MISTA 2009

Optimal Pump Scheduling For Water Distribution Systems

Lindell Ormsbee • Srini Lingireddy • Don Chase

Abstract: Three different mathematical formulations for the scheduling of pumps to minimize energy consumption while still satisfying operational constraints are presented. Each formulation requires the solution of a highly nonlinear optimization problem that requires the iterative solution of an external numerical model for evaluating and satisfying the operational constraints. Two different algorithms for solving the problem are examined.

1 Introduction

Perhaps one of the most important, yet most widely overlooked components of the urban infrastructure in the United States, is the public water supply system. However, over the last several decades and in particular since 9/11, many utilities are beginning to move to the use of sophisticated computer technologies to not only provide critical real time information about the state of their systems, but also to improve their overall operations. One such area of focus has been in the optimal scheduling of pumps so as to minimize cost and improve operational conditions.

Researchers have been exploring the area of optimal pump operations for several decades [11]. In its most general form, the classical non-linear constrained optimization problem as applied to pump operations may be formulated as follows:

Minimize or Maximize: $$F(X : q, p) \tag{1}$$

Lindell Ormsbee
University of Kentucky
Email: lormsbee@engr.uky.edu

Srini Lingireddy
Fluid Hammer Consultancy Service
Email: Srini@fluidhammer.com

Donald Chase
University of Dayton
Donald.Chase@notes.udayton.edu

Subject to: $$q(X) = 0 \quad (2)$$

$$p(X) \geq 0 \quad (3)$$

$$X_{\min} < X < X_{\max} \quad (4)$$

where F(X:q,p) represents the objective function (in terms of cost) to be minimized, q(X) represents the explicit system constraints (conservation of mass and energy in terms of flow q), and p(X) represents the implicit bound constraints (minimum pressures) to be satisfied. Finally, X_{min} and X_{max} represent explicit bound constraints on the decision variables of the optimal control formulation (duration of pump operations). The system constraints can be represented explicitly by the use of a simplified inductive model of the water distribution system [15] or a deductive model [10]. The exact form of the objective function and the associated decisions variables depend upon the nature of the problem formulation.

2 System Constraints

Because of the complexity of the system constraints, they are normally extracted out of the formal optimization formation and handled externally through use of a simulation program as shown in Figure 1. When using this structure, a vector of decision variables X is selected that explicitly satisfies the explicit bound constraints. This vector is then passed to the simulation program where the system constraints are satisfied and a vector of resulting pipe flowrates q and junction pressures p are determined. Once determined these values are passed back to the optimization program in order to check for feasibility of the implicit bound constraints (i.e. minimum pressures - p), and if feasible, the resulting value of the objective function associated with the decision vector X.

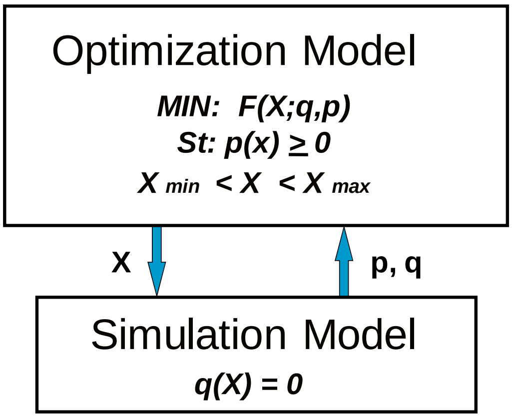

Figure 1. Problem Disaggregation Structure

The system constraints for this type of problem can be illustrated by considering the simple pipe distribution system shown in Figure 2. This system contains two tanks, five pipes, three junction nodes and one pump. The direction of flow in each of the pipes is indicated by the arrow on the line segment. In this case, a typical problem might be to determine the operating times of the pumps at the pump station over the course of an operating period (e.g. a day) so as to meet the varying demands M_t at junction node 3 while maintaining adequate operating pressures at each junction node (which is related to insuring that the water in the tank never drops below a critical level).

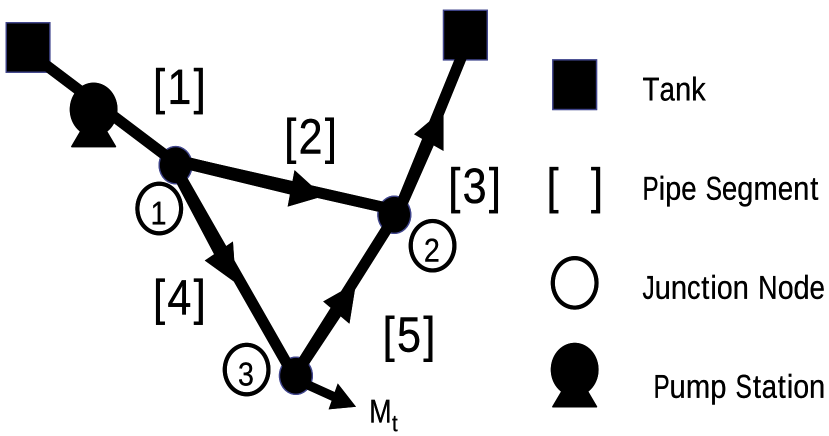

Figure 2. Example Pipe Network

The majority of pumps used in the water industry are constant speed pumps, in which the discharge is a function of the water levels in the tank. During those times in which the pump station discharge exceeds Mt, the excess will fill the tank. During those times in which the pump station discharge is less than Mt, then the tank will drain. The decision of what pumps to turn on and how long they should run will be driven by both the nodal demand Mt and the tank level (which controls the associated junction pressures). In many cases, pump stations will consist of different size pumps with different flow versus pressure relationships. Thus, the decision of which pump or combination of pumps to run during a particular period can become rather complicated. This is further complicated by the fact that the discharges of multiple pumps cannot be simply added together. Instead, the combined flows are a nonlinear function of the energy loss through the system (as determined by a system energy curve) and the associated tank level. For example, consider the situation of two pumps as shown in Figure 3. Each pump will have a different discharge versus pressure curve (known as a pump characteristic curve) as shown in the accompanying graph. The actual discharge that a particular pump will produce is determined by the intersection point of the characteristic curve and the system energy loss curve. When pump A is operating, the discharge is Qa. When pump B is operating the discharge is Qb. However, when both pumps are operating, the discharge is Qa+b. Technically, the system energy loss curve will shift up or down depending upon the water level in the tank, so things can become somewhat complicated.

