WATER RESOURCES RESEARCH, VOL. 49, 3038–3043, doi:10.1002/wrcr.20211, 2013

Tracking trade transactions in water resource systems: A node-arc optimization formulation

Tohid Erfani,[1] Ivana Huskova,[1] and Julien J. Harou[1]

Received 12 November 2012; revised 14 February 2013; accepted 24 March 2013; published 2 May 2013.

[1] We formulate and apply a multicommodity network flow node-arc optimization model capable of tracking trade transactions in complex water resource systems. The model uses a simple node to node network connectivity matrix and does not require preprocessing of all possible flow paths in the network. We compare the proposed node-arc formulation with an existing arc-path (flow path) formulation and explain the advantages and difficulties of both approaches. We verify the proposed formulation model on a hypothetical water distribution network. Results indicate the arc-path model solves the problem with fewer constraints, but the proposed formulation allows using a simple network connectivity matrix which simplifies modeling large or complex networks. The proposed algorithm allows converting existing node-arc hydroeconomic models that broadly represent water trading to ones that also track individual supplier-receiver relationships (trade transactions).

Citation: Erfani, T., I. Huskova, and J. J. Harou (2013), Tracking trade transactions in water resource systems: A node-arc optimization formulation, *Water Resour. Res.*, *49*, 3038–3043, doi:10.1002/wrcr.20211.

1. Motivation and Previous Work

[2] Trading water rights is an appealing option in many river basins where no further resource development is possible for hydrological, ecological, or financial reasons [*Chong and Sunding*, 2006]. Most water trading is done by pairwise exchange of a water right or abstraction license between willing buyers and sellers. The local and regional hydrologic and economic impacts of trading and the impacts of different regulatory policies will depend on which specific transactions occur: what volumes are traded between which abstractors. One impediment of water trading is transaction cost which varies depending on the trading parties thus influencing which transactions occur in practice [*Young*, 1986]. Most hydroeconomic modeling studies [*Harou et al*., 2009] that consider the implementation and benefits of water trading use models that represent regional or sectoral movement of water but do not track individual transactions [e.g., *Ward and Lynch*, 1996; *Draper et al*., 2003; *Harou et al*., 2010]. This is appropriate for regional-scale policy studies looking at broad impacts of water trading. However, for water managers interested in designing specific trading rules and regulations or evaluating more detailed hydrologic or economic impacts of trading for a particular water resource system, a model that represents individual transactions is valuable.

[1]Department of Civil, Environmental and Geomatic Engineering, University College London, London, UK.

Corresponding author: J. J. Harou, Department of Civil, Environmental and Geomatic Engineering, University College London, Gower Street, London WC1E 6BT, UK. (j.harou@ucl.ac.uk)

0043-1397/13/10.1002/wrcr.20211

[3] A formulation that can track individual trades between water right holders will, for any arc in the network (stream reach, canal, or pipeline), specify the distinct supplier-recipient transactions and other components of flow. *Cheng et al*. [2009] presented such a formulation that could "clearly describe water deliveries by identifying the relationship between suppliers and receivers." In their seminal contribution, "the water rights owner, water quantity, water location, and associated flow path of each delivery action are represented explicitly in the results rather than merely as an optimized total flow quantity in each arc of a distribution network." This allows representing transaction costs and trading rules in detail, customized if necessary for each pair of water right holders engaged in a potential transaction.

[4] The *Cheng et al*. [2009] model can be considered as a multicommodity network flow (MCNF) problem. The MCNF problem can be formulated in two ways, namely, *node-arc* and *arc-path* [*Ouorou et al*., 2000]. In the node-arc formulation proposed in this paper the flow decision variables are the units of commodity transferred through an arc (link between two nodes). In the arc-path approach the decision variables are units transferred along a "path" consisting of multiple connected arcs. *Cheng et al*. [2009] use an arc-path formulation; in their study all supplier-receiver paths are predefined using an iterative algorithm and stored in a sparse binary *flow path* matrix. The iterative algorithm exploits the information of an *incidence matrix* where the entering (+1) and leaving (−1) arcs (columns) to and from each node (rows) build up the matrix structure. To formulate the optimization problem, $x_{r,t}^{i,j}$ is the decision variable for water delivery x via a flow path r with delivery relationship between supplier i and receiver j at time step t. The model constraint set includes the regular mass balance constraints for all source nodes, demand nodes, and reservoirs. To accommodate reservoirs in the arc-path formulation,

storage nodes are substituted by one dummy supplier, one receiver, and two extra arcs. The solution is decoded using the flow path matrix to determine the flow in each arc of the network.

