminant is given by:
\[
\sum_{i \in I} F_{SO,SI} (i, j, s) C_{SO} (i, q) + \sum_{k \in KNM} F_{NM,SI} (k, j, s) C_{NM} (k, q, s) \\
+ \sum_{k' \in K_M} (F_{MP,SI} (k', j, s) C_{MP} (k', q, s) + F_{MR,SI} (k', j, s) C_{MR} (k', q, s)) \\
\leq F_{SI} (j) C_{\text{max}} (j, q), \quad \forall j \in J, \forall q \in Q, \forall s \in S
\]

where \( C_{\text{max}} (j, q) \) is MAIC for a regenerator \( k \) for each of the contaminants \( q \). For a discharge sink, its quality constraint is given by:

\[
\sum_{i \in I} F_{SO,SI} (i, j, s) C_{SO} (i, q) + \sum_{k \in KNM} F_{NM,SI} (k, j, s) C_{NM} (k, q, s) \\
+ \sum_{k' \in K_M} (F_{MP,SI} (k', j, s) C_{MP} (k', q, s) + F_{MR,SI} (k', j, s) C_{MR} (k', q, s)) \\
\leq F_{W} (s) C_{\text{max}} (j, q), \quad j = \text{Discharge}, \forall q \in Q, \forall s \in S
\]

### 3.3.3 Operational Constraints

The operational constraints link the second-stage operating flow variables in every scenario with their corresponding first-stage design flow variables to guarantee operating feasibility. From a physical viewpoint, these constraints serve to accommodate the various possible operating policies by ensuring that the interconnections, which are essentially pipelines, are operable. Thus, they are designed in such a way that the maximum design flows have to be greater than their operational counterparts in each scenario:
\[ F_{AB}(\alpha, \beta, s) \leq F_{AB}^{\text{max}}(\alpha, \beta), \]
\[ \forall (\alpha, A) \in (I, SO) \cup (K_{NM}, NM) \cup (K_{M}, MP) \cup (K_{M}, MR). \]
\[ \forall (\beta, B) \in (J, SI) \cup (K_{NM}, NM) \cup (K_{M}, MP) \cup (K_{M}, MR). \]
\[ \forall s \in S. \]

In addition, linear logical constraints on selection of the interconnections link their existence, as given by \( y_{AB}(\alpha, \beta) \), with the corresponding maximum design flows \( F_{AB}^{\text{max}}(\alpha, \beta) \):

\[ F_{AB}^{L}(\alpha, \beta) y_{AB}(\alpha, \beta) \leq F_{AB}^{\text{max}}(\alpha, \beta) \leq F_{AB}^{U}(\alpha, \beta) y_{AB}(\alpha, \beta), \]
\[ \forall (\alpha, A) \in (I, SO) \cup (K_{NM}, NM) \cup (K_{M}, MP) \cup (K_{M}, MR), \]
\[ \forall (\beta, B) \in (J, SI) \cup (K_{NM}, NM) \cup (K_{M}, MP) \cup (K_{M}, MR). \]

where \( F_{AB}^{L}(\alpha, \beta) \) and \( F_{AB}^{U}(\alpha, \beta) \) are the lower and upper bounds on maximum flows, respectively. The constraints imply that if an interconnection is optimally selected, then its design flows can take values between the specified bounds; otherwise, these values are zero.

4 Solution Methodology

4.1 Two-Phase Solution Approach

A large number of scenarios are often required to capture the underlying uncertainty of a water network synthesis problem, causing the model to suffer from the curse of dimensionality. To handle this challenge of high computational load, we employ a sequential two-phase solution approach as shown in Figure 4 in which the first phase (Phase I) generates a risk-neutral solution.
Based on Phase I solution, the VaR parameter is estimated using a scenario-ordering procedure and subsequently utilized in Phase II to compute the intended risk-averse solution.

**PHASE I: Synthesis under uncertainty**

**STEP 1. Scenario generation using Monte Carlo simulation-based pseudorandom number generation**
Discretize uncertainty space by random sampling of removal ratios with fixed lower & upper limits for a small no. of scenarios NS

**STEP 2. Solution of risk-neutral multisenario model**
Compute optimal risk-neutral solution $x^*$ for NS

**Solution found in reasonable time?**

**PHASE II: Synthesis under uncertainty with risk management**

**STEP 3. Scenario reduction procedure [optional]**
Reduce computational burden by determining minimum no. of scenarios $NS_{min}$

$x^*$ and $NS_{min}$

**STEP 4. Simulation of Value-at-Risk (VaR) via scenario ordering**
Estimate VaR from cumulative distribution function for desired confidence level $\beta$