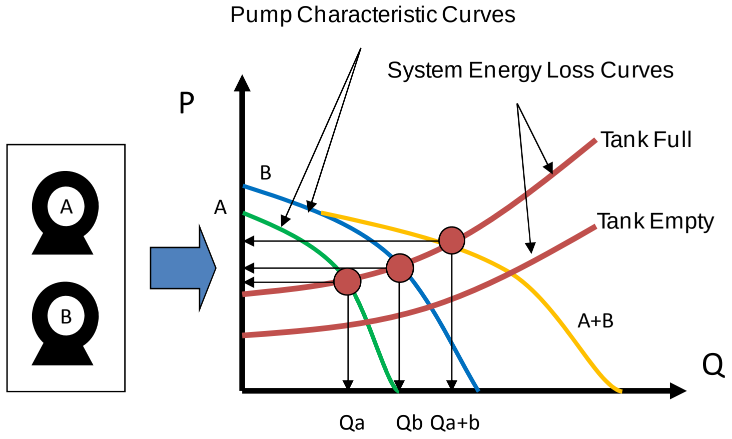

Figure 3. Pump Characteristic Curves and System Energy Loss Curves

The system energy loss curve represents a conceptual approximation of the flow versus pressure relationship that exists in a real water distribution network. In actual practice, this relationship must be determined by solving a series of nonlinear energy equations that are used to describe the physics of the system. For example, if we use the simple system shown in figure 2, then we are required to write one energy equation for each loop or energy path that exists in the network. An energy path is simply a path or series of pipes that connect any two tanks (technically any two points in the system with a known boundary condition – i.e. water level). For this simple system, the energy equation for the simple loop may be written as:

$$K_2Q_2^{1.852} - K_4Q_4^{1.852} - K_5Q_5^{1.852} = 0 \tag{5}$$

Likewise the energy equation for the path may be written as:

$$K_1Q_1^{1.852} - Z_P / Q_1 + K_2Q_2^{1.852} + K_3Q_3^{1.852} = \Delta E \tag{6}$$

where Q_i is the flowrate in pipe i, Z_p coefficient which is a function of the horsepower of the pump, ΔE is the difference in elevation between the water levels of the two tanks, and K_i is an energy loss term for each pipe that may be expressed as:

$$K_i = \frac{4.73L_i}{C_i^{1.852}D_i^{4.87}} \tag{7}$$

where L = the pipe length in feet, D = the pipe diameter in feet, and C = an empirical roughness coefficient that depends on pipe age and pipe material (e.g. C = 130 for a new ductile iron pipe).

Since equations 5 and 6 are nonlinear in terms of Q, they cannot be solved directly but they may be solved iteratively using a Taylor's series approximation. For our simple system, the resulting equations become:

$$1.852[K_2Q_2^{.852} - K_4Q_4^{.852} - K_5Q_5^{.852}]\Delta Q = K_2Q_2^{1.852} - K_4Q_4^{1.852} - K_5Q_5^{1.852} \tag{8}$$

$$[1.852K_1Q_1^{.852} + Z/Q_1^2 + 1.852K_2Q_2^{.852} + 1.852K_3Q_3^{.852}]\Delta Q =$$

$$= K_1Q_1^{1.852} - Zp/Q_1 + K_2Q_2^{1.852} + K_3Q_3^{1.852} + \Delta E \tag{9}$$

If we express equations 8 and 9 in matrix notation, then the set of equations can be written as:

$$\begin{bmatrix} G_I\,(\mathbf{Q}_I) & 0 \\ 0 & G_{II}\,(\mathbf{Q}_{II}) \end{bmatrix} \begin{bmatrix} \Delta Q_{\mathbf{I}} \\ \Delta Q_{\mathbf{II}} \end{bmatrix} = \begin{bmatrix} -F_I(\mathbf{Q}_I) \\ -F_{II}(\mathbf{Q}_{II}) \end{bmatrix}$$

Figure 4. Matrix structure for simple network example

where G_I and F_I represents the function on the right hand side and left hand side of equation 8 respectively, and Q_I represents the vector of flows. Likewise, G_{II} and F_{II} represent the functions on the right hand side and left hand side of equation 9 respectively, and Q_{II} represents the vector of flows. Reducing further, the equations can now be expressed as:

$$[G]\Delta Q = \{F\} \tag{10}$$

Thus, for this simple example, each time the simulation model is accessed by the optimization model, initial estimates of the flows in each pipe must be generated so that conservation of mass is preserved at each junction node (e.g. $Q_1 - Q_2 - Q_4 = 0$). Once this is done, the K coefficients are determined and substituted into the equations. Next, flow adjustment factors (i.e.ΔQ) are calculated for each loop or path as follows:

$$\Delta Q = [G]^{-1}\{F\} \tag{11}$$

Once these are determined, each of the individual flows in each pipe may be determined using a recursive equation. For example, for pipe 1 the flows are updated as follows:

$$Q_{1i+1} = Q_{1i} + \Delta Q_I \tag{12}$$

For pipes that are common to more than one equation (e.g. pipe 2), the recursive equation is expressed as follows:

$$Q_{2i+1} = Q_{2i} + \Delta Q_I + \Delta Q_{II} \tag{13}$$

Once all the flows have been updated, they are then used to update the K coefficients and the process is repeated until the ΔQ's converge to zero.

