Multi-agent management of integrated food-energy-water systems using stochastic games: from Nash equilibrium to the social optimum

Milad Memarzadeh 1, Scott Moura 2 and Arpad Horvath 2

1 USRA, Data Sciences Group, NASA Ames Research Center, Moffett Field, CA, 94035 United States of America
2 Department of Civil and Environmental Engineering, University of California, Berkeley, CA, 94720 United States of America
E-mail: memarzadeh.milad@gmail.com

Keywords: food-energy-water systems, game theory, multi-agent management

Abstract

System-level integration and optimization of food-energy-water systems (FEWS) require coordination of multiple agencies and decision-makers and incorporating their interdependence. In general, such coordination might be hard to achieve. As a result, the literature on FEWS management either optimizes the operations for one sector (or one decision-maker), or models interdependence among the sectors without optimizing their operations. In this article, we develop a novel multi-agent management optimization approach that is able to incorporate stochasticity and uncertainty in the system’s dynamics and interdependence of the water and energy resources for food production. The proposed method is the first attempt to utilize fundamentals of decision and game theories to optimize operations of multi-agent FEWS. We specifically focus on differentiating between (1) cooperative decision optimization of the operations, where all decision-makers cooperate to achieve the best outcome for the whole system, the social optimum, and (2) non-cooperative decision-making of the agents, the Nash equilibrium. Illustrating with a real-world case study of FEWS in Ventura County, California, we show the difference between the cooperative and non-cooperative decision making in terms of long-term expected cost of managing the system. We further show how the extra costs associated with utilizing the renewable sources of water and energy could be incentivised, so that the non-cooperative solution (the Nash equilibrium) would naturally converge to the best outcome for the whole system (the social optimum).

1. Introduction

Food-energy-water systems (FEWS) are among the most crucial yet complex interdependent systems and sustainability of the food production relies on coordination of multiple agencies and decision-makers (Daher et al 2019, Kurian et al 2018, Madani 2010). However, the coordination among different decision-makers might be hard to achieve, as the end-goal, or policy constraints and regulations for each decision-maker is different and often conflictory. Moreover, beyond this conflictory decision-making process, there are difficulties corresponding to the interdependence of the food production, energy and water utilization and their incorporation in the operation optimization (Albrecht et al 2018, Karan and Asadi 2018, Liu et al 2016, Namany et al 2019). Incorporating such interdependence in multi-agent management of resources available for food production is a challenging task. As a result, the literature on FEWS management either focuses on single-agent decision-making process (assuming that there is one decision-maker in the whole process) and identifies optimal strategies for such management, or focuses on the flow of information among the different sectors and decision-makers, ignoring the optimization of such multi-agent decision-making process.

Single-agent decision optimization of the FEWS operations has a longer history and is more studied. In this literature, one needs to identify objectives of management, the constraints to the manager, strategies available to her, utilities corresponding to the operational costs, profits, and environmental impacts, as well as the effects of exogenous variables such as environmental variations (Memarzadeh et al 2019). Once these are quantified, several approaches
have been used to identify the near-optimal management strategies and outcomes of such implementations on the FEWS operations in long-term, including mathematical programming (Bieber et al 2018, Zhang et al 2018), life-cycle assessment (Bell et al 2018, Wang et al 2017, Sherwood et al 2017) and scenario planning (Ramaswami et al 2017, Chaudhary et al 2018, Karan et al 2018).

Another area of research develops methods based on network and graph theory to model and quantify the flow of information and resources among different sectors involved in the whole FEWS operations. Kurian et al (2018) and Daher et al (2019) focus on the inter-connection and communication among different sectors in the FEWS governance network, while Givens et al (2018) emphasizes the social aspects of FEWS and its effect on the network’s dynamics. Tsolas et al (2018) utilizes a graph-theoretic approach to optimize the flow of energy and water within a FEWS network, as well as designing the network topology optimally. Liang et al (2019) models and quantifies the material and energy flow among different sectors of FEWS and their inter-connection.

In this article, we develop a novel multi-agent management approach based on fundamentals of decision and game theories. The proposed approach is able to incorporate stochasticity and uncertainty in the system’s dynamics and interdependence of the water and energy resources for food production. We specifically focus on differentiating between (1) cooperative decision optimization of the operations, where all decision-makers cooperate to achieve the best outcome for the whole system, the social optimum, and (2) non-cooperative decision-making of each decision-maker in response to their best knowledge of other decision-makers’ goals and intentions, the Nash equilibrium. We quantify the advantage of these strategies to a scenario where each decision-maker ignores the existence of others and acts individually. We further show the difference between the social optimum and the Nash equilibrium on multi-agent management of a real-world case study of FEWS in Ventura County, California.

2. Methodology

In this section, we propose a scalable algorithm for optimizing operations of FEWS managed by multiple decision-makers. In order to build a deeper understanding of the methodology, we introduce a problem of multi-agent management of a FEWS based on the data obtained from a case study in Ventura County, California (Bell et al 2018).