$x^*$ (as initial values), $NS_{min}$, and VaR

**STEP 5. Solution of risk-averse multisenario mean-CVaR stochastic program**
Compute optimal risk-averse solution $x^{**}$ for desired $\beta$ and risk factor $\theta$ for $NS_{min}$ or a representative no. of scenarios

**STOP: Optimal risk-averse solution**

Figure 4. Proposed two-phase solution approach for water network synthesis under uncertainty with risk management
4.1.1 Phase I: Risk-Neutral Model

The goal of this step is to obtain a solution with minimum cost without being averse to taking risks, which implies that the result obtained is one in which risky decisions are not penalized. Solution of such a risk-neutral multiscenario model will be used subsequently to estimate the VaR parameter. If a solution of the model cannot be found in reasonable time, we may consider generating a smaller but still representative set of scenarios to solve the multiscenario model.

We pose a multiscenario risk-neutral program based on our proposed model in Khor et al. (2012b) which employs a linear membrane regenerator technology representation, but reformulated to address uncertainty in the removal ratio parameters. The formulation results in a mixed-integer quadratically-constrained quadratic program (MIQCQP), a special form of mixed-integer nonlinear program (MINLP), as presented in the rest of this section. $K_{NM}$ denotes the set of non-membrane regenerators, while the set $K_{M}$ of membrane regenerators contains ordered pairs of the permeator MP and the rejector MR.

4.1.1.1 Scenario Generation

We first employ a Monte Carlo approach using pseudorandom number generation to generate scenarios that approximate the original full space of the probability distribution of the uncertain parameters. The values of the uncertain parameters in each scenario and their probability of occurrence are randomly sampled using pseudorandom numbers and are therefore independent of the first-stage decisions. Alternatively, the scenarios can be considered as equiprobable. In this
regard, our approach differs from previous work (e.g., Koppol and Bagajewicz (2003)) in the sense that we consider the scenarios as constructed from random variables, instead of historical observations that may be susceptible to noisy data. As will be shown in the discussion of the numerical results in section 5, the quality of the data that constitute the scenarios is essential to ensure a reliable solution is obtained.

A scenario is thus fully described by the values of the uncertain parameters with its associated probability, and provides a set of candidate optimal solutions exhibiting varying risk propensity. The number of all scenarios is given by the number of scenarios NS for each uncertain parameter raised to the power of the number of uncertain parameters NP, hence the problem is of combinatorial nature. In this regard, deciding a priori on the number of scenarios to be generated is a tradeoff between the computational load and the need to sufficiently capture the underlying uncertainty of the problem. This calls for a combination of the knowledge of a modeler and the experience of a decision-maker (who is typically the desired end-user) as well as past practices reported in the literature for a similar class of problems.

4.1.1.2 Objective Function

The objective function of the model involves minimizing the first-stage costs of the piping interconnections (PI) and the expected costs of the second-stage recourse operations:

\[
\min \sum_{s \in S} \left( p(s) \left( OC_{PI}(s) + OC_{FW}(s) + OC_{W}(s) \right) \right) 
\]

(21)
where $\delta$ is the annualized capital charge factor that represents tradeoffs between capital cost (CC$_{PI}$) and operating cost (OC$_{PI}$) of the interconnections. Components of the first-stage interconnection costs are taken as linear functions of the 0–1 variables denoting their existences and the continuous maximum flow variables.

On the other hand, components of the expected operating costs are assumed to be linear functions of the recourse operational flows (for OC$_{PI}$), as well as cost penalties due to recourse actions involving freshwater consumption (OC$_{FW}$) and wastewater treatment in the ETS (OC$_{W}$). Expressions for these cost components on an annualized basis are defined as follows:

\begin{equation}
CC_{PI} = \sum_{(\alpha,A),(\beta,B) \in N} b_{A,B}(\alpha,\beta) F_{A,B}^{max}(\alpha,\beta) + d_{A,B}(\alpha,\beta) y_{A,B}(\alpha,\beta), \quad (22)
\end{equation}

\begin{equation}
OC_{PI}(s) = \sum_{(\alpha,A),(\beta,B) \in N} c_{A,B}(\alpha,\beta) F_{A,B}(\alpha,\beta,s), \quad (23)
\end{equation}

\begin{equation}
OC_{FW}(s) = c_{FW} F_{FW}(s) H, \quad \forall s \in S, \quad (24)
\end{equation}

\begin{equation}
OC_{W}(s) = c_{W} F_{W}(s) H, \quad \forall s \in S, \quad (25)
\end{equation}

where $b_{A,B}$, $c_{A,B}$, $d_{A,B}$, $c_{FW}$, and $c_{W}$ are the respective cost coefficients. $H$ indicates hours of plant operation per annum.
4.1.2 Phase II: Risk-Averse Model Formulation

The consideration for incorporating risk in a two-stage stochastic program is to account for the variability, besides the expected value, of the second-stage function as a measure of the desired intensity of the recourse operations. In this work, we adopt CVaR as a proxy to represent risk in the objective function. CVaR has received increasing applications in process systems engineering in recent work (Colvin and Maravelias, 2011; Tometzki and Engell, 2011; Verderame and Floudas, 2010; 2011).