It should be recognized that most water distribution systems are much more complicated than the simple network shown in Figure 2. Further, in simulating the performance of such a system over time (e.g. a day), the previous solution methodology must be repeated for each time step in the simulation. In this case, the flows determined at the beginning of each time step are then used to update the boundary conditions for use in solving for the flows at the next time step. In order to minimize the error associated with this type of Euler integration, a maximum time step of one hour may be necessary. When applied to typical water distribution system, the computational requirements associated with each "function call" from the optimization model can become significant.

3 Objective Function

The objective of the extended period operation problem is to minimize the total energy consumption charges associated with operating a set of pumps over the course of an operating horizon, while simultaneously satisfying any service and reliability- related requirements of the system. In a typical water distribution system, the energy consumption costs incurred by pumping water depends mainly on the rate at which water is pumped, the associated pump head, the duration of pumping, the unit cost of electricity, and the different combined efficiencies of various pump combinations. Mathematically the objective function may be expressed as:

$$\text{Minimize } Z = \left(\frac{0.746\gamma}{550}\right)\sum_{t=1}^{T} R_t \sum_{i=1}^{I}\left[\frac{Q_{t,i}H_{t,i}X_{t,i}}{e_{t,i}}\right] \tag{14}$$

where: Z = the total energy cost to be minimized ($)

$Q_{t,i}$ = the average flowrate associated with pump i during time t (cfs)

$H_{t,i}$ = the average head associated with pump i during time t (ft)

note: the term "head" is a common term in pump hydraulics used to express pressure in terms of an equivalent height of water.

$X_{t,i}$ = the duration of the time pump i is operating during interval t (hr)

$e_{t,i}$ = the average wire to water efficiency associated with pump i during time t

R_t = the electric rate during time t ($/Kw-hr)

γ = the specific weight of water (lb/ft^3)

I = the total number of pumps included in the optimization

T = the total number of time intervals in the operating horizon

For a given network configuration and an associated set of initial boundary conditions (including the vector of initial tank levels E and the vector of system demand loadings M), the average discharge Q_i, pump head H_i, and pump efficiency e_i, associated with a particular pump i can be expressed as a function of the characteristics of the pump itself plus the characteristics of another pump which may be operating during the same time periods as pump i. Since the set of pumps operating during a particular period, t, can be explicitly defined by the duration of time each pump in the set is operating, (i.e. if $X_{t,i} = 0$ then pump i is not operating during time period t and if $X_{t,i} > 0$ then the pump is operating during period t), then the pump discharge, the pump head, and the pump efficiency can be expressed as implicit functions of the vector of total pump durations for a particular time interval [12]. As a result, equation 14 may now be expressed as:

$$\text{Min } Z = \sum_{t=1}^{T}\sum_{i=1}^{I} f\left[Q_{t,i}(\mathbf{X}_t, \mathbf{M}_t, \mathbf{E}), \mathbf{H}_{t,i}(\mathbf{X}_{t,s}, \mathbf{M}_t, \mathbf{E}_t), \mathbf{e}_{t,i}(\mathbf{X}_{t,s}, \mathbf{M}_{t,i}, \mathbf{E}_{t,i}), \mathbf{X}_{t,s}, \mathbf{R}_t\right] \tag{15}$$

As a result, the objective function can be expressed solely in terms of the vector of the individual pump operating times. As will be discussed in the following sections, the exact nature of the pump operating times will be dependent upon the problem formulation.

2 Implicit Scheduling Formulation

Historically, the optimal pump scheduling problem has been either formulated as an implicit control problem or an explicit control problem. In the implicit formulation, pump station discharge, supply pressure, or tank water levels are treated as the decision variables of the optimal control problem. In the explicit formulation, the actual pump operating times (either individually or compositely) are treated as the decision variables.

The implicit formulation will typically require the solution of two sub-problems. The first sub-problem involves determining an optimal decision trajectory. The optimal decision trajectory can be defined as that series of pump station discharges, supply pressures, or tank water levels which, over the course of the operating horizon, result in a minimal total operating cost.

The second sub-problem involves determining the specific pump operations which produce the optimal decision trajectory. The difficulty associated with finding specific pump combinations is compounded by the fact that many combinations capable of producing the desired decision trajectory may exist. Furthermore, from all the possible combinations capable of producing the desired trajectory, the combination which results in minimal operating costs must be found.

Historically, the implicit approach has been used to develop control strategies for single tank systems. In such cases, the tank water level has been used as the implicit decision variable and the resulting formulation is typically solved using dynamic programming [7], [8], [13], [16]. Due to the curse of dimensionality, dynamic programming solutions are normally restricted to problems involving no more than three storage tanks. For systems with several tanks, researchers have either employed dynamic programming along with an associated site specific spatial decomposition scheme [3], [21] or nonlinear programming along with an alternative implicit decision variable [9].

3 Discrete Explicit Pump Scheduling

In the discrete explicit approach, the actual times of the operation of each pump are treated as the decision variables [6], [14]. In this case, the objective function may be characterized in terms of operational costs and the associated state variables (such as flow or pressure) can then be modeled using a much more robust model of the water distribution system that can be linked to the optimization algorithm via iterative subroutine calls [19].

3.1 Restricted Formulation.

The restricted formulation of time as a decision variable was originally proposed by Chase and Ormsbee [5] and later applied by Chase [4] and Brion [2]. In this approach, a pump is forced to begin operating at the beginning of a pre-determined timer interval (e.g. every four hours), in effect, restricting the time a pump can begin operating. The decision variable in this case is the duration of time each pump operates during a particular time interval. For example, suppose decision variable $X_{3,2}$ has a value of 2.50 and a four hour time interval is used. Under the restricted approach, pump 2 would be turned on at the beginning of time interval #3 and turned off 2.5 hours into the interval. In other words if the operating horizon begins at midnight, pump 2 would be placed on line at 8:00 a.m. and turned off at 10:30 a.m. The values taken on by the decision variables are bounded between 0 (pump off) and Δt (duration of time interval). Figure 5 provides a pictorial representation of the decision variables in the restricted approach.

