On cooling-water systems design for South African industry: two recent developments

Thokozani Majazi*† and Nongezile Nyathi*

Introduction

Inherently large quantities of energy and water are used in industrial chemical processes. Since the late 1970s, significant efforts have been put into devising systematic procedures for energy optimization in grassroots and retrofit design. One outcome was pinch analysis,1 which has been applied successfully to various industries worldwide. This development was followed by the introduction of similar techniques for water minimization in chemical processing industries from the late eighties to very recently.2–11 While this technology has not been fully refined, various companies around the world have started reap ing the benefits of its application. The main drawback of previous developments is that the objectives of mass and heat integration are treated in a dichotomous manner.

Work has also been done on cooling-water systems relating to the reliability of cooling towers,12–15 the optimum sizing of these towers,16 energy conservation,17 cooling-water treatment18 and associated operational issues. These activities have not considered the influence of the cooling-water network on the performance of the cooling tower. It is only recently that combined heat and mass integration have been studied. A graphical technique to maximize cooling-tower performance by minimizing water supply to the cooling-water network has, however, been applied.19 The latter involves all those operations that use cooling water, can be realized only by considering the system as a whole. Traditional approaches have focused separately on either the cooling tower or the operational network. Optimality, in the context of this paper, refers to minimum cooling-water flowrate to the network or maximum return temperature to the source of the cooling water (a cooling tower). Only systems with at least two cooling towers instead a single one are considered here, to highlight the complexity of a typical cooling-water system. The first exercise is based on a graphical technique in which targeting for the minimum cooling water precedes design of the cooling-water network to achieve the target. The second exercise uses mathematical modelling to optimize a superstructure that entails all possible topological arrangements of the cooling-water network. An industrial case study involving a South African explosives manufacturing plant is used to demonstrate the effectiveness of both techniques. Cooling-water savings of more than 20% were realized with modest capital investment.

This paper presents two recent developments in the targeting and design of cooling-water systems using process integration. The basis of this work is the observation that true optimization of any cooling-water system, comprising a cooling tower and a network of operations that use cooling water, can be realized only by considering the system as a whole. Traditional approaches have focused separately on either the cooling tower or the operational network. Optimality, in the context of this paper, refers to minimum cooling-water flowrate to the network or maximum return temperature to the source of the cooling water (a cooling tower). Only systems with at least two cooling towers instead a single one are considered here, to highlight the complexity of a typical cooling-water system. The first exercise is based on a graphical technique in which targeting for the minimum cooling water precedes design of the cooling-water network to achieve the target. The second exercise uses mathematical modelling to optimize a superstructure that entails all possible topological arrangements of the cooling-water network. An industrial case study involving a South African explosives manufacturing plant is used to demonstrate the effectiveness of both techniques. Cooling-water savings of more than 20% were realized with modest capital investment.

Problem statement

The problem addressed in this paper can be stated as follows. Given (i) a set of cooling towers with different supply temperatures, (ii) maximum design capacity for each cooling tower, (iii) a set of cooling waters operating with limited temperature requirements, and (iv) the duty requirements for each of the operations using cooling water, determine the minimum amount of cooling water required by the overall cooling-water network as well as the optimal amount of water needed for each of the cooling towers, without compromising their respective performances. The limiting temperatures refer to maximum inlet and outlet water temperatures for each of the operations that use cooling water. These are determined by the characteristics of each operation.