CVaR is a risk measure originally employed to reduce the probability that an investment portfolio will incur high losses. It is closely related to VaR that measures the maximum expected loss in the value of a risky entity at a certain confidence interval (typically 95% or 99%) over a given period under normal market conditions. CVaR is the expected loss given that the actual loss exceeds some VaR threshold at the same confidence level (Rockafellar and Uryasev, 2002; Szego, 2002). For example, at a one-month 95% confidence interval, VaR reports a single value with 95% certainty that that is the value of the maximum expected loss. CVaR measures the expectation that the value is greater than VaR. Within a water network design setting, for instance, if VaR for a network cost (capital and operational) is $1 million at a one-month 95% confidence interval, this implies that there is a 5% probability that the cost will drop more than $1 million over any given month. CVaR is the expected loss in the network cost that is greater than $1 million over the same duration for the same confidence interval.
Rockafellar and Uryasev (2000, 2002) propose the following linear (convex) approximation of CVaR for a discrete probability distribution function associated with the first- and second-stage decision variables \( x \) and \( y(s) \), respectively, which can be interpreted as a weighted average between VaR and the losses exceeding VaR:

\[
\text{CVaR} \approx \text{VaR} + \frac{1}{1-\beta} \sum_{s=1}^{NS} p(s)\left( f(x, y(s)) - \text{VaR} \right)
\]

where \( \beta \) denotes confidence level.

### 4.1.2.1 Simulation of Value-at-Risk (VaR)

The value of the risk measure VaR for a problem can be simulated in an offline manner by developing its associated cumulative distribution function (CDF) plot or risk curve (Barbaro and Bagajewicz, 2004) by using the risk-neutral solution. VaR is expressed as:

\[
\text{VaR}(x, y) = \min_{x, y} \left\{ f(x, y(s)) \big| G(f) \geq \alpha \right\}
\]

where \( f(x, y(s)) \) is the objective function of the optimization problem. \( G(f) \) is the cumulative distribution function plot for the objective, and it is straightforward to determine VaR from such a plot (Santoso et al., 2005; Webby et al., 2007).
The resultant convex mean–CVaR objective function is formulated as follows using a weighting method for examining the tradeoffs between the two conflicting objectives of expected cost and risk (Khor et al., 2011):

$$\min \delta CC_{pl} + \sum_{s \in S} p(s)(OC_{pl}(s) + OC_{FW}(s) + OC_{W}(s)) + \theta CVaR(s)$$  \hspace{1cm} (28)

where $\theta$ is a decision maker-defined adjustable non-negative weights of risk factors that facilitate the tradeoff between expected cost with risk. A larger value of $\theta$ indicates a higher propensity towards risk-taking and vice versa. In a more general setting, alternative risk measures besides CVaR such as downside risk (Eppen et al., 1989) and financial risk (Barbaro and Bagajewicz, 2004) are also applicable.

The complete formulation of a risk-averse multiscenario water network synthesis under uncertainty model is given by:

$$\min \delta CC_{pl} + \sum_{s \in S} p(s)(OC_{pl}(s) + OC_{FW}(s) + OC_{W}(s))$$

$$+ \theta \left[ \text{VaR} + \frac{1}{1-\beta} \sum_{s=1}^{NS} p(s) \left( f(s) - \text{VaR} \right) \right]$$  \hspace{1cm} (29)

s.t. constraints (1)–(41).
4.1.2.2 Scenario Reduction

It is acknowledged from numerical experiments that a well-controlled choice of sample size can significantly reduce computational time and improve the accuracy of solutions. Thus, where appropriate, a scenario reduction procedure can be applied to reduce the high computational load suffered as a result of handling the potentially prohibitively large number of scenarios generated. This procedure is particularly useful to facilitate the simulation to determine VaR. A number of scenario reduction methods are available in the literature, e.g., see the work of Romisch (2009) and a statistical-based approach employed in You et al. (2009). Further references are available on this subject (e.g., Law and Kelton (2000)).

4.2 Global Optimization

The model is nonconvex in the bilinear terms arising from contaminant mixing that may result in multiple local optima (Quesada and Grossmann, 1995). Thus, implementing a global optimization approach is required to obtain a certificate of optimality and reliability to the solution.