[5] Section 2 describes the proposed node-arc optimization model formulation which accomplishes the same tasks as the *Cheng et al*. [2009] arc-path formulation without requiring enumeration of all flow paths. Section 2 then compares the node-arc and arc-path formulations. In section 3, a hypothetical water distribution network is used to compare the structure and results of both approaches.

2. *Node-arc* MCNF Formulation

[6] The proposed model is a node-arc MCNF formulation where flows on arcs between the nodes (with an extra index recording the delivery relationship) are the decision variables. Consider a directed network $G = (N, A)$ with N being the set of nodes and A being the set of arcs (links) which connect each pair of nodes. We associate with each arc a flow $x_{i,j}^{k,t}$ indicating the quantity of water owned originally by node k transferred through arc $i \rightarrow j$ for time period t. The water network architecture is defined with a (0,1)-*connectivity* matrix, **CO**. A "1" in the matrix means node i (rows) is connected to node j (columns) with a flow direction from i to j. In this model, a supplier is a node that can provide water, and the receiver is a node that can make use of the supplier's water. Each supplier has a *type* tag to distinguish its water from the other suppliers. The k index in $x_{i,j}^{k,t}$ refers to this water type which records supplier-receiver relationships. Various objective functions that optimize the water exchange and delivery strategy may be used. The constraint equations described next are designed to recreate the *Cheng et al*. [2009] model using the node-arc formulation.

2.1. Model Constraints

2.1.1. Conservation Mass Balance

[7] In our network model, sources provide water, and demand nodes use the water. The mass balance equations for source nodes are

$$\mathrm{SO}_i^{k,t} = \sum_{\substack{j \\ \mathbf{CO}_{i,j}=1}} x_{i,j}^{k,t} + \mathrm{ST}_j \times \mathrm{Res}_j^{k,t+1}, \quad \forall i \in \text{Source}, \quad k \in \text{Type}, \quad t \in \text{Time}, \tag{1}$$

where $\mathrm{SO}_i^{k,t}$ is the inflow to source i at time t of water type k. **CO** shows whether or not $i \rightarrow j$ is an arc in the network, and $\mathrm{Res}_j^{k,t+1}$ is the water type k stored at reservoir j at time t; ST_j equals one if the connecting node to the source is a reservoir. At water demand node i, the mass balance equation to satisfy the quantity of water demanded, DE_i^t, is

$$DE_i^t = \sum_{\substack{j \in \text{Type} \\ \mathbf{CO}_{j,i}=1}} \sum_{k} x_{j,i}^{k,t} + Def_i^t \quad \forall i \in Demand, \quad t \in \text{Time}, \tag{2}$$

where Def_i^t is water deficit at time t for demand i. To avoid deficits, Def_i^t can be penalized in the objective function. For junction nodes where flow in equals flow out,

$$\sum_{\substack{j \\ \mathbf{CO}_{i,j}=1}} x_{i,j}^{k,t} = \sum_{\substack{j \\ \mathbf{CO}_{j,i}=1}} x_{j,i}^{k,t}, \quad \forall i \in \text{Junction}, \ k \in \text{Type}, \ t \in \text{Time}, \tag{3}$$

is introduced. At network terminal points, flow goes to sinks represented by the following constraint:

$$\sum_{k \in \text{Type}} \sum_{\substack{i \\ \mathbf{CO}_{i,j}=1}} x_{i,j}^{k,t} = \mathrm{Wd}_j^t \ \ \forall j \in \text{Discharge}, t \in \text{Time}, \tag{4}$$

where Wd_j^t is the amount of discharged water. One can penalize spilled water to minimize discharge out of the system.