Targeting

Targeting involves the setting of cooling-water supplies, from the cooling tower, to the operational network. A cooling-water composite curve is constructed20 for single water sources by combining all individual profiles into a single curve within prescribed temperature intervals. However, there are two scenarios that can arise in the presence of multiple cooling-water sources. If these sources have very high design capacities and high maximum return temperatures, all the cooling water used will be provided by the source with the lowest supply temperature (primary source). Thus, the water-supply line will be similar to that for systems with a single source. This observation is demonstrated in Fig. 1, which shows that the flowrate for the cooling-water source with a lower supply temperature ($T_s$) is less than that with a higher temperature. Thus, less water will be required if it is provided solely from the primary source.
However, the lower flowrate is associated with a higher return temperature for the same heat duty (Fig. 1). It is, therefore, evident that in a situation where flowrates are unlimited and return temperature not critical, the source with the lowest temperature should always be the sole supplier. This renders the other supplies redundant. The targeting is thus performed, as reported in the literature, for single cooling-water sources, where the flowrate is minimized by ensuring that the water-supply line forms a ‘pinch’ with the limiting cooling-water composite curve.\(^{19}\) Any water supply beyond that which forms a pinch with the limiting composite curve does not satisfy the cooling requirements of the system.

In cases where the minimum water requirement for the cooling-water system is higher than the design capacity of the primary source, however, the cooling water from the hotter source has to be used to supplement the primary cooling-water supply. In this particular case, targeting is no longer straightforward. The steps then followed in targeting for the cooling-water supply are outlined below.

**Step 1: Primary cooling-water source**

Maximize the amount of cooling water from the primary source by adjusting the gradient of the cooling-water supply line until its value corresponds to the maximum flowrate. This step is illustrated in Fig. 2. At this point, supplementary water from sources at higher temperatures is required to meet the target for the entire cooling-water system.

**Step 2: Intermediate cooling-water sources**

All the subsequent water-supply lines represent the overall flowrate of the previous water sources and the one being targeted at that stage. The supply line for the intermediate cooling-water sources is set by either the maximum capacity of the source or the pinch point. If the maximum capacity is reached before a pinch is formed with the limiting cooling-water composite curve, then the water source with the next closest temperature is used to supplement the cooling requirement. However, if the pinch is formed before the maximum capacity is reached, the rest of the water sources need not be used (they are redundant) because the cooling requirement will be satisfied. This step is illustrated in Fig. 3. As shown in the figure, the gradient of the cooling-water supply decreases with the addition of an extra water source as it represents the sum of all the cooling-water flowrates from all the sources that have been used.

**Step 3: Cooling-water source with the highest supply temperature**

If a pinch is not formed while targeting for any of the intermediate cooling-water sources, the supply with the highest temperature has to be used. The overall amount of cooling water used by the entire system is set by minimizing the quantity from the last source. This is done by ensuring that the final supply line forms a pinch with the cooling-water composite curve. The flowrate obtained from this line represents the overall cooling-water requirement for the system. This is shown in Fig. 4. Once the target has been set using the graphical technique, the following mathematical model is then applied to describe the corresponding network design.

---

**Mathematical model**

The mathematical model for optimization of a cooling-water system, as presented in this paper, is based on the superstructure shown in Fig. 5. The explanation for the symbols used appears at the end of the paper under Nomenclature. Subscripts, instead of brackets, are used in the notation for clarity. The mathematical optimization formulation is derived from this superstructure by

---

**Fig. 1.** Cooling-water targets for cooling-water sources with very high capacities.

**Fig. 2.** Setting the target for the primary cooling-water source.

**Fig. 3.** Setting the targets for intermediate cooling-water sources.
writing mass and energy balances across the existing operations. The following two cases are considered.

**Case A**

An unspecified cooling-water return temperature to the cooling tower, without a dedicated source and sink for any operation using the water—that is, pre-mixing of streams from cooling towers and post-mixing of streams, from operations needing cooling, back to the cooling towers—is allowed.

**Case B**

Here, unspecified cooling-water return temperature to the cooling tower, with a dedicated source and sink for each water-consuming operation—that is, the pre- and post-mixing of streams—is not allowed.

The objective for both cases is the minimization of the overall cooling-water supply from all sources, without violating any of the design and operational constraints within the cooling-water network.