4.2.1 Natural Variable Bounds

For the natural bounds, nonnegativity conditions are enforced for all continuous variables:
\[ F_{A,B}(\alpha, \beta, s) \geq 0, \]
\[ \forall (\alpha, A) \in (I, SO) \cup (K_{NM}, NM) \cup (K_M, MP) \cup (K_M, MR), \]
\[ \forall (\beta, B) \in (K_{NM}, NM) \cup (K_M, MP) \cup (K_M, MR) \cup (J, SI), \]
\[ \forall s \in S, \]

\[ F_A^F(k, s) \geq 0, \quad \forall (k, A) \in (K_{NM}, NM) \cup (K_M, M), \forall s \in S, \]  
\[ (31) \]

\[ C_A(k, q, s) \geq 0, \quad \forall (k, A) \in (K_M, MP) \cup (K_M, MR), \forall q \in Q, \forall s \in S, \]  
\[ (32) \]

\[ C_A^F(k, q, s) \geq 0, \quad \forall (k, A) \in (K_{NM}, NM) \cup (K_M, M), \forall q \in Q, \forall s \in S. \]  
\[ (33) \]

The nonnegativity constraints also obviate the possibility of reverse flows in a single-choice interconnection.

For the 0–1 integer variables, integrality conditions are enforced:

\[ y_{A,B}(\alpha, \beta) \in \{0, 1\}. \]  
\[ (34) \]

### 4.2.2 Tighter Variable Bounds

It is imperative to specify tight lower and upper bounds for all variables in solving nonconvex problems to obtain resulting tight relaxations. Both solvers of our choice (BARON and GloMIQO) share a common emphasis on the importance of supplying good variable bounds,
mainly for constructing convex relaxations for the nonconvex bilinear terms in executing the solution procedure. We use equation-based variable upper bounds as summarized in Table 1 for a more rigorous representation in the same manner as Ahmetović and Grossmann (2011) and Misener and Floudas (2010). The variable lower bounds are taken to be the minimum physically-feasible flowrates for controllability or economical purposes, which can be taken as 2 ton/h (Chakraborty, 2009).

### Table 2. Variable upper bounds

<table>
<thead>
<tr>
<th>Equation</th>
<th>Flow variable</th>
<th>Origin</th>
<th>Destination</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>$F_{SO,SI}(i, j, s)$</td>
<td>Source except freshwater</td>
<td>Sink</td>
<td>Minimum of the two entities</td>
</tr>
<tr>
<td>27</td>
<td>$F_{SO,A}(i, k, s)$ where $A = \text{regenerator} \in {\text{NM, MP, MR}}$</td>
<td>Source except freshwater</td>
<td>Regenerator</td>
<td>Equal to the source flow $F_{SO}(i,s)$</td>
</tr>
<tr>
<td>28</td>
<td>$F_{A,SI}(k, j, s)$</td>
<td>Regenerator</td>
<td>Sink</td>
<td>Equal to the sink flow $F_{SI}(j,s)$</td>
</tr>
<tr>
<td>29</td>
<td>$F_{A,B}(k, k', s)$</td>
<td>Regenerator</td>
<td>Other regenerator</td>
<td>Maximum of the total source flows except freshwater and total sink flows</td>
</tr>
<tr>
<td>30</td>
<td>$\sum_{i \in I} \sum_{j \in J} F_{SO,SI}(i, j, s) + F_{A,SI}(k, j, s)$</td>
<td>Source and regenerator</td>
<td>Discharge</td>
<td>Equal to total source flows $\sum F_{SO}(i,s)$</td>
</tr>
<tr>
<td>31</td>
<td>$F_{FW}(s)$</td>
<td>Freshwater</td>
<td>Regenerators and sinks</td>
<td>Equal to total sink flows except discharge</td>
</tr>
<tr>
<td>32</td>
<td>$F_{W}(s)$</td>
<td>Sources and regenerators</td>
<td>Discharge</td>
<td>Equal to total source flows except freshwater</td>
</tr>
</tbody>
</table>

$$F_{SO,SI}^U(i, j, s) = \min_{\forall i \in I, \forall j \in J, \forall s \in S} (F_{SO}(i, s), F_{SI}(j, s)), \ \forall i \in I, \forall j \in J, \forall s \in S. \quad (35)$$