We define state at each time step \( t \) (\( s_t \in S \)) as the amount of water (\( w_t \in W \)) and energy (\( e_t \in E \)) resources (from conventional sources) available for agricultural production, \( S = W \times E \). There are \( N = 4 \) agents, each corresponding to one of the four main crops in the region, strawberry, lemon, avocado, and celery. Each agent \( i \) makes decisions at each time step \( t \) on what sources of water (\( a_{w,t}^i \) = conventional, \( Conv_w \), or recycled, \( Rec_w \)) and energy (\( a_{e,t}^i \) = conventional, \( Conv_e \), or renewable, \( Ren_e \)) to use. The conventional water source in the region comes from run-offs into the nearby river as well as local wells, and the conventional energy source is mostly natural gas (Bell et al 2018). The renewable resources correspond to recycled water and wind energy, which is assumed to only cover 25% of total water and energy demand of these crops in the region.

The amount of conventional water (\( w_t \)) available for irrigation changes in time according to

\[
 w_{t+1} = w_t - \sum_{i \in \{1, \ldots, N\}} a_{w,t}^i \lambda - r_t^w - w_e \cdot 1_{Conv}(a_{e,t}) + w_w \cdot 1_{Rec}(a_{w,t}) + z_{w,t} \tag{1}
\]

where \( w_t \) is the available water at time step \( t \), \( \lambda \in \{\text{Spring, Summer, Fall, Winter}\} \) is the variable representing the seasonal changes, \( a_{w,t}^i \) is the water demand for crop \( i \in \{\text{strawberry, lemon, avocado, celery}\} \) at time step \( t \) and in season \( \lambda \), and \( r_t^w \) is the seasonal rainfall (data obtained from the Western Regional Climate Center: https://wrcc.dri.edu) for Ventura County.

We model the inter-connection of the water and energy resources as a deterministic function of agents’ decisions. If any of the agents chooses to utilize the conventional energy source (\( 1_{Conv}(a_{e,t}) \)), we assume that \( w_t \) amount of available water will be consumed for generation and distribution of the energy (assumed to be 20% of the maximum available amount). On the other hand, if any agent chooses to utilize the recycled water (\( 1_{Rec}(a_{w,t}) \)), we assume that this translates to a deterministic boost in the amount of \( w_t \) to the total available water for irrigation (which is maximum of 25% in Ventura County (Bell et al 2018)).

In many real-world scenarios, changes in the amount of available water from time \( t \) to \( t + 1 \) are not deterministic since a lot of other factors (that we might not have data and knowledge of to incorporate) are involved. Some examples of such factors are effects of evaporation and transpiration on available water, or errors in water and energy consumption as well as demand for crop production. As a result, researchers include a degree of stochasticity to incorporate these into account. \( z_{w,t} \) serves as a factor to incorporate potential stochasticity that is not formulated here due to lack of knowledge and data. We assume the stochasticity to be Normal distribution with a known standard deviation, truncated at zero to avoid negative values. \( z_{w,t} \sim N(0, \sigma_{z_{w,t}}) \).

The amount of conventional energy (\( e_t \)) changes similar to the water in each time interval as follows:
where $e_t$ is the available energy at time step $t$, $d_{t,i}$ is the energy demand for crop $i$ at time step $t$. Changes in the available energy do not depend on seasonal changes in this case study due to lack of data, however, extension to include such seasonal dependence is straightforward.

Interconnections between water and energy resources are also similarly modeled as before. If any agent chooses to consume the recycled water $(1_{\text{Rec},}(a_{t,i}))$, $e_{w}$ amount of energy is consumed for recycling water (assumed to be 20% of the maximum amount), and if each agent utilizes the renewable energy source $(1_{\text{Ren}},(a_{t,i}))$, a deterministic boost in the amount of $e_t$ will be added to the available energy for crop production (assumed to be 25%). Similar to equation 1, $c_{t,i}$ takes into account the stochasticity that is not modeled directly due to lack of data.

The seasonal water and energy demands for these four crops in Ventura County are obtained from Bell et al (2018). Moreover, the costs, $u_i(s, a_i)$ associated with the actions of each agent $(a_i)$ are comprised of energy cost, GHG emissions (environmental costs) and operational costs and are obtained from Memarzadeh et al (2019). Although the costs associated with actions of the agents are formulated individually, they depend on the state of the system $(S = W \times E)$, which is changing according to actions taken by all agents. This implies that the ‘fate’ of the agents as well as the FEWS is coupled in this process, even though their choices of actions are independent of one another. Each agent has to come up with a strategy $\pi^t$ for managing her crop, which identifies what sources of water and energy resources to utilize depending on the current status of available water and energy and season, $\pi^t : (S, \Lambda) \rightarrow \Lambda_i$. For example, an ignorant agent might always utilize the conventional sources of water and energy independent of the current state and season, $\pi^{\text{ignorant}} = \{\text{Conv}_{w}, \text{Conv}_e\}$. Once each agent has identified her strategy, $\Pi_i = \{\pi^t_i\}$, the expected overall cost of managing the FEWS over its entire life cycle for each agent $i$ (depending on the current state, $s_t$, and season, $\lambda_t$), $V^\Pi_t(s_t, \lambda_t)$ can be calculated as:

\[
V^\Pi_t(s_t, \lambda_t) = u(s_t, \lambda_t, \pi^t(s_t, \lambda_t)) + \gamma \mathbb{E}_{s_{t+1} \sim f(s_t, \lambda_t, \pi(s_t, \lambda_t))} [V^\Pi_t(s_{t+1}, \lambda_{t+1})]
\]  

where $V \sim f(\lambda)$ is a stochastic function governing the seasonal changes, and $s' \sim f(s, \lambda, \pi(s, \lambda))$ is a stochastic function governing dynamics of the system’s state, as defined by equations (1-2). Depending on whether agents cooperate with each other or act in a non-cooperative manner, the optimum solution in multi-agent management (the optimal value in equation (3)) can switch from the best outcome for the whole system (the social optimum) to the non-cooperative solution (the Nash equilibrium). The social optimum in this formulation also corresponds to the Pareto optimal solution (Censor 1977). To better understand the difference between these two solutions, we illustrate it in a famous prisoner’s dilemma (Poundstone 1992) problem in the appendix A for an interested reader. In this article, (1) we develop a scalable algorithm for finding the non-cooperative solution (the Nash equilibrium), which is a solution that naturally occurs in the FEWS case study, (2) illustrate the difference between the non-cooperative solution and the social optimum, and (3) study how the utility functions of agents can be modified or incentivised in this case study so that the Nash equilibrium converges to the social optimum.

### 2.1. Finding the non-cooperative solution

As mentioned before, the non-cooperative solution is the Nash equilibrium in the multi-agent decision-making problem and is the natural solution in the FEWS example. The exact mathematical definition of the Nash equilibrium is outlined in appendix A. The equilibrium represents a set of management strategies for each agent $(\pi^t_i)$ for $i \in 1, ..., N$ and implies that it is not worthwhile for any of the agents to deviate from this equilibrium. As we illustrated on a simple example in appendix A, under this non-cooperative assumption, the Nash equilibrium is not necessarily the best strategy for the entire system. As we shall see later, if cooperation is allowed among the agents, the multi-agent management achieves the social optimum (the best outcome for the system). We will discuss this matter later with several examples to illustrate the difference between the social optimum and the Nash equilibrium.

In order to find the Nash equilibrium, we have developed a dynamic programming approach called Nash policy iteration, inspired by the Nash Q-learning algorithm (Hu and Wellman 2003). Figure 1 shows the complete algorithm for finding the Nash equilibrium in any multi-agent decision making scenario, where the dynamics of the system’s state and costs of actions for each agent can be formalized. In figure 1, $NashV^\Pi(s, \lambda)$ is the long-term total management cost of the Nash equilibrium solution at state $(s, \lambda)$ at iteration $w$. Given that the state is fixed, finding this value requires finding the Nash equilibrium of a general-sum (sometimes also called non-zero sum) normal-form game. This requires solving a non-convex non-linear optimization problem as depicted in appendix C. This is what the function SolveGame would perform, and we provide details of the optimization problem and how to find the Nash equilibrium to this problem in appendix C. An important implementation detail is that the quality of policies obtained by the algorithm depends on the stopping criteria $\varepsilon$ and $\text{thr}$. $\varepsilon$ provides the convergence property for evaluating an equilibrium, which is usually fixed at a very small number, $10^{-3}$. 

\[
e_{t+1} = e_t - \left[ \sum_{i \in \{1, ..., N\}} d_{t,i} \right] - e_w \cdot 1_{\text{Rec},}(a_{t,i}) + e_w \cdot 1_{\text{Ren},}(a_{t,i}) + c_{t,i}
\]
Figure 1. The Nash policy iteration algorithm for finding the non-cooperative solution (the Nash equilibrium).

thr controls the quality of the final selected policy at the Nash equilibrium. We have fixed thr at 10%, which means that if the policy for 90% of the state space does not deviate from the previous iteration, then we consider the policy to be converged to the Nash equilibrium.

2.2. Finding the social optimum

The social optimum solution means finding the joint strategy that is the best for the system in the long term, if none of the agents deviate from this strategy and if they all act in a cooperative manner. It is aligned with the definition of Pareto optimal solution in multi-objective problems (Censor 1977). The difference between the Nash equilibrium and the social optimum is that in Nash, agents are not aware of each others actions and try to find a strategy that is the best response to the best strategies of others, given what they know about the utilities/costs of other agents and dynamics of the system. On the other hand, in the social optimum, the assumption is that agents coordinate on what strategy to take, which would lead to the best outcome for the entire system (and not necessarily for each agent individually). This means that although the Pareto optimal solution might not be unique, the social optimum is always unique (Bellman 1957) and corresponds to the solution that is the best for the system according to the definition of the utility function. Social optimum can be found easily by framing the multi-agent problem as a single-agent decision optimization with an action profile that consists of actions of all managers involved in the process. Memarzadeh et al (2019) provides details of how to find a solution in the single-agent decision optimization problem.