As can be seen from Figure 5, the total number of decision variables required to solve the optimal control problem is equal to the product of the number of pumps and the number of time intervals making up the operating horizon. As the size of the time interval is decreased, the closer the problem becomes to one of continuous pump operation. Unfortunately a decrease in the size of the time interval results in a proportional increase in the number of time intervals which, in turn, causes an increase in the number of decision variables. An increase in the number of decision variables can be undesirable since the computation time required to solve the optimal control problem will increase.

3.2 Unrestricted Formulation.

In an effort to reduce the number of decision variables, yet continue to pose the problem within the framework of an explicit time model, the restricted formulation can be modified to yield an unrestricted formulation. In the unrestricted formulation, the decision variables become the starting and ending time for the pumps. Figure 6 shows a pictorial representation of the decision variables for the unrestricted formulation.

Unlike the restricted formulation, in the unrestricted formulation there are fewer conditions as to when a pump may start or stop. Instead of forcing a pump to begin operating at the start of a particular time interval, pumps are allowed to start operating at any time during the operating horizon. Likewise, instead of shutting down a pump during a given time interval perhaps only to have it operating again at the start of the next time step, under the unrestricted approach a pump may stop operating at any time during the operating horizon. Consequently, feasible values of the decision variables for the unrestricted approach are bounded between 0 (pump off) and T where T is the operating horizon, usually 24 hours.

3.3 Discussion of Unrestricted and Restricted Formulations

A nice feature of the unrestricted approach is that it more closely replicates actual pump operation. In other words, system operators will typically place pumps on line in response to certain system parameters such as a low tank level, low pressure, an increase in system demands, a fire, etc. Such changes in system indices and/or system loadings are not likely to occur at the start of a pre-defined time interval, unless of course, the time interval is extremely small.

The restricted formulation allows pumps to operate several times per day. This can be accommodated in the unrestricted formulation by increasing the number of decision variables. In this case, a pair of decision variables are associated with each pump duty cycle. If a pump is allowed to operate only once per day, then there is one pump duty cycle per day and the number of decision variables is twice the number of pumps, i.e. one decision variable for the starting time and one for the ending time. If all pumps are allowed to operate twice per day, then there are two pump duty cycles for each pump. Following the convention that a pair of decision variables is associated with each pump duty cycle, then the total number of decision variables would be four times the number of pumps.

In general, the number of decision variables associated with the unrestricted formulation can be found using the following equation:

$$\mathrm{NDV} = \sum_{i=1}^{I}(2 * DC_i) \tag{16}$$

where NDV is the total number of decision variables in the unrestricted explicit formulation, DC_i, is the number of daily duty cycles allowed for pump i, and I is the number of pumps included in the optimization.

With the unrestricted formulation, additional bound constraints are needed to assure that the starting time for a pump is less than the stopping time for the same pump. If multiple duty cycles are allowed, then constraints are needed to insure that the starting time for particular pump duty cycle is greater than the stopping time for the pump's previous duty cycle. The additional constraints required for the unrestricted formulation can be expressed with the following equations:

$$\mathbf{X}_{i+I} - \mathbf{X}_{i} > 0 \quad \text{i= 1, 2…..I} \tag{17}$$

$$\mathbf{X}_{n,I} - \mathbf{X}_{n-1,i+I} < 0 \quad \text{i= 1, 2……I; n= 2, 3…..DC} \tag{18}$$

where $\mathbf{X}_{i}$ is the starting time for pump i, $\mathbf{X}_{i+1}$ is the stopping time for pump i, $\mathbf{X}_{n,i}$ is the starting time for pump I for the nth duty cycle, and $\mathbf{X}_{ni+I}$ is the stopping time for pump i during the nth duty cycle. Such constraints are not necessary for the restricted approach since the decision variable is the duration of time a pump operates and since the formulation keeps pump operation within a single time interval.

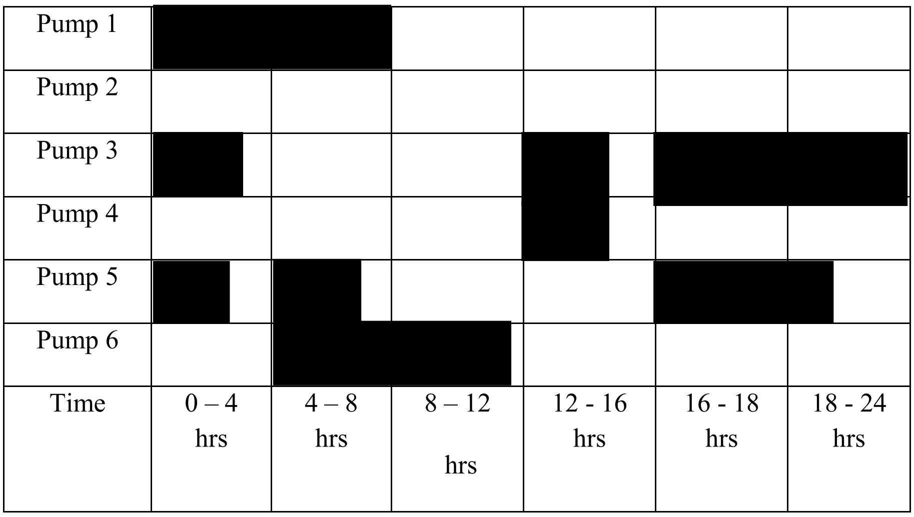

Figure 5 Schematic Representation of Decision Variables for the Restricted Approach