Note that the right-hand-side minimization operation involves computing the minimum value over the stipulated controlling indices of $i, j,$ and $s.$

$$F_{SO,A}^U(i, k, s) = F_{SO}(i, s), \ \forall (k, A) \in (K_{NM}, NM) \cup (K_{M}, MP) \cup (K_{M}, MR), \forall i \in I, \forall s \in S, \quad (36)$$
\[ F_{A,SI}^U (k, j, s) = F_{SI}^U (j, s), \quad \forall (k, A) \in (K_{NM}, NM) \cup (K_M, MP) \cup (K_M, MR), \forall j \in J, \forall s \in S, \tag{37} \]

\[ F_{A,B}^U (k, k', s) = \max \left( \sum_{i \in I} F_{SO}^U (i, s), \sum_{j \in J} F_{SI}^U (j, s) \right), \]
\[ \forall (k, A) \in (K_{NM}, NM) \cup (K_M, MP) \cup (K_M, MR), \tag{38} \]

\[ \forall (k', B) \in (K_{NM}, NM) \cup (K_M, MP) \cup (K_M, MR), \]
\[ k \neq k', \forall s \in S, \]

\[ \sum_{i \in I} \sum_{k \neq k'} \left( F_{SO,SI}^U (i, j, s) + F_{A,SI}^U (k, j, s) \right) = \sum_{i \in I} F_{SO}^U (i, s), \quad j \in \{ \text{Discharge} \}, \forall s \in S, \tag{39} \]

\[ F_{FW}^U (s) = \sum_{j \in J \setminus \{ \text{Discharge} \}} F_{SI}^U (j, s), \quad \forall s \in S, \tag{40} \]

\[ F_{W}^U (s) = \sum_{i \in I} F_{SO}^U (i, s), \quad \forall s \in S. \tag{41} \]

The concentration variables are generally bounded from above by the maximum source concentration for the corresponding contaminant \( q \):

\[ C_{A}^{FU} (k, q, s) = \max_{i \in I} C_{SO} (i, q), \quad \forall (k, A) \in (K_{NM}, NM) \cup (K_M, M), \forall q \in Q, \forall s \in S, \tag{42} \]

\[ C_{NM}^{U} (k, q, s) = \max_{i \in I} C_{SO} (i, q), \quad \forall k \in K, \forall q \in Q, \forall s \in S, \tag{43} \]
\[ C_{MP}^{U}(k,q,s) = (1 - R(k,q,s)) \max_{\forall i \in I} C_{SO}(i,q), \quad \forall k \in K, \forall q \in Q, \forall s \in S, \quad (44) \]

\[ C_{MR}^{U}(k,q,s) = R(k,q,s) \max_{\forall i \in I} C_{SO}(i,q), \quad \forall k \in K, \forall q \in Q, \forall s \in S. \quad (45) \]

### 4.2.3 Heuristic-Based Cuts

We append the heuristic-based logic cuts introduced in Khor et al. (2012b) as additional constraints into the model to increase the convergence. These linear logical constraints are derived generically based on design and structural specifications of water network synthesis problems and may be applicable to related process synthesis problems.

Importantly, the model given by the objective function in equation (30) subject to constraints (9)–(31) corresponds to a risk-neutral formulation. In the next section, we present an extension of the formulation to handle risk management.

### 5 Computational Results

A numerical example is presented that is inspired from a real-life case study of a petroleum refinery water network in Malaysia. Seven water-using operations are involved that comprise three sources as well as freshwater and three sinks. Two types of water regeneration technologies are considered, namely a cartridge (carbon) filter and a single-stage reverse osmosis network (RON) regenerator with uncertain removal ratios for two contaminants, i.e., oil and grease.
(O&G) and total suspended solids (TSS). Data for the problem is given in Tables 2 to 5, which are invariant over all scenarios except for the freshwater and discharge flows, as indicated.

Table 3. Data on water sources

<table>
<thead>
<tr>
<th>Water Source $i$</th>
<th>Flowrate $F_{SO} \text{ (ton/h)}$</th>
<th>Concentration $C_{SO}$ (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Oil and Grease</td>
</tr>
<tr>
<td>Process area</td>
<td>23.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Blowdown losses</td>
<td>3.500</td>
<td>1.00</td>
</tr>
<tr>
<td>Sidestream filter backwash losses</td>
<td>1.800</td>
<td>1.00</td>
</tr>
<tr>
<td>Freshwater</td>
<td>(variable)</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Table 4. Data on water sinks

<table>
<thead>
<tr>
<th>Water Sink $j$</th>
<th>Flowrate (ton/h)</th>
<th>Maximum Allowable Inlet Concentration (MAIC) (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Oil and Grease</td>
</tr>
<tr>
<td>Cooling tower</td>
<td>25.60</td>
<td>25.00</td>
</tr>
<tr>
<td>Boiler</td>
<td>115.0</td>
<td>0.10</td>
</tr>
<tr>
<td>Desalter</td>
<td>28.40</td>
<td>25.00</td>
</tr>
<tr>
<td>Discharge</td>
<td>(variable)</td>
<td>10.00</td>
</tr>
</tbody>
</table>