3. Results and discussion

In this section we evaluate the performance of the proposed multi-agent management on two examples of real-world interdependent FEWS operations in Ventura County, California. We obtained data of the major agricultural operations specifically for four main crops in the region—strawberry, lemon, avocado, and celery - which on average account for 33% of California’s total crop production and 29.5% of US
production, with a gross value of B $ 1.18 (Ross 2015). We assume that each crop’s production is managed by a different agent, so in this case study we have four agents (N = 4), one for each crop. Since the demands of water and energy for production of these crops are different in each season (Bell et al 2018), we expect the non-cooperative decision making (the Nash equilibrium) to be significantly different with respect to the cooperative decision making (the social optimum). The reason for this is that in the non-cooperative setting, each agent is only concerned with minimizing the costs of producing her own crops, while in the cooperative setting, the goal is to minimize the cost of producing entire crops.

Figure 2 visualizes the individual policies implemented for each crop if each agent independently had access to the entire amount of water and energy available in the region. As can be seen, the policies are significantly different from crop to crop, which is due to the different seasonal demand of water and energy for production. Axes correspond to the available amount of conventional energy and water resources, and different shapes denote different management actions. The general trend is that managers tend to utilize recycled water (green triangle and magenta cross) more aggressively in the high water-demand seasons compared to low water-demand seasons. For example, in the case of strawberry, the manager uses the renewable water source 44% more in high water-demand seasons compared to low water-demand seasons (these differences are 64% for lemon, 20% for avocado, and 70% for celery).

3.1. Simplified example

In this section, we comprehensively compare the management strategy suggested by the Nash equilibrium and by the social optimum on a simplified FEWS operation. This simplified example will help us build insights about the convergence of the Nash equilibrium to the social optimum that can be further clarified later. We assume that the amount of energy available for food production is infinite, hence the shared and deteriorating state corresponds only
to conventional water, \( S = w \). We discretize the normalized available water resource, \( w \in [0,1] \) into 21 states with a step of 0.05. Only two crops, strawberry and lemon are considered here, so there are two agents involved in the decision-making process. The action profiles of each agent are as discussed before, \( A_i = \{A_{w_i}^S, A_{e_i}^S\} \), where \( A_{w_i}^S \in \{\text{Conv}_{w_i}, \text{Rec}_{w_i}\} \) and \( A_{e_i}^S \in \{\text{Conv}_{e_i}, \text{Ren}_{e_i}\} \). The exogenous variable \( \lambda \) correspond to the seasonal changes, \( w = e = 20\% \), \( w = e = 25\% \), and \( \sigma_c = 5\% \). Each time step is assumed to be one day to characterize the seasonal changes.

Figure 3 compares the average implemented action based on 100 independent forward simulations for managing the FEWS operations of the simplified example. Using the management strategy evaluation approach (Smith 1994), we are comparing the strategies of (1) the social optimum, which can be obtained by combining the decision-making of all agents into a single decision-making of one top-agent, which minimizes the costs of operation for the entire system (as described in section 2.2), (2) the Nash equilibrium, which can be obtained using the algorithm in figure 1, and (3) the individual strategies, which can be obtained by solving the problem for each agent independently, ignoring the effect of other agents in the decision-making process (Memarzadeh et al. 2019). We can observe that the implemented strategy by the Nash equilibrium is significantly close to the social optimum (the actions agree on average 76% of time), while on the other hand the individual strategies are significantly different (individual strategies only agree with the social optimum 40% of time on average). To calculate these numbers, we simulate the strategies identified by Nash equilibrium, social optimum, and individual policies in 100 independent forward simulations, and calculate the portion of times where these policies agree with one another in a similar state (similar season and available amount of conventional water and energy).

Figure 4 shows (A) the trajectory of the water state, as well as (B) the long-term discounted costs of managing the system, based on each implemented strategy for each season. As it can be seen, the individual approach results in significant depletion of the water state in low-rainfall seasons, Summer and Fall, resulting in accumulating penalties due to not meeting the water demand for food production and excessive costs of 200% and 225% for Summer and Fall compared to Spring, respectively. On the other hand, the most interesting result is that the strategy obtained by the Nash equilibrium performs identically to the social optimum in terms of cost (figure 4(B)), however, depletes the state of water more aggressively in Spring and Summer (figure 4(A)). This is an outstanding result, showing that under the presented formulation of FEWS in Ventura County for the simplified example, the Nash equilibrium solution is naturally converging to the social optimum. The reason for this natural conversion can be evaluated based on the additional cost of utilizing the recycled water resource for strawberry and lemon.

As reported in table B1 in Appendix B, on average, the extra cost of utilizing recycled water for strawberry and lemon are 10% and 13%, respectively. Since these values are close to each other, we expect the outcome of non-cooperative decision-making to be close to the outcome of the cooperative setting. As we will see later, this is not the case when we evaluate this for all four crops in the region.