**Case A: mathematical constraints**

This case is subject to the following constraints. Constraint (1) stipulates that the total cooling-water supply is the sum of the cooling water from each cooling tower \( n \).

\[
CW = \sum_{n \in N} OS(n). \tag{1}
\]

Constraints (2) and (3) ensure that the inlet and outlet cooling-water flowrates of cooling tower \( n \) are equal.

\[
OS(n) = \sum_{i \in I} CS(n, i), \quad \forall n \in N \tag{2}
\]

\[
OS(n) = \sum_{i \in I} CR(i, n), \quad \forall n \in N \tag{3}
\]

The total flowrate to operation \( i \) that uses cooling water is the sum of the flowrates from all the cooling towers and reused water from all other operations \( i' \), as given by constraint (4).

\[
\sum_{n \in N} CS(n, i) + \sum_{i' \in I} FR(i', i) = F_{in}(i). \quad \forall i \in I \tag{4}
\]

The total flowrate from operation \( i \) is the sum of return water flowrate from that unit to each cooling tower \( n \) and reused water to all other operations \( i' \), as given in constraint (5).

\[
F_{out}(i) = \sum_{n \in N} CR(i, n) + \sum_{i' \in I} FR(i', i'). \quad \forall i \in I \tag{5}
\]

Constraint (6) gives the energy balance across operation \( i \).

\[
Q(i) = F_{in}(i) \cdot D_{in}(i) = F_{out}(i) \cdot D_{out}(i) - Q(i). \quad \forall i \in I \tag{6}
\]

Constraint (7) defines the inlet water temperature for operation \( i \).

\[
T_{in}(i) = \frac{\sum_{i' \in I} FR(i', i) T_{out}(i') + \sum_{n \in N} CS(n, i) T(n)}{F_{in}(i)}. \quad \forall i \in I \tag{7}
\]

Since water is conserved by each operation that uses cooling water, constraint (8) is necessary:

\[
F_{in}(i) = F_{out}(i). \quad \forall i \in I \tag{8}
\]

The following equation ensures that the cooling-tower design capacity is not violated:

\[
OS(n) \leq OS^U(n). \quad \forall n \in N \tag{9}
\]

Constraints (1)–(7) are sufficient for case A. However, it is noteworthy that constraints (6) and (7) entail bilinear terms, which are naturally nonconvex, thereby rendering the overall formulation a nonconvex nonlinear programming (NLP) problem. Fortunately, there exists a condition which allows exact linearization of these constraints, as elaborated below.

**Linearization**

A problem with a similar mathematical structure to that exhibited by constraints (1)–(9) has been examined, albeit in water utilization systems for wastewater minimization. They demonstrated that a water network in which each operation operates at maximum outlet concentration will always require minimum water and postulated this as the necessary condition for optimality. Similarly, it can be shown in this case that a cooling-water network in which every operation has maximum outlet water temperature will require minimum flowrate from the water supply. This therefore becomes the necessary condition for optimality in cooling-water systems obeying the constraints of case A. Imposing this condition on constraints (6) and (7) allows the overall model to be linearized without any loss of optimality. In essence, this guarantees a globally optimal solution.

This is achieved by substituting constraint (7) into (8) to elimi-
nate the bilinear term, $F_u(i)T_{\text{supply}}(i)$, and replacing the continuous variable $T_{\text{in}}(i)$ with the parameter $T_{\text{out}}^U(i)$. Thus, constraints (6) and (7) are replaced by constraint (10).

$$Q(i) + c_p \sum_{i \in I} F(i')T_{\text{out}}^U(i') + c_p \sum_{n \in N} CS(n,i)T(n) = \begin{cases} 0 & \forall i \in I \end{cases}$$

To ensure that the maximum inlet concentration is obeyed, the following constraint is necessary.

$$F_u(n) \leq F_{\text{in}}^U(i), \quad \forall i \in I$$

where,

$$F_{\text{in}}^U(i) = \frac{Q(i)}{c_p(T_{\text{in}}^U(i) - T_{\text{out}}^U(i))}. \quad \forall i \in I$$

The objective function for this scenario is the minimization of overall flowrate, $CW$, from cooling-water supply as shown in Equation (1).