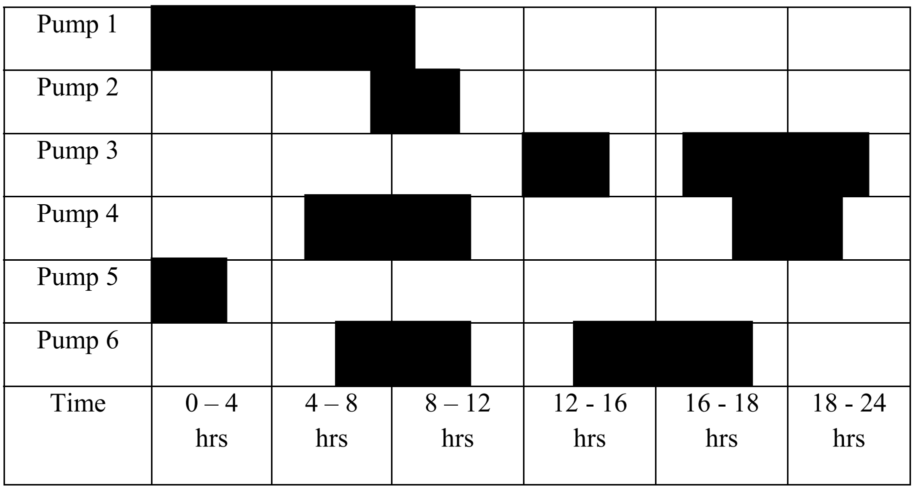

Figure 6. Schematic Representation of Decision Variables for the Unrestricted Approach

Two separate explicit control formulations have been proposed: 1) a restricted formulation and 2) an unrestricted formulation. Under the restricted formulation, pumps are forced to begin operating at the start of a time step and the value of the decision variable is equal to the duration of time the pump operates. Under the unrestricted formulation, pumps are allowed to start and stop operating at any time within the operating horizon and the decision variables represent the actual starting and stopping times for each pump duty cycle. With respect to the restricted formulation, the unrestricted approach allows more freedom of operation and more closely resembles actual pump operation. However, the unrestricted approach requires the use of additional constraints to regulate pump operation within practical limits.

As a result, the addition of new constraints results in a slight modification of the optimal control problem originally shown in equations 1-4. The modified optimal control problem for the explicit formulation may now be expressed as:

Minimize or Maximize: $$F(X) \tag{19}$$

Subject to: $$G_j(X) \leq 0 \qquad j= 1, 2, 3....J \tag{20}$$

$$H_k(X) = 0 \qquad k= 1, 2, 3....K \tag{21}$$

$$U_m(X) \leq 0 \qquad m= 1, 2, 3....M \tag{22}$$

$$X_{\min} \leq X \leq X_{\max} \tag{23}$$

where all the terms are the same as before and U(X) represent bound constraints used to keep pump starting and stopping times within practical limits when using the unrestricted approach. If constraints (20) – (23) are included directly in the optimization problem as binding constraints, then the result for most real problems would be a large scale nonlinear problem. The scale of the problem can be reduced by handling the constraints in the following manner. The implicit system constraints, H(X), are solved through the use of a simulation model [20]. Depending upon the size of the mathematical model describing the water distribution systems, accommodating the implicit constraints through the use of a simulation model can substantially reduce the number of constraints which would otherwise be solved by the optimization algorithm. The implicit bound constraints, G(X), and the bound constraints associated with the unrestricted approach, U(X), can be included in the objective function as a penalty term so as to allow solution using an unconstrained optimization method. Alternatively, depending upon the type of solution algorithm employed, they may also be explicitly satisfied outside of the algorithm. Finally, the explicit bound constraints can be used to assign values of decision variables when the variables exceed their bounds. For example, if during the course of the optimization a value taken on by a decision variable is greater than the upper bound of the variable, X_{max}, then the value of the decision variable is set equal to the upper bound. Similarly if a value of a decision variable is less than the lower bound, X_{min}, then the value of that variable is set equal to its lower bound.

4 Composite Explicit Pump Scheduling

As discussed previously, the exact form of the decision variable will be dependent upon whether a restricted or an unrestricted formulation is employed. In applying the restricted approach, the normal operating horizon (typically 24 hours) is divided in to T separate time intervals (e.g. 4 hours) and the pump operating time for each pump in each time interval is determined. In the unrestricted approach, a specific number of pump duty cycles is specified for each pump and the beginning and ending times of each duty cycle is determined. For distribution systems with multiple pump stations, and with each pump station containing numerous pumps, both formulations can result in an excessive number of decision variables. One way to significantly reduce the total number of decision variables would be to develop a single decision variable for each pump station for each time interval that relates the particular set of pumps in operation during that period. Such a formulation can be obtained by ordering the various available pump combinations associated with each pump station on the basis of unit cost. A single decision variable can then be developed for each pump station s and each time interval t of the form X_{st} = II.CC where II = is an integer and corresponds to the identification number of the pump combination that operates CC% percent of the time interval. By definition, it will also be assumed that combination II-1 operates the remaining (1-CC) percent of the time interval. (The combination in which II = 0 corresponds to the null combination or the decision to run no pumps). Modification of the original objective function to accommodate the proposed formulation yields the following objective function:

$$\text{Min}\, Z = \sum_{t=1}^{T} \sum_{i \in s}^{I_s} f\left[Q_{t,i}(\mathbf{X}_{t,s}, \mathbf{M}_t, \mathbf{E}_t), \mathbf{H}_{t,i}(\mathbf{X}_{t,s}, \mathbf{M}_t, \mathbf{E}_t), \mathbf{e}_{t,i}(\mathbf{X}_{t,s}, \mathbf{M}_{t,i}, \mathbf{E}_{t,i}), \mathbf{X}_{t,s}, \mathbf{R}_t \right] \quad (24)$$

The objective function as expressed in Eq. 24 is subject to the same three sets of constraints as with the previous explicit formulations. These include: (1) a set of implicit system constraints, (2) a set of implicit bound constraints, and (3) a set of explicit decision variable constraints. While both the implicit system and bound constraints will have an identical form as before, the explicit decision variable constraints are different for the new formulation as a result of the use of a new set of decision variables. In this case the decision variable for each pump station for a particular time interval will be restricted between a lower value of zero (corresponding to no pumps in operation) and an upper value related to the maximum number of pump combinations available for that pump station.