3.2. Complex FEWS example

In this section, we apply the proposed multi-agent management on the FEWS operation of four main crops in Ventura County, California. The level of complexity of this example is much higher than the previous example: first, we relax the assumption of the infinite energy source, and as a result the state space includes the available conventional energy resource as well, \( S = w \times E \); second, the problem involves four agents (one per crop), which contribute to the fate of the system through the strategies that they implement. We discretize both water and energy states each into 21 states, which result in total state space of 441 states, \( |S| = |W| \times |E| = 441 \). The action profile of each agent is the same as before: each agent has access to four different actions corresponding to utilizing the conventional or renewable sources of water and energy. Other parameters are fixed as before (refer to section 3.1). Figure 5 visualizes the results for management of the four crops based on strategies identified by (1) the social optimum, (2) the Nash equilibrium, and (3) the individual approach.

In this example, long-term expected cost of management is different with respect to figure 4, and the purpose is not to compare these two graphs with each other. The general trend of both social optimum and Nash equilibrium outperforming the individual approach still holds for this complicated example. However, we observe that as the problem complexity increases (and as a result the complexity of the management strategy increases), the performance of the Nash equilibrium diverges away from the social optimum (compared to the simplified example in section 3.1). On average, the Nash equilibrium results in excessive 110% cost of management compared to the social optimum, as illustrated in figure 5, while this excessive cost for the individual policies is 440%.

As discussed in section 3.1, this divergence of the Nash equilibrium from the social optimum can be studied in the extra costs associated for utilization of renewable water and energy resources for different crops. For example, in the FEWS case study in Ventura County, on average, the extra cost of utilizing the recycled water for crops of strawberry, lemon, avocado and celery are 10%, 13%, 20%, and 46%, respectively (table B1 in the appendix B). Let us focus on an example of strawberry and celery; looking at
Figure 3. This figure shows the average trajectory, based on 100 independent forward simulations, of implemented strategies for managing the system according to the social optimum, Nash equilibrium, and individual strategies.

Figure 4. This figure shows (A) trajectory of the water state for the system managed by strategies obtained by the social optimum, Nash equilibrium, and the individual agents, and (B) the corresponding long-term expected cost of managing the system.
the individual strategies of these agents (figure 2), we realize that it is not optimal for the strawberry agent to utilize the recycled water resource in Winter, due to high amount of rainfall and low seasonal demand for the production. Given this, in the non-cooperative setting (the Nash equilibrium), the best strategy for the strawberry agent is to utilize the conventional source of water. On the other hand, the celery agent, which has a higher demand of water for production in Winter, need to take advantage of the recycled water resource, despite significant higher cost of utilization compared to strawberry. As a result, the cost associated with managing the system in the non-cooperative fashion is much higher.

In the social optimum setting, there is no sense of individualism, and as a result all agents cooperate to achieve the best outcome for the entire system. Hence, although it is not optimal individually for the strawberry agent to utilize the recycled water resource in Winter, she deviates from her own optimal strategy and utilizes the recycled water, so the conventional resources of water can be used by those agents that have higher cost of utilizing the recycled water and higher demand of water in Winter. In this way, the total cost of managing the system is expected to be much lower than the non-cooperative setting, as illustrated by figure 5. Further research needs to be done to evaluate how the costs associated with utilization of the renewable energy and water resources can be revised or incentivized, so that the Nash equilibrium would naturally converge to the social optimum, a case that we observed happening in section 3.1.

4. Conclusions

We have developed a novel method for multi-agent management of food-energy-water systems (FEWS), basing upon fundamentals of decision and game theories. The proposed approach is able to incorporate stochasticity and uncertainty in the system’s dynamics and interdependence of the water and energy resources available for food production. We specifically focus on differentiating between (1) cooperative management, where all decision-makers cooperate to achieve the best outcome for the whole system, the social optimum, and (2) non-cooperative decision-making, the Nash equilibrium.

We illustrated the method on a real-world case study in Ventura County, California, and compared the benefits of adopting management strategies obtained by the Nash equilibrium with respect to individual management strategies, where the existence of other agents are ignored. Specifically we showed in figure 5 that the Nash equilibrium can perform significantly better than that individual management (110% of extra cost compared to the social optimum, compared to 440% extra costs in the case of individual management). We also discussed how the costs associated with utilization of the renewable energy and water resources can be revised or incentivized so that Nash equilibrium would naturally converge to the social optimum, which was illustrated in a simplified example as well (figure 4).

This article is the first attempt to deploy game theory for multi-agent management of the FEWS.
operations that is scalable and applicable to real-world scenarios. There are many other important factors that need to be integrated in such multi-agent management that were not studied in this article, including sectoral interdependence among the three sectors involved. Moreover, the current method only focuses on management at a crop level, however, one can extend the proposed framework to include hierarchical decision-making and sectoral decision-making at multiple scales. Another future direction for our effort is incorporating the effects of climate change and environmental variations. Although in this article we include effects of seasonal changes, the effect of climate variations such as rise in temperature or variations in precipitation are important factors that need to be incorporated in the future direction of this research.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1739676. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Data availability statement

The data that support the findings of this study are either included within the article or openly available from Bell et al (2018) and Memarzadeh et al (2019).