Constraints (1)–(5) and (8)–(12) constitute an overall LP model for case A. The formulation presented can be readily extended to situations stipulated in the following case.

**Case B**

In this case, the return water temperatures for each water source are not specified. However, each operation that uses cooling water has a dedicated source and sink, i.e. pre-mixing of cooling water supply and post-splitting of cooling-water return are not allowed. The objective function remains the same, i.e. the total cooling-water supply from all sources has to be minimized. In this scenario, the same constraints used in case A are used to determine the minimum cooling-water supply. However, the following additional constraints are necessary to prevent supply cooling-water pre-mixing and return cooling-water post-splitting. Prevention of pre-mixing implies that no more than one water source can supply any operation to be cooled. It is noteworthy, however, that recycled and reused water is allowed to provide any additional required cooling water in any operation. On the other hand, prevention of post-splitting implies that any operation to be cooled can return water to no more than one cooling-water sink, with some of the return water allowed for reuse and recycling.

Prevention of pre-mixing is achieved by constraints (13) and (14). Constraint (13) stipulates that the supply flowrate from a particular source $n$ to a particular operation $i$ cannot exceed a maximum. Constraint (14) ensures that operation $i$ can be supplied with cooling water by a maximum of only one source $n$.

$$CS(n,i) \leq F_u^U(i)yr(n,i), \quad \forall i \in I, n \in N$$

$$\sum_{i \in I} yr(n,i) \leq 1, \quad \forall n \in N$$

Similarly, to prevent post-splitting of cooling-water return, constraints (15) and (16) are necessary. Constraint (16) states that the return flowrate from a particular operation $i$ to a particular cooling-tower $n$ cannot exceed the maximum water flowrate into operation $i$. Constraint (16) ensures that a particular operation $i$ can supply at most only one cooling tower.

$$CR(i,n) \leq F_{\text{in}}^U(i)yr(i,n), \quad \forall i \in I, n \in N$$

$$\sum_{n \in N} yr(i,n) \leq 1, \quad \forall i \in I$$

To ensure that the source and the sink are the same for a given water-using operation, as might be deemed necessary by the layout of the plant, constraints (17) and (18) are necessary.

$$yr(i,n) \leq yr(n,i) + (2 - yr(n,i) - yr(i,n)), \quad \forall i \in I, n \in N$$

$$yr(n,i) \geq yr(n,i) - (2 - yr(n,i) - yr(i,n)), \quad \forall i \in I, n \in N$$

Evidently, the model for case B, composed of Equations (1)–(5) and (8)–(16), is a mixed integer linear programming (MILP) model for which global optimality can be readily ensured for solvable models.

To demonstrate the application and effectiveness of the proposed mathematical formulation, an actual case study is presented below. All calculations were performed with a 1.2 GHz Pentium M processor with 512 MB RAM.

**Industrial case study**

The industrial case study presented here is based on a South African explosives (nitrates) manufacturing facility. The process involved is supplied by one counter-current, induced-draft cooling tower. The cooling-water network is comprised predominantly of heat exchangers with a parallel-series configuration as shown in Fig. 6. The limiting cooling-water data of the case study are shown in Table 1.

The manufacturing process requires 3900 t/h of cooling water with about 59% (2290 t/h) used by the absorption tower; 41% (1610 t/h) of the cooling water supplied by the cooling tower is used by the rest of the heat exchangers before it is returned to the cooling tower at a higher temperature. The temperature of water returned to the cooling tower is 34°C and the supply temperature is 24°C. Some of the cooling water is lost in the cooling tower through evaporation, windage and blowdown. Blowdown is necessary to prevent accumulation of nitrates and other salts in the system. Water lost is replaced by a freshwater supply.