In applying the proposed algorithm to a specific distribution system, the desired operating horizon (typically 24 hours) is once again divided into a discrete set of time intervals. A separate decision variable for each time interval is then assigned to each pump station. To initiate the algorithm a separate vector for each time interval is randomly generated or explicitly specified which contains the values of the decision variables for each pump station in the system. As a result, any potential solution will consist of a set of N vectors where N – the number of time intervals which constitute the operating horizon. To insure a feasible solution, the initially specified or generated set of decision vectors must satisfy the explicit bound constraints.

Similar to the previous explicit control formulations, the proposed composite formulation also uses a disaggregated solution methodology. That is, once an initial set of decision vectors is obtained, it is then passed down to a network simulation model [20] for use in explicitly satisfying the implicit system constraints and for use in evaluating the implicit bound constraints. The values of the resulting state variables (i.e. flowrate, pressure, kilowatt consumption, etc.) are then passed back to the optimization algorithm for use in quantifying the objective function and identifying any violations in the implicit bound constraints. This information is then used to generate an improved set of decision vectors which automatically satisfies the explicit bound constraints and which seeks to minimize the objective function.

Once generated, the improved set of decision vectors is then passed back down to the simulation algorithm for subsequent evaluation. This process is then repeated until a specified level of algorithmic convergence is obtained.

To permit increased model flexibility, the computational time interval used in the simulation model is not restricted to be equal to the interval associated with the decision variables. For example, if a 12 hour time interval is used for the decision variables a much smaller time step (i.e. 2 hour, 4 hour, etc.) can be used as the time interval in the simulation model.

5. Solution Methodologies

Various optimization methods have been employed by researchers in solving the discrete explicit pump scheduling problem. These methods have varied from traditional gradient based methods to more exotic evolutionary methods (e.g. genetic algorithms). More recently Tufail [18] and Tufail and Ormsbee [19] have proposed using a modified and improved version of a relatively straightforward direct search method [1]. The approach, called the Shuffled Box Complex Method, has a potential advantage over genetic algorithms in that both implicit and explicit constraints may be handled directly without the use of a penalty formulation. Similar to genetic algorithms, their method also pursues an optimal solution along multiple simultaneous search paths, thereby improving the efficiency of the original method.

While the composite explicit pump scheduling problem can also be solved using the same algorithms discussed above, Ormsbee and Reddy [12] found that the unique nature of the problem yields itself to solution via a simple heuristic. Because of the nature of the selected decision variable, the least cost solution to the unconstrained objective function (i.e. Eq. 24) is explicitly known. That is, the least cost solution is one in which no pumps are operated. As a result, the optimization problem reduces to one of finding the set of decision vectors which produce solutions on the composite constraint boundaries that are as close to the origin of the n dimensional solution space as possible. For a given initial set of decision vectors an improved set may be obtained by simply contracting the scalar values in the vectors toward the origin. In the event that a contraction results in a constraint violation, then the vector that produced the constraint violation can be subsequently expanded until a feasible solution is obtained. By continuing to bisect the resulting search vector, a set of decision variables can be obtained which will result in a solution which lies just on the constraint boundary. In the event the initial solution violates an implicit bound constraint, then the solution may be expanded away from the origin along the resulting search direction until a feasible solution is obtained. Due to the nature of the explicit bound constraints, a feasible solution will eventually be found.

It should be recognized that application of such an approach will only result in the best solution that lies on the search direction located between the given initial solution and the origin of the dimensional solution space. However, additional feasible solutions may be obtained by simply replicating the methodology using additional sets of initial decision vectors. Once the final set of feasible solutions is obtained an "optimal" solution may be obtained by simply selecting the best solution from among the resulting feasible solutions.

More recently, Teegavarapu et al. [17] combined the general search heuristic with a genetic algorithm (GA) by using the GA to continue to generate a new set of potentially feasible solutions. Once again, the algorithm is restricted to that class of problems where the unconstrained optimal solution is known (e.g. where all the pump operating times are zero). In this case, once the new population of solutions is generated, each solution is checked for feasibility. Those solutions that are not feasible are projected up to the constraint boundary from below while those solutions that are feasible are now projected down to the constraint boundary from above (Figure 7). This will then insure that each offspring in the current population is both feasible and locally optimal (e.g. relative to the vector that originates from the origin of the solution space). Once these solutions are generated, the GA is now engaged and a new population is generated (Figure 8). Once again, some of these solutions will be feasible and some will be infeasible. As before, both sets of solutions are projected to the constraint boundary, and the process is continued (Figure 9). Eventually, the GA will drive the solutions to that region on the constraint boundary that is closest to the origin of the solution space and thereby yield an optimal solution. (Figure 10).

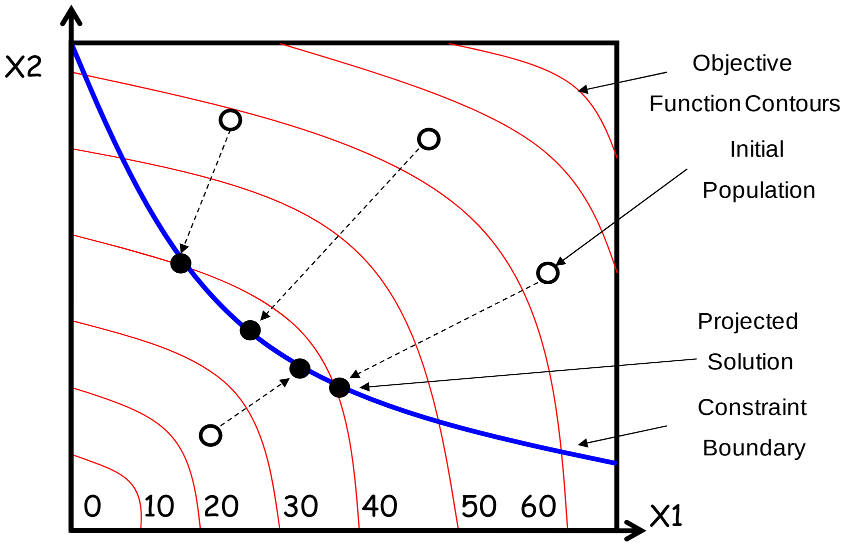

Figure 7. Projection of Initial Solutions to Constraint Boundary (2 dimensional example)