2.1.2. Storage Balance

[8] Mass balance in storage nodes is applied to track storage (transferring a volume over time rather than space). This leads to

$$\sum_{\substack{i \\ \mathbf{CO}_{i,j}=1}} x_{i,j}^{k,t} + \mathrm{Res}_j^{k,t} = \mathrm{Res}_j^{k,t+1} + \sum_{\substack{i \\ \mathbf{CO}_{j,i}=1}} x_{j,i}^{k,t}, \quad \forall j \in \text{Reservoir}, k \in \text{Type}, t \in \text{Time}. \tag{5}$$

2.1.3. Transportability

[9] The flow in each arc is restricted by minimum flows and maximum capacity. This introduces

$$\mathrm{MinF}_{i,j} \le \sum_{\substack{k \in \text{Type} \\ \mathbf{CO}_{i,j}=1}} x_{i,j}^{k,t} \le \mathrm{MaxF}_{i,j} \quad \forall i,j \in \text{Node}, t \in \text{Time}, \tag{6}$$

where $\mathrm{MinF}_{i,j}$ and $\mathrm{MaxF}_{i,j}$ are the minimum and maximum flows in each arc, respectively.

2.2. Comparison of the Formulations

[10] The proposed *node-arc* formulation differs from the *arc-path* one in that the decision variables are arc flows rather than a flow *path* from a source to a demand node. The proposed formulation represents storage with a single node (without additional arcs) and as it does not use flow paths, does not require enumerating all possible flow paths prior to optimization. The following points highlight the differences between the formulations:

[11] (1) The incidence matrix (required by the *arc-path* formulation) for a water network with n nodes has $n \times e$ entries. The parameter e, the number of arcs, ranges from $n - 1$ to $n(n-1)/2$ depending on the network topology. In the proposed formulation, the connectivity matrix is $n \times n$, and its dimension is unchanged regardless of network architecture. This implies the connectivity matrix has an upper bound of n^2 entries, while the incidence matrix can reach up to n^3.

[12] (2) The two network data matrices required by the *arc-path* approach (incidence and flow path matrices) may become cumbersome to generate if exchanges are possible between many sites in large or complex networks; they must also be regenerated after each change to network topology. In networks with few paths, the *arc-path* model works well since this formulation may lead to simple decompositions due to the structure of constraints. However, finding all possible paths in a directed graph can be a cumbersome extra step [*Higashiyama and Ariyoshi*, 1984] for large pseudogrid networks with multiple sources and numerous possible paths between each pair of nodes. The number of paths grows exponentially with the problem dimensions regardless of network structure, so finding all paths in a directed network may be impractical for certain large networks [*Ouorou et al.*, 2000]. Additionally, when water right holders can sell water in the arc-path formulation, all sellers need a unique dummy node (as a source node) and a new connecting arc; this further increases the dimension of the incidence matrix.

[13] (3) If a node is added into the network or the functionality of a node is changed, the incidence matrix is augmented with a new node and a couple of arcs. The flow path matrix, however, needs to be redefined based on the type of the added node (whether it is receiver, a supplier, or a junction). This requires reidentifying the paths from and to the existing nodes as well as for the new node. In the node-arc approach the connectivity matrix is augmented with a row and a column to represent the extra added node into the water network. The same situation applies if a node is removed from the network.

[14] (4) While in the arc-path formulation a two-way flow arc can be included by introducing two separate arcs with opposite direction which increases the dimensionality of the incidence matrix, the bidirected flow arc in the proposed node-arc formulation can be added by altering the connectivity matrix without introducing two new arcs.

[15] (5) Depending on network topology the proposed node-arc model may be larger and therefore potentially slower to solve. The *arc-path* formulation may yield a mathematical program with fewer decision variables depending on the number of flow paths. The proposed *node-arc* formulation has a supplementary mass balance constraint for each junction node in the network.

[16] It should be noted that the capacity constraints on each arc can be applied in the same way with both formulations. Next we test the model on a hypothetical water network.

3. Application

[17] To demonstrate the proposed formulation the hypothetical water network problem presented by *Cheng et al.* [2009] is solved using both formulations. The network and flow allocation problem shown in Figure 1 are described by the following matrices (blanks represent zeros):

$$
\mathbf{CO} = \begin{array}{c} \\ c1 \\ c2 \\ 4 \\ 6 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15 \end{array}
\begin{array}{cccccccccccc} c1 & c2 & 4 & 6 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\
 & & 1 & & & & & & & & & \\
 & & & 1 & & & & & & & & \\
 & & & & & 1 & & & & & & \\
 & & & & & & 1 & & & & & \\
 & & & & & & & & & & & \\
 & & & & 1 & & & & 1 & & & \\
 & & & & & 1 & & & & 1 & & \\
 & & & & & & & & & & & \\
 & & & & & & & 1 & & 1 & & \\
 & & & & & & & & & & 1 & 1 \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \end{array}
$$