Appendix A. Illustration of Nash equilibrium and social optimum on prisoner’s dilemma

A.1. Nash equilibrium

In the non-cooperative management, where each agent selects her actions in a non-cooperative manner (and secretly), and is only interested in minimizing her own costs, the solution is known as Nash equilibrium, defined below:

**Definition 1.** We shall say that \( \{a_1^*(s), a_2^*(s), \ldots, a_N^*(s)\} \) is a Nash equilibrium if,

\[
\forall (a_1^*(o), \ldots, a_i^*(o), \ldots, a_N^*(o)) \in A_i \ni (s, \lambda) \ni \forall s \in S, \forall \lambda \in \Lambda, \forall a_i(s) \in A_i, \forall i \in \{1, 2, \ldots, N\}
\]

\[
(A1)
\]

The definition of Nash equilibrium implies that it is not economically worthwhile for any of the agents to deviate from this equilibrium.

A.2. Prisoner’s dilemma

In this section, we illustrate the difference between Nash equilibrium and the social optimum solution on a well-known and simple problem of prisoner’s dilemma. In this problem, two persons are being interrogated as suspects of a crime. Each of the suspects can either Confess (denoted as C) or Not Confess (denoted as NC), so \( A_1 = A_2 = \{C, NC\} \). There is no communication among the two suspects, while they are both aware of their own and other player’s payoff, which reported in the payoff bi-matrix in figure A1.

The values in parentheses represent the payoff to player 1 and 2, respectively, under each possible strategy. Please note that this is not a stochastic game, rather a static normal-form game, where the game is played in only one state, and terminates afterward. To better understand the payoff in the notations used in the manuscript, when both players choose to not confess, \( a_1 = a_2 = NC \), then \( Q_1(s, a_1, a_2) = Q_2(s, a_1, a_2) = -3 \).

As it can be seen, if suspect 1 confesses, and suspect 2 does not confess, then suspect 1 goes to jail for 4 years while suspect 2 goes free. The same scenario applies to suspect 2 as well. Now the interesting part is that if both suspects confess, then they both go to jail for 1 year, while if both do not confess, they both go to jail for 3 years. It is clear by looking at the payoff bi-matrix that the social optimum solution, defined as the best outcome of the game for both players, is for both of them to confess and go to jail for one year. This corresponds to the social optimum solution (best outcome for all) and also the Pareto optimal solution (Censor 1977). However, due to the fact that the suspects do not cooperate with each other, if you solve either the optimization problem of equation (C3), or solve the problem using the iterative algorithm in figure C2 in the appendix C, the Nash equilibrium of the game (or the natural solution) would be \( a_1 = a_2 = NC \) with the value of \(-3\) for both suspects (the results of optimization would be \( x_1^* = x_2^* = 0 \), \( p^* = q^* = -3 \)). The reason for this dilemma lies in the definition of the Nash equilibrium, formalized in *Definition 1*. The best response for each player to the best strategy of the other player is to not confess. This is a very simple example that

\[
\begin{array}{ccc}
\text{Suspect 1} & a_2 = C & a_2 = NC \\
\hline
\text{Suspect 2} & a_1 = C & (-1, -1) & (-4, 0) \\
& a_1 = NC & (0, -4) & (-3, -3)
\end{array}
\]

Figure A1. The payoff matrices of both players in the prisoner’s dilemma game.
Memarzadeh

1 are the utility matrices

\[ \{ -A_S \text{ and } = \} \]

\( \forall x \in \mathbb{R} \)

\( \sum \in \mathbb{R} \)

\[ \lambda \in \{ \text{Spring, Summer, Fall, Winter} \} \]

\( d \)

\( r \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\( \in \mathbb{R} \)

\[ \max_{x_i \in \Lambda_i} \{ Q_i | A_i \}, x_{i,-i} \}

\[ Q_{-i}x_{i,-i} \leq p_{|A_i|} \]

\[ Q_{-i}x_{i,-i} \leq q_{|A_i|} \]

\[ \sum x_{i,j} = 1, \sum x_{i,-i} = 1 \]

\[ x_{i,j} \geq 0, \forall j \in \{ 1, \ldots, |A_i| \}, x_{i,-i} \geq 0, \forall j \in \{ 1, \ldots, |A_{-i}| \} \]

\[ (C3) \]

\[ (C2) \]

\[ \text{Iterative method of solving a game} \]

\[ \text{Inputs: } Q_i(s, \lambda, a_i), s \in S, \lambda \in \Lambda, a_i \in A_i, \forall i \in \{ 1, 2, \ldots, N \} \]

\[ \text{Initialize } \delta_i \in A_i \text{ randomly, } \forall i \in \{ 1, 2, \ldots, N \} \]

\[ \forall k \in \{ 0, 1, \ldots, \}, \forall i \in \{ 1, 2, \ldots, N \}: \]

\[ \delta_i^{k+1}(s, \lambda) = \arg\max_{a_i} Q_i(s, \lambda, a_i) \]