In the assessment of the network, it is evident that the cooling water has two sources, i.e. the cooling tower (primary source) and the absorption tower (secondary source), as shown in Fig. 7. This system is unique in that the cooling water required by the secondary source is supplied by the primary source. The absorption tower is treated as a secondary source rather than one of the operations that uses cooling water because its water

<table>
<thead>
<tr>
<th>Heat exchanger</th>
<th>$T_{\text{supply}}$ (°C)</th>
<th>$T_{\text{target}}$ (°C)</th>
<th>Duty (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>28</td>
<td>10 700</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>44</td>
<td>16 700</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>42</td>
<td>13 500</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>29</td>
<td>300</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>44</td>
<td>1100</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>46</td>
<td>4400</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
<td>44</td>
<td>-2000</td>
</tr>
</tbody>
</table>
supply is always fixed. As a process requirement, the amount of cooling water needed by the absorption tower is fixed and cannot be manipulated. In this case, therefore, the objective is to maximize the amount of cooling water that is reused from the secondary source, while minimizing the amount of water from the primary source. This is the direct opposite of the case presented above. The analysis is, therefore, reversed as detailed in the following sections.

Graphical optimization: targeting

The case study analysed is a special system in that, to obtain an optimal solution, the amount of cooling water from the secondary source should be maximized instead of maximizing the water from the primary source as required in those systems discussed earlier. In this case, the secondary cooling-water source has a fixed water requirement and this is supplied by the primary source, as shown in Figs 6 and 7. To minimize the overall amount of cooling water required from the cooling tower, the water supplied directly from the tower to the operations to be cooled should be minimized, since the stream to the absorption tower is fixed. This can be achieved only by maximizing the amount of cooling water reused from the secondary cooling-water source, i.e. the absorption tower. The procedure for constructing the cooling-water supply line for this case is described below. Water using operation 7 is not included in the construction of the limiting cooling-water curve because in that unit, cooling water is used as a hot stream. This is characterized by a higher inlet water temperature than the exit temperature. The limiting cooling-water composite curve is constructed considering only the operations where the cooling water is used as a cold stream, which is heated from a lower to a higher temperature.

The amount of cooling water used from the primary source is minimized by allowing its supply line to form a pinch, with the composite curve at a temperature equal to or below that of the subsequent cooling water source. Figure 8 illustrates targeting when two cooling-water sources serve operations requiring cooling water. From this figure, it can be seen that as the amount of cooling water from the primary source is decreased, the water required from a secondary source is increased as shown by the arrows. The limit is set when the cooling-water supply line for the primary source forms a pinch with the limiting cooling-water composite curve. In Fig. 8, this occurs when the cooling-water supply line is exactly on top of the limiting cooling-water composite curve. Further reduction in primary cooling water would not be feasible as the composite curve forms the boundary between the feasible and infeasible regions. Targeting in this manner ensures that as much cooling water as possible is used from the secondary cooling-water source.

The cooling-water target set for the case study is indicated in Fig. 9, which shows that the supply line for the primary cooling-water source is directly superimposed on the composite curve. The overall target shown in Fig. 9 for part of the network is 1800 t/h. This includes the water obtained directly from the primary source (694 t/h) and part of that which is used in the secondary source (1106 t/h). The cooling water from the secondary source, that is not reused anywhere else in the network (1184 t/h), is recombined with water from the rest of the network before being returned to the cooling tower (primary source). Thus, the temperature shown at the end of the cooling-water supply line in this case does not reflect the return temperature to the cooling tower. The return temperature is obtained by calculating the temperature of the mixture of water from the secondary source (that is not reused anywhere else in the system) and that from the rest of the network. In this case, the final temperature is 37°C.

Mathematical optimization: design

The mathematical model was based on case A, with the absorption tower treated as an ordinary operation using cooling water at fixed flowrate. The results obtained were exactly equal to those using the graphical technique. The mathematical model entails 42 linear constraints and 61 continuous variables, i.e. a linear program (LP). The solution was obtained in 0.01 CPU seconds.