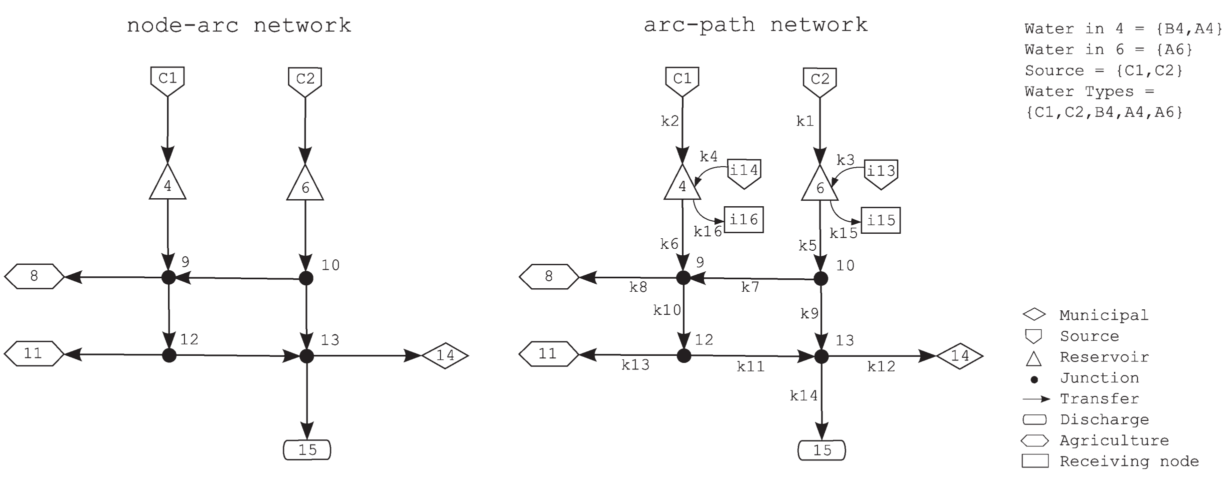

Figure 1. Hypothetical water distribution network from *Cheng et al.* [2009]. Suppliers i14, i13, receivers i16, i15, and sets of extra arcs k4, k16 and k3, k15 are required to accommodate reservoirs 4 and 6 in arc-path formulation. An example of a decision variable in the node-arc formulation is $x_{10,9}^{C2,t}$ (flow during time t from node 10 to node 9 originating at node C2). An example of an arc-path decision variable is $x_{r1,t}^{C2,14}$ (flow along flow path $r1$: C2 → 6 → 10 → 13 → 14 at time t). Flow path $r1$ is defined in the flow path matrix **FP**.

$$
\mathbf{IN} = \begin{array}{c} \\ c2 \\ c1 \\ 6 \\ 4 \\ 10 \\ 9 \\ 13 \\ 12 \\ 8 \\ 11 \\ 14 \\ 15 \\ 13 \\ 14 \\ 15 \\ 16 \end{array}
\begin{array}{cccccccccccccccc}
k1 & k2 & k3 & k4 & k5 & k6 & k7 & k8 & k9 & k10 & k11 & k12 & k13 & k14 & k15 & k16 \\
-1 & & & & & & & & & & & & & & & \\
 & -1 & & & & & & & & & & & & & & \\
1 & & 1 & & -1 & & & & & & & & & & -1 & \\
 & 1 & & 1 & & -1 & & & & & & & & & & -1 \\
 & & & & 1 & & -1 & & -1 & & & & & & & \\
 & & & & & 1 & 1 & -1 & & -1 & & & & & & \\
 & & & & & & & & 1 & & 1 & -1 & & -1 & & \\
 & & & & & & & & & 1 & -1 & & -1 & & & \\
 & & & & & & & 1 & & & & & & & & \\
 & & & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & 1 & & & & \\
 & & & & & & & & & & & & & 1 & & \\
 & & -1 & & & & & & & & & & & & & \\
 & & & -1 & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & 1 & \\
 & & & & & & & & & & & & & & & 1
\end{array}
$$