\[ \Delta_i = \sum_{(s, \lambda) \in \Lambda} \left( \delta_i^{k+1}(s, \lambda) - \delta_i^k(s, \lambda) \right) \]

\[ \text{Until } \Delta_i < \text{thr} \]

\[ \text{Output the policy and value at equilibrium:} \]

\[ \pi_i^*(s, \lambda) = \left[ \delta_i^{k+1}(s, \lambda), \ldots, \delta_i^0(s, \lambda) \right] \]

\[ \text{NashV}_i^*(s, \lambda) = Q_i(s, \lambda, \pi_i^*(s, \lambda)) \]

\[ \text{Figure C2. The iterative algorithm for solving the general-sum normal-form game.} \]


to be consistent with the extension to more than 2 players,

\[ Q_i(s, \lambda, a_i, a_{-i}) = u_i(s, a_i, a_{-i}) + \gamma \sum_{\lambda^i \in \Lambda} p(\lambda^i | \lambda) \sum_{s' \in S} p(s' | s, \lambda^i, a_i, a_{-i}) \text{NashV}_i^*(s', \lambda^i) \]

\[ (C2) \]

where \( s \in S, \lambda \in \Lambda, a_i \in A_i \), and \( a_{-i} \in A_{-i} \). Defining Q-values, then \( Q_i \) and \( Q_{-i} \) are the utility matrices for agent \( i \) and \( i' \), respectively, depending on different actions that they can take. Now finding the Nash equilibrium of such general-sum game requires solving the following non-convex non-linear optimization problem (Mangasarian and Stone 1964, Filar and Vrieze 1997),

\[ \max_{x, x_{-i} \in \mathbb{R}^{|A|}} x_i T x_{i,-i} + x_{i,-i} x_{-i} T - p - q \]

\[ \text{s.t. } Q_i x_i \leq p |A_i| \]

\[ Q_{-i} x_{i,-i} \leq q |A_i| \]

\[ \sum x_{i,j} = 1, \sum x_{i,-i} = 1 \]

\[ x_{i,j} \geq 0, \forall j \in \{ 1, \ldots, |A_i| \}, x_{i,-i} \geq 0, \forall j \in \{ 1, \ldots, |A_{-i}| \} \]

\[ (C3) \]

illustrates how the Nash equilibrium is different with respect to the best outcome of the game, the social optimum.

Appendix B. Additional figures and tables

Appendix C. Finding the Nash equilibrium of a general-sum normal-form game

In this section, we provide details of how to find the Nash equilibrium of a general-sum normal-form game, which is a requirement for the algorithm in figure 1 (function SolveGame). For simplicity of the notations, we focus on the problem with only two agents here, \( \{ i, \neg i \} \), but the formulations easily generalize to more than two agents. Let us assume that \( A_i \) and \( A_{-i} \) are the action spaces of agents \( i \) and \( \neg i \), respectively, and define the Q-value for agent \( i \) as follows (we have assumed that the other agent is \( \neg i \))

| Crop          | Energy | GHG   | Operation
|---------------|--------|-------|-----------
| Strawberry    | 10%    | 14%   | -7%       |
| Lemon         | 12%    | 7%    | 22%       |
| Avocado       | 17%    | 9%    | 34%       |
| Celery        | 54%    | 59%   | 25%       |

Table B2. Variable used in this article and their definition.

| Variable | Definition                          |
|----------|-------------------------------------|
| s \in S  | Entire domain of state variables.  |
| w \in W  | Available conventional source of water for crop production. |
| e \in E  | Available conventional source of energy for crop production. |
| N        | Number of agents involved in the decision making process. |
| a_i \in A_i | Entire domain of actions available to agent i. |
| \lambda | Exogenous variable defining seasonal changes. |
| d        | Variable representing demands of water and energy. |
| r        | Exogenous variable defining precipitation. |
| \zeta | Variable defining stochasticity. |
| \pi | Management strategy chosen for the system by agent i. |
| u_i \in U | Utility variable quantifying the quality of manager i’s actions. |
| V_i      | Long-term expected cost of managing the system for agent i. |
| NashV    | Long-term expected cost of managing at the Nash equilibrium. |
| \gamma   | Discount factor, relating future costs to their net present value. |

Table B1. Additional costs associated with using the recycled water resource in terms of energy cost (MJ/kg of the crops produced), GHG emissions (kgCO2/kg of the crops produced), and operational costs ($/kg of the crops produced) (source: (Bell et al 2018)).

| Crop          | Energy | GHG   | Operation |
|---------------|--------|-------|-----------|
| Strawberry    | 10%    | 14%   | -7%       |
| Lemon         | 12%    | 7%    | 22%       |
| Avocado       | 17%    | 9%    | 34%       |
| Celery        | 54%    | 59%   | 25%       |
solving the game (Robinson 1951), as illustrated in figure C2.