Table 5. Data on cartridge filter regenerator

<table>
<thead>
<tr>
<th>Contaminant</th>
<th>Removal ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil and grease</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total suspended solids</td>
<td>0.63</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Data on reverse osmosis network regenerator

<table>
<thead>
<tr>
<th>Contaminant</th>
<th>Removal ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil and grease</td>
<td>0.45</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Total suspended solids</td>
<td>0.97</td>
<td>0.98</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Economic data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual operating time $H$</td>
<td>8,760 h/y</td>
</tr>
<tr>
<td>Unit cost for freshwater $C_{FW}$</td>
<td>$1.00/t</td>
</tr>
<tr>
<td>Unit cost for effluent treatment $C_W$</td>
<td>$1.00/t</td>
</tr>
<tr>
<td>Annualized capital charge factor $\delta$</td>
<td>0.1/y</td>
</tr>
<tr>
<td>Capital cost coefficient for purchasing individual piping $b_{A,B}(\alpha, \beta)$</td>
<td>$100$</td>
</tr>
<tr>
<td>Parameter</td>
<td>Value</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Capital cost coefficient for installing individual piping (d_{A,B}(\alpha, \beta))</td>
<td>$6</td>
</tr>
<tr>
<td>Operating cost coefficient for pumping water in piping (c_{A,B}(\alpha, \beta))</td>
<td>$0.006/t</td>
</tr>
</tbody>
</table>

5.1 Scenario Generation

In actual practical situations, only a subset of the uncertain parameters is random. In this work, as alluded to earlier, we consider uncertainty in the left-hand side constraint coefficients of each of the removal ratios \(R(k, q, s)\) for a contaminant \(q\) of a regenerator \(k\) in a scenario \(s\). We apply a Monte Carlo-based pseudorandom number generation technique for the uncertain \(R\) values that involves sampling of finitely many, mutually exclusive scenarios from a discrete distribution developed based on historical data analysis. Each scenario is assigned an equiprobability value for occurrence of \(p\), equals to 1/NS.

5.2 Risk-Neutral Model Solution

We solve the problems using a state-of-the-art commercial global solver, BARON 11.9.1 (Tawarmalani and Sahinidis, 2005) and a new commercial global MIQCQP solver, GloMIQO 2.1 (Misener and Floudas, 2012), both from GAMS 24.0.2. It is noteworthy that the purpose here is not to compare the performance of the two solvers, but rather to verify the solution obtained from one solver against the other in an independent manner.

Table 6 summarizes the model size and solution statistics in solving the risk-neutral model. Both BARON and GloMIQO may incorporate redundant constraints derived from the reformulation–linearization technique (RLT) (Sherali and Alameddine, 1992) to enhance computational...
procedure. The resulting network are then shown in Figure 5 and Figure 6 for solutions obtained using BARON and GloMIQO, respectively, for a relative optimality tolerance $\epsilon_r$ of $10^{-3}$ (0.1%).

Table 8: Model size and computational statistics for the risk-neutral model with 100 scenarios

<table>
<thead>
<tr>
<th>Computing platform</th>
<th>GAMS 24.0.2 on computing cluster with 70 nodes mostly on 12-core 3.47 GHz Intel® Xeon® X5690, 4–128 GB of RAM (For more information, refer to: wiki.ce.ic.ac.uk/tiki-index.php?page=The+Linux+Cluster (last accessed on 12 July 2013))</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of continuous variables</td>
<td>2498</td>
</tr>
<tr>
<td>No. of 0–1 variables</td>
<td>28</td>
</tr>
<tr>
<td>No. of constraints</td>
<td>2207 (plus any RLT equation)</td>
</tr>
<tr>
<td>No. of nonconvex bilinear terms</td>
<td>900</td>
</tr>
</tbody>
</table>

BARON ($\epsilon_r = 10^{-3}$): Total cost = $1,245,899/year (380.00 s)

Figure 5. $\varepsilon$-Optimal risk-neutral network computed by BARON for $\epsilon_r = 10^{-3}$
The results indicate that $\varepsilon_r = 10^{-3}$ is not suitable because the optimal solutions reported are not consistent. Moreover, the optimum network obtained with BARON involve impractical reuse of a regenerated RO rejector flow to supply the high purity requirement (i.e., with low contaminant concentrations) of boiler feedwater (see Figure 5). Hence, there is a need to consider tighter tolerance to provide a certificate of global optimality (e.g., $\varepsilon_r = 10^{-5}$ or $\varepsilon_r = 10^{-6}$, instead of $\varepsilon_r = 10^{-3}$). In Figure 7, we present result for $\varepsilon_r = 10^{-5}$ as obtained using GloMIQO since it generates a solution faster than BARON—interestingly, an optimal network suggests reuse of the permeator to supply the boiler feedwater.
5.3 Risk-Averse Model Solution