$$
\mathbf{FP} = \begin{array}{c} \\ r1 \\ r2 \\ r3 \\ r4 \\ r5 \\ r6 \\ r7 \\ r8 \\ r9 \\ r10 \\ r11 \\ r12 \\ r13 \\ r14 \\ r15 \\ r16 \\ r17 \\ r18 \\ r19 \\ r20 \\ r21 \\ r22 \\ r23 \\ r24 \end{array}
\begin{array}{cccccccccccccccc}
k1 & k2 & k3 & k4 & k5 & k6 & k7 & k8 & k9 & k10 & k11 & k12 & k13 & k14 & k15 & k16 \\
1 & & & & 1 & & & & 1 & & & 1 & & & & \\
1 & & & & 1 & & & & 1 & & & & & 1 & & \\
1 & & & & 1 & & 1 & 1 & & & & & & & & \\
1 & & & & 1 & & 1 & & & 1 & & & 1 & & & \\
 & & 1 & & 1 & & & & 1 & & & 1 & & & & \\
 & & 1 & & 1 & & & & 1 & & & & & 1 & & \\
 & & 1 & & 1 & & 1 & 1 & & & & & & & & \\
 & & 1 & & 1 & & 1 & & & 1 & & & 1 & & & \\
 & 1 & & & & 1 & & & & 1 & 1 & 1 & & & & \\
 & 1 & & & & 1 & & & & 1 & 1 & & & 1 & & \\
 & 1 & & & & 1 & & 1 & & & & & & & & \\
 & 1 & & & & 1 & & & & 1 & & & 1 & & & \\
 & & & 1 & & 1 & & & & 1 & 1 & 1 & & & & \\
 & & & 1 & & 1 & & & & 1 & 1 & & & 1 & & \\
 & & & 1 & & 1 & & 1 & & & & & & & & \\
 & & & 1 & & 1 & & & & 1 & & & 1 & & & \\
1 & & & & 1 & & 1 & & & 1 & 1 & 1 & & & & \\
1 & & & & 1 & & 1 & & & 1 & 1 & & & 1 & & \\
 & & 1 & & 1 & & 1 & & & 1 & 1 & 1 & & & & \\
 & & 1 & & 1 & & 1 & & & 1 & 1 & & & 1 & & \\
1 & & & & & & & & & & & & & & 1 & \\
 & & 1 & & & & & & & & & & & & 1 & \\
 & 1 & & & & & & & & & & & & & & 1 \\
 & & & 1 & & & & & & & & & & & & 1
\end{array}
$$

where **CO** is the connectivity matrix for the node-arc formulation, and **IN** and **FP** are the incidence and flow path matrices, respectively, for the arc-path model. The system has two sources (water type C) and two reservoirs (A and B). To identify the sources from which demand nodes receive water, we tag the water based on its origin. Five different water types are identified: C1 from source 1, C2 from source 2, A4 and B4 from reservoir 4, and A6 from

Table 1. Optimal Water Delivery Results of Node-Arc Model[a]

									Time		
Demand Node (Receiver)		Path of Transaction						Water Type (Supplier)	t_1	t_2	t_3
Agriculture 8		1	4	9	8			C1	0.1	0.1	0.1
		4	9	8				B4	0.25		
		4	9	8				A4	0.4	0.6	
		6	10	9	8			A6		0.05	0.65
Agriculture 11		4	9	12	11			B4	0.75		
		4	9	12	11			A4		0.75	
		6	10	9	12	11		A6			0.6
Municipal 14		2	6	10	13	14		C2	0.1	0.1	0.1
		4	9	12	13	14		A4	0.25		
	*	6	10	9	12	13	14	A6		0.25	0.25
	*	6	10	13	14			A6	0.4	0.4	0.4

[a]The symbol "*" represents different paths due to capacity constraint on arc 10 → 13.

reservoir 6. Water type C arrives from outside the agency's network, B is traded on a market, and A is owned by the agency and not paid for. The task is to track the water in the network while satisfying the 0.75×10^6 m^3 of water for each demand node in each time period. The initial storages at A4, B4, and A6 are 2, 1, and 3 Mm^3, while 0.1×10^6 m^3 of water is available from each source node (C1 and C2) at each time period. The objective is to minimize the cost of water delivery ($1 conveyance charge in each arc) with penalties for deficits ($100 for municipal demand node 14 and $10 for agricultural demand nodes 8 and 11) and discharge from the system ($80/$m^3$). Water transfers incur an additional charge of $1, $3, and $5/$m^3$ for the first, second, and third time periods, respectively. The same charge is applied if water type B is traded from reservoir 4. The capacity of 3×10^6 m^3 is enforced on all arcs except for 10 → 13 which is capped at 0.5 Mm^3, and reservoirs have a capacity of 3×10^6 m^3.