**ORCID iD**

Milad Memarzadeh  [https://orcid.org/0000-0001-7672-6033](https://orcid.org/0000-0001-7672-6033)

**References**

Albrecht T R, Croofof A and Scott C A 2018 The water-energy-food nexus: a systematic review of methods for nexus assessment *Environ. Res. Lett.* 13 043002

Bell E M, Stokes-Draut J R and Horvath A 2018 Environmental evaluation of high-value agricultural produce with diverse water sources: case study from Southern California *Env. Res. Lett.* 13 025007

Bellman R E 1957 *Dynamic Programming* (Princeton, NJ: Princeton University Press)

Bieber N, Ker J H, Wang X, Traintafyllidis C, van Dam K H, Koppelaar R H E M and Shah N 2018 Sustainable planning of the energy-water-food nexus using decision-making tools *Energy Policy* 113 584–607

Censor Y 1977 *Pareto optimality in multiobjective problems* Appl. Math. Optim. 4 41–59

Chaudhary A, Gustafson D and Mathys A 2018 Multi-indicator sustainability assessment of global food systems *Nat. Commun.* 9 848

Daher B, Hannibal B, Portney K E and Mohr R H 2019 Toward creating an environment of cooperation between water, energy, and food stakeholders in San Antonio *Sci. Total Environ.* 651 2013–26

Filare J and Vreeke K 1997 *Competitive Markov Decision Processes* (New York: Springer)

Givens J E, Padowski J, Guzman C D, Malek K, Witink-Huber R, Cosens B, Briscoe M, Boll J and Adam J 2018 Incorporating social system dynamics in the Columbia River basin: food-energy-water resilience and sustainability modeling in the Yakima River basin *Front. Environ. Sci.* 6 104

Hu J and Wellman M P 2003 *Nash Q-learning for general-sum stochastic games* *Machine Learning Res.* 4 1039–69

Karan E, Asadi S, Mohr R and Baadwin M 2018 Towards the optimization of sustainable food-energy-water systems: a stochastic approach *J. Cleaner Production* 171 662–74

Karan E and Asadi S 2018 Quantitative modeling of interconnections associated with sustainable food, energy and water (FEW) systems *J. Cleaner Production* 200 86–99

Kurian M, Portney K E, Rappold G, Hannibal B and Ghebrechorkos S Het al 2018 *Governance of Water-Energy-Food Nexus: A Social Network Analysis Approach to Understanding Agency Behaviour* Hulsman S, Ardakanian R (Eds) *Managing Water, Soil and Waste Resources to Achieve Sustainable Development Goals* (Berlin: Springer) pp 125–47

Liang S, Qu S, Zhao Q, Zhang X, Daigger G T, Newell J P and Mille S A et al 2019 Quantifying the urban food-energy-water nexus: the case of the Detroit metropolitan area *Environ. Sci. Technol.* 53 779–88

Liu Y, Hejazi M, Kyle P, Kim S H, Davies E, Miralles D G, Teuling A J, He Y and Niyogi D 2016 Global and regional evaluation of energy for water *Environ. Sci. Technol.* 50 9736–45

Madani K 2010 *Game theory and water resources* *J. Hydrol.* 381 225–38

Mangasarian O L and Stone H 1964 *Two-person nonzero-sum games and quadratic programming* *J. Math. Anal. Appl.* 9 384–55

Memarzadeh M, Moura S and Horvath A 2019 Optimizing dynamics of integrated food-energy-water systems under the risk of climate change *Environ. Res. Lett.* 14 074010

Namany S, Al-Ansari T and Govindan R 2019 Sustainable energy, water and food nexus systems: a focused review of decision-making tools for efficient resource management and governance *J. Cleaner Production* 225 610–26

Poundstone W 1992 *Prisoner’s Dilemma* (New York: Anchor Books)

Ramaswami A, Boyer D, Nagpure A S, Fang A, Bogra S, Bakshi B, Cohen E and Rao-Ghorpade A 2017 An urban systems framework to assess the trans-boundary food-energy-water nexus: implementation in Delhi, India *Environ. Res. Lett.* 12 025008

Robinson J 1951 An iterative method of solving a game *Ann. Math.* 54 296–301

Ross K 2015 *California Agricultural Statistics Review, 2014–2015* (Sacramento, CA: California Department of Food Agriculture)

Sherwood J, Clebeaux R and Carbajales-Dale M 2017 An extended environmental input–output lifecycle assessment model to study the urban food–energy–water nexus *Environ. Res. Lett.* 12 105003

Smith A D M 1994 Management Strategy Evaluation—the Light on the Hill *In Population Dynamics for Fisheries Management* Ed D A Hancock (Perth: Australian Society for Fish Biology) 249–53

Tsolas S D, Karim M N and Hasan M M F 2018 Optimization of water-energy nexus: a network representation-based graphical approach *Appl. Energy* 224 230–50

Wang S, Cao T and Chen B 2017 Urban energy–water nexus based on modified input–output analysis *Appl. Energy* 196 208–17

Zhang J, Campana P E, Yao T, Zhang Y, Lundblad A, Melton F and Yan J 2018 The water-food-energy nexus optimization approach to combat agricultural drought: a case study in the United States *Appl. Energy* 227 449–64