Results and discussion

The final cooling-water network design obtained for the case study is shown in Fig. 10. Table 2 shows the current cooling-water requirements as well as the reduced cooling-water
requirements obtained using the method developed.

Analysing the nitrates production facility using the new method results in a reduction of 23% in the amount of water for those operations requiring cooling as well as the overall need for cooling water. The return water temperature was slightly increased to 37°C.

This rise is insignificant compared to the benefit obtained from the reduction in the cooling-water flowrate, that is, the reduction in the utility costs of the process.

From Figs 6 and 10, it can be seen that no major changes are needed in the overall layout of the plant. A few additional water pipes only will have to be added. Considering the benefit of the reduced utility requirements, it would be worth making the necessary changes. However, the effects of an associated drop in pressure will have to be investigated before any decision is made.

Conclusion

We have presented graphical and mathematical techniques for minimizing the use of cooling water in situations where multiple water sources are involved. The main conclusion is that a true optimum can be attained only by considering all components of a cooling system holistically rather than discretely. Considering a single source with its dedicated set of operations to be cooled, in the presence of a similar arrangement within the cooling-water system, is likely to yield inferior results. Illustrating these techniques by means of a real-life case study has shown reduced cooling-water requirements of over 20%. This presents considerable opportunities for capital cost savings in grassroots and retrofit designs.

We thank the Water Research Commission for funding this research.

Nomenclature

Sets

\( I \) Operations using cooling water
\( N \) Cooling towers supplying the cooling-water network

Parameters

\( Q(i) \) Duty of cooled operation \( i \) (kW)
\( T(i) \) Cooling-water supply temperature from cooling tower \( i \) (°C)
\( OS(n) \) Maximum design capacity of cooling tower \( n \) (t/h)

Continuous variables

\( CS(n,i) \) Cooling-water supply to cooled operation \( i \) from cooling towers \( n \) (t/h)
\( CS(n,i) \) Cooling-water supply to cooled operation \( i \) from cooling towers \( n \) (t/h)
\( CR(n,i) \) Return cooling water from cooled operation \( i \) to cooling towers \( n \) (t/h)
\( FR(n,i) \) Cooling water reuse from cooled operation \( i \) to operation \( n \) (t/h)
\( FE(i) \) Total cooling water into cooled operation \( i \) including direct supply and reused water (t/h)
\( FE(i) \) Total cooling water into cooled operation \( i \) including direct supply and reused water (t/h)
\( Tiin(i) \) Inlet cooling-water temperature to cooled operation \( i \) (°C)
\( Tfout(i) \) Outlet cooling-water temperature from cooled operation \( i \) (°C)

Binary variables

\( yl(n,i) \) \( \{ \begin{align*}
1 & \text{ if there exists a direct cooling-water stream from source } n \text{ to operation } i \\
0 & \text{ otherwise}
\end{align*} \) 
\( yr(n,i) \) \( \{ \begin{align*}
1 & \text{ if there exists a return cooling-water stream from operation } i \text{ to source } n \\
0 & \text{ otherwise}
\end{align*} \)

Received 13 January 2005. Accepted 12 January 2007.


Fig. 10. Improved cooling-water network design for the case study.

Table 2. Current and improved cooling-water requirements.

<table>
<thead>
<tr>
<th>Heat exchanger</th>
<th>( T_{\text{supply}} ) (°C)</th>
<th>( T_{\text{return}} ) (°C)</th>
<th>Flowrate (t/h)</th>
<th>New flowrate (t/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>28</td>
<td>2290</td>
<td>2290</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>44</td>
<td>2390</td>
<td>895</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>42</td>
<td>1450</td>
<td>643</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>29</td>
<td>60</td>
<td>58</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>44</td>
<td>450</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>46</td>
<td>900</td>
<td>269</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
<td>34</td>
<td>1700</td>
<td>214</td>
</tr>
</tbody>
</table>