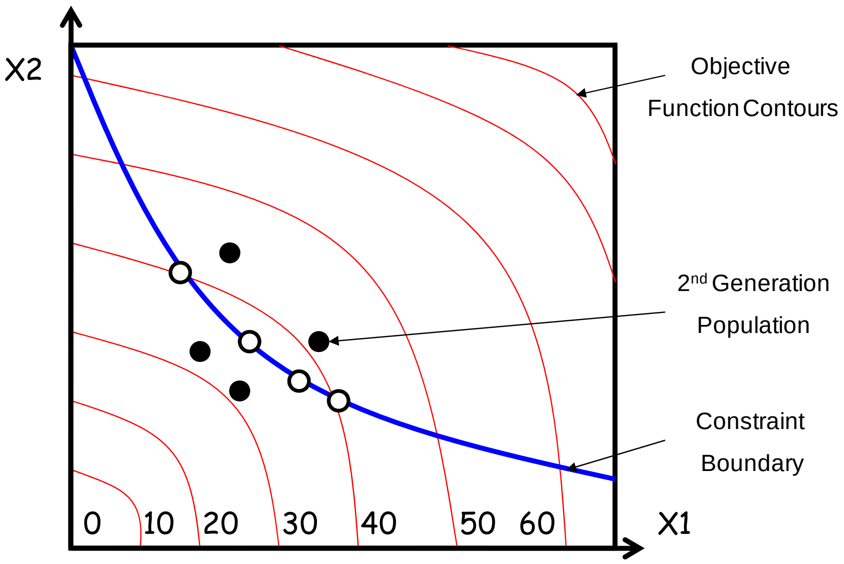

Figure 8. Generation of New Population (2 dimensional example)

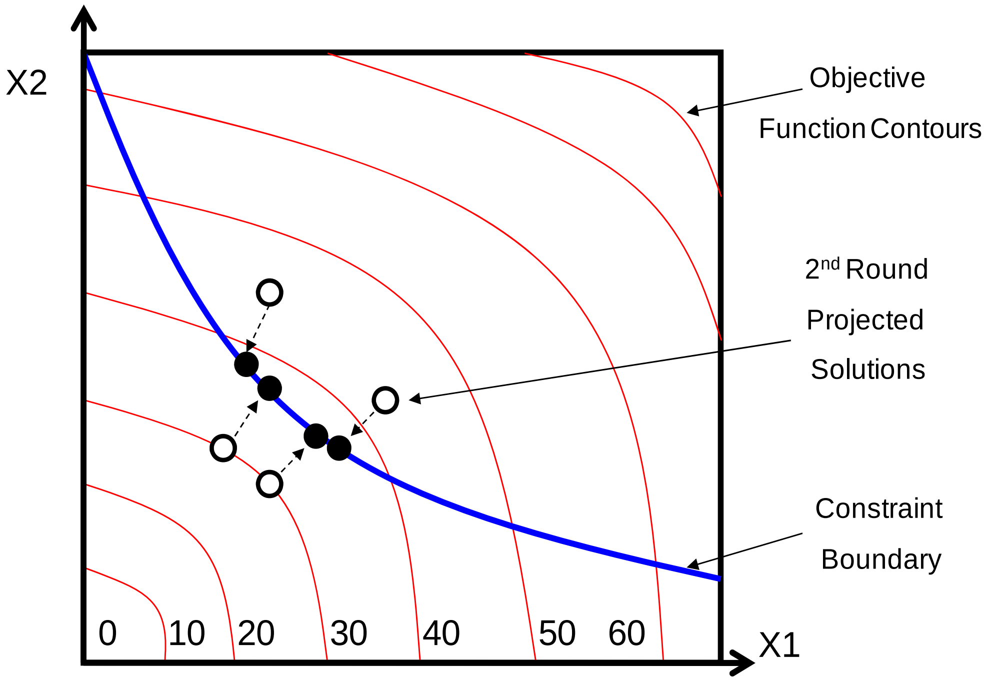

Figure 9. Projection of 2[nd] Round Solutions to Constraint Boundary (2 dimensional example)

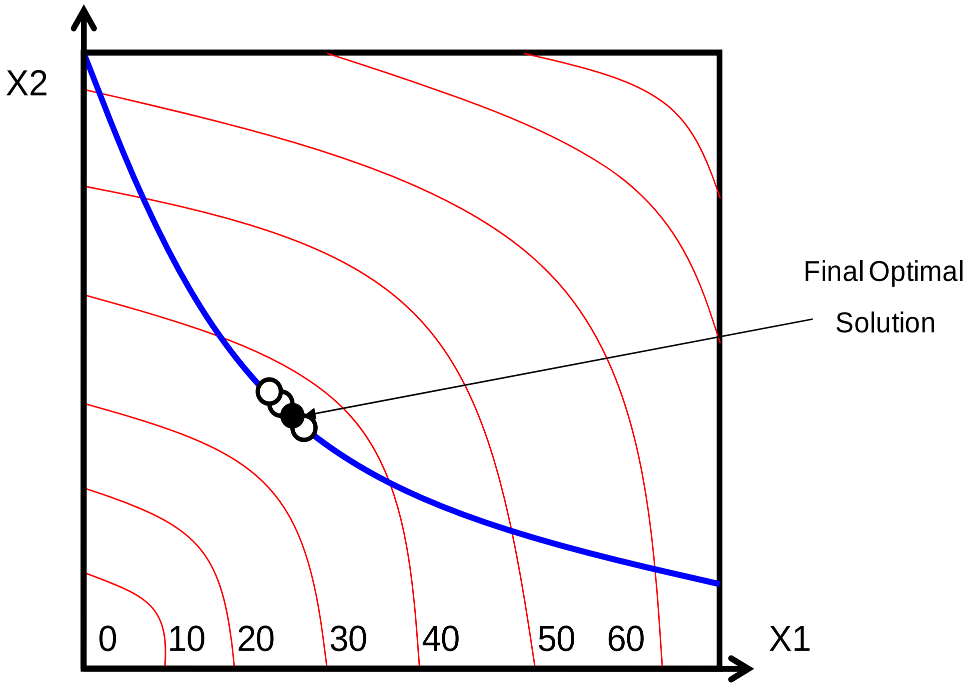

Figure 10. Collapse of Nth Round Solutions to Optimal Solution (2 dimensional example)