For the purpose of prototyping, we employ manual intervention to estimate the VaR parameter from a CDF by plotting cumulative probability against the objective function value for each scenario (i.e., using the optimal risk-neutral solution). As shown in Figure 8, this involves a scenario-ordering procedure of sorting the values in ascending order to generate a CDF plot. The plot is utilized to determine VaR at a chosen confidence level of 95% (i.e., at the corresponding 5% cumulative probability). It is noteworthy that this function evaluation procedure of CDF estimation for simulating VaR can be automated as necessary. While in principle, these are continuous curves as based on their integral-based definitions, they are presented here as discrete.
curves because the uncertain parameters are approximated through discretizing the parameter values to form the scenarios.

Figure 8. Cumulative distribution function (CDF) for determining VaR

In the final step, the risk-averse program with a mean–CVaR objective is solved to global optimality by accounting for a larger and more representative 100 scenarios. Table 9 summarizes the main parameters while Table 10 reports the model size and computational statistics.

Table 9. Parameters for risk-averse model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of scenarios</td>
<td>100</td>
</tr>
<tr>
<td>Confidence level β</td>
<td>0.95</td>
</tr>
<tr>
<td>VaR$_{0.05}$</td>
<td>$1,246,028$/year</td>
</tr>
</tbody>
</table>
Table 10: Model size and computational statistics for risk-averse model

<table>
<thead>
<tr>
<th>Computing platform</th>
<th>GAMS 23.8.1 on Windows 7 on Sony Vaio laptop with Intel Core i7-620M processor 2.66 GHz (with Turbo Boost up to 3.33 GHz) and 8 GB of RAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of continuous variables</td>
<td>4357</td>
</tr>
<tr>
<td>No. of 0–1 variables</td>
<td>28</td>
</tr>
<tr>
<td>No. of constraints</td>
<td>4948 and any RLT equation</td>
</tr>
<tr>
<td>No. of nonconvex bilinear terms</td>
<td>1804</td>
</tr>
</tbody>
</table>

Figure 9 and Figure 10 show the optimum network determined by GloMIQO for risk factors \( \theta = 1 \) and \( \theta = 5 \), respectively for \( \varepsilon_r = 10^{-5} \); similar results generally apply to other risk factors.

GloMIQO (\( \varepsilon_r = 10^{-5} ; \beta = 0.95 ; \theta = 1 \)): Total cost = $1,250,042/year (25.00 s)

Figure 9. \( \varepsilon \)-Optimal risk-averse network for risk factor \( \theta = 1 \) (\( \varepsilon_r = 10^{-5} \))
Finally, Figure 11 shows histograms illustrating the distributions of solutions for both the risk-neutral and the risk-averse models as computed by GloMIQO for each of the 100 equiprobable scenarios for the instance of $\theta = 1$. To investigate the incentive for considering risk, a comparison between the two model solutions reveal that while the expected total annualized cost (TAC) has increased for the risk-averse model, the risk of a higher or possibly very high TAC has been reduced compared to the risk-neutral model as shown in Figure 10 with similar trends observed for other risk factors. This result indicates that in accounting for risk, we gain in robustness (i.e., by reducing the risk of a very high total cost) but at the expense of a higher expected TAC, which is indeed intuitive (see also You et al. (2009)). Similar trends are observed for the results of other risk factors.
6 Concluding Remarks

The work in this article has presented an optimization-based framework for water network synthesis under uncertainty in contaminant removal ratios with risk management considerations. The proposed stepwise solution strategy involves solving risk-neutral and risk-averse formulations of recourse-based multiscenario two-stage stochastic MIQCQP. We first consider a risk-neutral model that offers a tractable approximate solution by using a small number of scenarios, which also consequently reduces the computational load in an ensuing scenario-ordering scheme for simulating the value-at-risk (VaR) parameter. Ultimately, these steps feed
into a risk-averse model that adopts a convex risk measure of conditional VaR (CVaR). We handle the model nonconvexities due to the presence of bilinear terms by employing GAMS 24.0.2/GloMIQO 1.0.0, a new global MIQCQP solver and verifying the solutions with GAMS 24.0.2/BARON 11.9.1, a state-of-the-art solver for the same problem class. Our computational experiments show that water network synthesis under uncertainty with risk management problems can be solved to global optimality in reasonable CPU time.