[18] Table 1 summarizes results. An objective function of 25 M is the same as *Cheng et al*. [2009] though with a different allocation indicating that the problem has multiple optima. The results show that the proposed formulation can identify seller-buyer relationships and model flow and storage. For example, in the second time period, 0.1 unit of water is transferred without cost (since it is water type A) from reservoir 4 to demand node 14 through 4 → 9 → 12 → 13 → 14. The paths marked by * in Table 1 show the model tracks the water types adapting to the restriction on arc 10 → 13. Although the arc-path model is smaller in terms of constraints, the proposed node-arc formulation solves the problem with a 12×12 connectivity matrix rather than using a 16×16 incidence matrix and a 24×16 flow path matrix (Table 2). The arc-path model solution requires postprocessing to obtain arc flows after optimization whereas the node-arc formulation obtains them directly.

4. Conclusions

[19] In this paper we have proposed an optimization model formulation for tracking transactions such as trades in water resource networks. The proposed node-arc formulation is an alternative to the arc-path multicommodity flow network model proposed by *Cheng et al*. [2009]. Both models can represent supplier-receiver relationships that occur, for example, in water right trading. The proposed formulation uses a simple network connectivity matrix. This implies existing hydroeconomic models that use connectivity matrices but currently do not track transactions can relatively easily be modified to do so by appending a new index to the flow decision variable which represents water's origin and hence its ownership. Tagging ownership allows supplier-receiver transaction tracking. The proposed formulation typically leads to larger constraint sets than the arc-path approach, but the network topology data are easier to manage, and enumerating all flow paths between source and demand nodes is not required. Both formulations were tested on a small hypothetical water distribution network and obtained the same objective function value.

Table 2. Differences Between the Linear Programs and Input Data Matrices Generated by Each Formulation for the Example Network

	Model Size[a]		Network Matrices		
	Variables	Constraints	Connectivity (**CO**)	Incidence (**IN**)	Flow path (**FP**)
Arc-path	79	37	–	16×16	24×16
Node-arc	75	56	12×12	–	–

[a]Per time step.

[20] **Acknowledgments.** The first author thanks Hamid Mokhtar for technical advice. Two anonymous reviewers provided comments that improved this paper.

References

Cheng, W.-C., N.-S. Hsu, W.-M. Cheng, and W. W. G. Yeh (2009), A flow path model for regional water distribution optimization, *Water Resour. Res.*, *45*, W09411, doi:10.1029/2009WR007826.

Chong, H., and D. Sunding (2006), Water markets and trading, *Annu. Rev. Environ. Resour.*, *31*, 239–264.

Draper, A. J., M. W. Jenkins, K. W. Kirby, J. R. Lund, and R. E. Howitt (2003), Economic-engineering optimization for California water management, *J. Water Resour. Plann. Manage.—ASCE*, *129*(3), 155–164.

Harou, J. J., M. Pulido-Velazquez, D. E. Rosenberg, J. Medellin-Azuara, J. R. Lund, and R. E. Howitt (2009), Hydro-economic models: Concepts, design, applications, and future prospects, *J. Hydrol.*, *375*(3–4), 627–643.

Harou, J. J., J. Medellin-Azuara, T. Zhu, S. K. Tanaka, J. R. Lund, S. Stine, M. A. Olivares, and M. W. Jenkins (2010), Economic consequences of optimized water management for a prolonged, severe drought in California, *Water Resour. Res.*, *46*, W05522, doi:10.1029/2008WR007681.

Higashiyama, Y., and H. Ariyoshi (1984), An algorithm for generation of s-t acyclic graphs, *Electron. Commun. Jpn. (Part I: Commun.)*, *67*(12), 10–18.

Ouorou, A., P. Mahey, and J. P. Vial (2000), A survey of algorithms for convex multicommodity flow problems, *Manage. Sci.*, *46*(1), 126–147.

Ward, F. A., and T. P. Lynch (1996), Integrated river basin optimization: Modeling economic and hydrologic interdependence, *Water Resour. Bull.*, *32*(6), 1127–1138.

Young, R. A. (1986), Why are there so few transactions among water users, *Am. J. Agric. Econ.*, *68*(5), 1143–1151.