6 Summary

Three different explicit formulations of the optimal pump scheduling problem have been presented. The resulting formulations may be solved by using either unconstrained methods along with penalty terms or constrained methods that explicitly incorporate the constraints via the mechanisms involved in the algorithm. In either case, the implicit system constraints can be solved directly using an external simulation program which is linked to the optimization algorithm via subroutine calls. Given the complexity of the system being modeled (i.e. water distribution systems), it is felt that the provided formulations as well as the explicitly constrained GA may have applicability to other complex scheduling problems as well.

7 References

1. Box, M. J. A new method for constrained optimization and a comparison with other methods. *Computer Journal,* Vol. 8, No. 1, 42-52 (1965).

2. Brion, L., M., *Methodology for Optimal Pumping Stations in Water Distribution Systems,*" Ph.D., Dissertation, University of Texas at Austin (1990).

3. Carpentier, P., and G. Cohen, "Decomposition, Coordination and Aggregation in the Optimal Control of a Large Water Supply Network," *Proc. of the 9th Triennial IFAC World Congress,* Budapest, pp. 3207-3212, (1984).

4. Chase, D., V., "*A Computer Program for Optimal Control of Water Supply Pump Stations: Developing and Testing,"* USACERL Technical Report N-90/14, U.S. Army Corps of Engineers, Construction Engineering Research Laboratory, Champaign, Illinois, pp. 1-99 (1990).

5. Chase, D., and L. Ormsbee, "Optimal Pump Operation of Water Distribution Systems with Multiple Storage Tanks," *Proceedings of the Amer. Water Works Association. Computer Specialty Conf*., Denver, CO, pp. 205-214 (1989).

6. Chase, D., and Ormsbee, L., "An Alternate Formulation of Time as a Decision Variable To Facilitate Real-Time Operation of Water Supply Systems", *Proceedings of the 18th Annual Conference of the ASCE Water Resources Planning and Management Division*, New Orleans, LA, pp. 923-927 (1991).

7. Coulbeck, B., "Optimisation of Water Networks," *Transactions of the Institute of Measurement and Control*, Vol. 6, No. 5, pp. 271-279 (1984).

8. DeMoyer, R., and L. Horowitz, *A Systems Approach to Water Distribution Modeling and Control,* Lexington Books, pp. 143 (1975).

9. Lansey, K., and Q. Zhong, "A Methodology for Optimal Control of Pump Stations," *Proceedings of the ASCE Water Resources Planning and Management Specialty Conference,* Fort Worth, TX, pp. 58-61 (1990).

10. Ormsbee, L.E. (1985). "OPNET: A Nonlinear Algorithm for Hydraulic Networks", *Proceedings of the ASCE Water Resources Planning and Management Conference: Computer Applications in Water Resources,* Buffalo, New York, pp. 739-748 (1985).

11. Ormsbee, L.E., and Lansey, K.E., "Optimal Operation of Water Distribution Systems, *ASCE Journal of Water Resources Planning and Management*, Vol. 120, No. 2, pp. 237-252 (1994).

12. Ormsbee, L.E., and Lingireddy, S. L., "Nonlinear Heuristic for Pump Operations," *ASCE Journal of Water Resources Planning and Management*, Vol. 121, No 4, pp. 302-309 (1995).

13. Ormsbee, L.E., Walski, T., Chase, D., Sharp, W., "Methodology for Improving Pump Operation Efficiency," ASCE Journal of Water Resources Planning and Management, Vol., 115, No. 2, pp. 148-164 (1989).

14. Ormsbee, L., Reddy, S., and Chase, D., "Comparison of Three Nonlinear Control Algorithms for the Optimal Operation of Water Supply Pumping Systems, *Proceedings of the International Conference on Integrated Computer Applications for Water Supply and Distribution*, Leicester, England, pp. 259-271 (1993).

15. Ormsbee, L., and Reddy, S., "Pumping System Control Using Genetic Optimization and Neural Networks," *Proceedings of 7th IFAC/IFORS/IMACS Symposium on Large Scale Systems: Theory and Applications*, London, England, (1995).

16. Sabet, M., and O. Helweg, "Cost Effective Operation of Urban Water Supply System Using Dynamic Programming," *Water Resources Bulletin*, AWRA, Vol. 21, No. 1, pp. 75-81 (1985).

17. Teegavarapu, R., Bradley, S., and Ormsbee, L., "Probabilistic Goal Driven Watershed Management", *Proceeding of the World Environmental and Water Resources Congress,* Tampa, May 15-19, (2007).

18. Tufail, M., *Optimal Water Quality Management Strategies for Urban Watersheds Using Macro-Level Simulation Models Linked with Evolutionary Algorithms,* Submitted in Partial Fulfillment of the Doctor of Philosophy Degree, University of Kentucky, Lexington, Kentucky, 40506 (2006).

19. Tufail, M., and Ormsbee, L.E., "Optimal Water Quality Management Strategies for Urban Watersheds Using Macro-Level Simulation and Optimization Models, *ASCE Journal of Water Resources Management,* Vol. 135, No. 4, pp. 276-285, (2009).

20. Wood, D. J., User's Manual – Computer Analysis of Flow in Pipe Networks Including Extended Period Simulations, Department of Civil Engineering, University of Kentucky, Lexington, KY, (1980)

21. Zessler, U., and U. Shamir, "Optimal Operation of Water Distribution Systems," *ASCE Journal of Water Resources Planning and Management*, Vol. 115, No. 6, pp. 735-752 (1989).