Acknowledgment

We sincerely thank Ruth Misener and Christodoulos A. Floudas for providing us with a free beta version of the GloMIQO license, as well as Michael R. Bussieck of GAMS Corporation for providing us with a free GAMS license that supports the GloMIQO library. Cheng Seong Khor is grateful for financial support from the Commonwealth Scholarship Plan, Centre for Process Systems Engineering (CPSE) of Imperial College London, Universiti Teknologi PETRONAS, and Ministry of Higher Education (MOHE) of Malaysia (Fundamental Research Grant Scheme (FRGS) number FRGS/1/10/TK/UTP/03/16).

Notations

Sets and Indices

$I$ set of sources $i$

$J$ set of sinks $j$

$K$ set of all types of regenerators $k$ where $K = K_M \cup K_{NM}$

$K_{NM}$ set of non-membrane separation-based regenerators $k$

$K_M$ set of membrane separation-based regenerators $k$

$Q$ set of contaminants $q$

Parameters
\( \text{SO} \), \( \text{M} \), concentration of contaminant \( q \) in outlet of source \( i \) in scenario \( s \) (mg/L)

\( \text{max} \), \( \text{M} \), maximum concentration of contaminant \( q \) at inlet to sink \( j \) (mg/L)

\( F_{\text{SO}}(i) \), \( \text{M} \), flow from outlet of source \( i \) (ton/h)

\( F_{\text{FW}}(s) \), \( \text{M} \), flow from outlet of freshwater source in scenario \( s \) (ton/h)

\( F_{\text{SI}}(j) \), \( \text{M} \), flow to inlet of sink \( j \) (ton/h)

\( F_{\text{W}}(s) \), \( \text{M} \), flow to inlet of discharge sink in scenario \( s \) (ton/h)

\( H \), \( \text{M} \), annual operating time of the water systems plant (hour/year)

\( R(k,q) \), \( \text{M} \), removal ratio of contaminant \( q \) in membrane regenerator \( k \) (dimensionless)

\( \lambda(k) \), \( \text{M} \), liquid-phase recovery factor of membrane regenerator \( k \) (dimensionless)

\( b_{\alpha,\beta} \), \( \text{M} \), capital cost for purchasing of piping interconnection between an origin entity \( A \) (as indexed by \( \alpha \)) and a destination entity \( B \) (as indexed by \( \beta \)) ($/ton)$

\( c_{\alpha,\beta} \), \( \text{M} \), operating cost of piping interconnection between an origin entity \( A \) (as indexed by \( \alpha \)) and a destination entity \( B \) (as indexed by \( \beta \)) ($/ton)$

\( d_{\alpha,\beta} \), \( \text{M} \), capital cost coefficient for installation of piping interconnection between an origin entity \( A \) (as indexed by \( \alpha \)) and a destination entity \( B \) (as indexed by \( \beta \)) ($/ton)$

**Continuous variables** (flows in t/h, concentrations in mg/L)

\( F_{\alpha,\beta}(s) \), \( \text{M} \), flow from an origin entity \( A \) (as indexed by \( \alpha \)) and a destination entity \( B \) (as indexed by \( \beta \)) in scenario \( s \)

\( F_{\alpha,\beta}^{\text{max}}(s) \), \( \text{M} \), maximum flow from an origin entity \( A \) (as indexed by \( \alpha \)) and a destination entity \( B \) (as indexed by \( \beta \)) in scenario \( s \)

\( F_{NM}(k,s) \), \( \text{M} \), flow of feed stream to a non-membrane regenerator \( k \) in scenario \( s \)

\( F_{M}(k,s) \), \( \text{M} \), flow of entry point to a membrane regenerator \( k \) in scenario \( s \)

\( C_{\alpha,q,s} \), \( \text{M} \), concentration of contaminant \( q \) in feed stream to non-membrane regenerator \( k \) (or entry point to membrane regenerator) in scenario \( s \)

\( C_{NM}(k,q,s) \), \( \text{M} \), concentration of contaminant \( q \) in outlet of non-membrane regenerator \( k \) in scenario \( s \)

\( C_{MP}(k,q,s) \), \( \text{M} \), concentration of contaminant \( q \) in outlet of permeator of membrane regenerator \( k \) in scenario \( s \)

\( C_{MR}(k,q,s) \), \( \text{M} \), concentration of contaminant \( q \) in outlet of rejector of membrane regenerator \( k \) in scenario \( s \)

**Binary variables**

\( y_{\alpha,\beta} \), \( \text{M} \), existence of interconnection from an origin entity \( A \) (as indexed by \( \alpha \)) and a destination entity \( B \) (as indexed by \( \beta \))
Subscripts

NM  index for non-membrane regenerators
MP  index for permeator of membrane regenerators
MR  index for rejector of membrane regenerators

References


