Valuing water quality tradeoffs at the farm level: An integrated approach

Moriah Bostian*[1], Gerald Whittaker[2], Brad Barnhart[2], Rolf Färe[3], and Shawna Grosskopf[3]

[1]*Department of Economics, Lewis & Clark College, Portland, OR USA*

[2]*Agricultural Research Service (USDA), Corvallis, OR USA*

[3]*Department of Economics, Oregon State University, Corvallis, OR USA*

September 17, 2014

Abstract

This study evaluates the tradeoff between agricultural production and water quality for individual producers using an integrated economic-biophysical hybrid genetic algorithm. We apply a multi-input, multi-output profit maximization model to detailed farm-level production data from the Oregon Willamette Valley to predict each producer's response to a targeted fertilizer tax policy. Their resulting production decisions are included in a biophysical model of basin-level soil and water quality. We use a hybrid genetic algorithm to integrate the economic and biophysical models into one multiobjective optimization problem, the joint maximization of farm profits and minimization of Nitrate runoff resulting from fertilizer usage. We then measure the tradeoffs between maximum profit and Nitrogen loading for individual farms, subject to the fertilizer tax policy. We find considerable

*Corresponding Author. Email: mbbostian@lclark.edu; Tel.: 1(503)768-7626

variation in tradeoff values across the basin, which could be used to better target incentives for reducing Nitrogen loading to agricultural producers.

Keywords: Agri-Environmental Policy, Integrated Modeling, Tradeoff Analysis, Pollution Tax, Nitrogen Loading, Genetic Algorithm

JEL Codes: Q15, Q51, Q53, Q58

1 Introduction

Environmental policy analysis, particularly the analysis of policies that are targeted to a distinct group of decision makers or to a particular geographic region, can be refined by integrating economic and biophysical models. Examples of integrated economic and biophysical models for agriculture include modeling the biophysical outcomes of alternative economic scenarios (Secchi and Babcock, 2007; Jha et al., 2010) or the solution to a single-objective economic optimization model (Schönert et al., 2011; Uthes et al., 2010) and linking both single and multiobjective economic optimization models to biophysical models in a model chain (Hillyer et al., 2003; Liu et al., 2002; Moore and Tindall, 2005; Rabotyagov et al., 2010c; Volk et al., 2008; Whittaker et al., 2005).

In the model chain approach, information passes only in one direction, so that the optimal decision at any point in the chain is constrained by any previous decisions or outcomes in the chain. A simultaneous optimization of all objectives can inform the calculation of tradeoffs between multiple objectives. Several studies employ genetic algorithms to simultaneously optimize multiple objectives by allowing information to pass between each objective in both directions (Bekele and Nicklow, 2005; Arabi et al., 2006; Rabotyagov et al, 2010a; 2010b). These studies illustrate the use of genetic algorithms to calculate the Pareto optimal frontier for both economic and environmental objectives.

We build on the use of genetic algorithms for agri-environmental policy analysis by integrating a realistic biophysical model with a detailed economic optimization model that more fully endogenizes each producer's response to the search for an optimal targeted policy. Our use of genetic algorithm methods to more freely integrate the economic and biophysical models is detailed in a related study of targeted policy design (Whittaker et al., 2014).

This approach contributes to the existing literature in several important ways. First, we include both a detailed, spatially explicit biophysical model and a complete model of profit maximization, with minimal restrictions to solution values and without imposing a production technology relationship. Second, we apply an adaptive modeling framework to allow for two-way feedback between our economic and environmental objectives. This framework fully endogenizes fertilizer usage, making economic cost endogenous and better updating the search for the efficient tax rate. The resulting policy generates a set of Pareto optimal tradeoffs that can be evaluated across objectives. Third, we evaluate the resulting tradeoffs for individual producers, in addition to the aggregate basin-level tradeoffs.

This integrated economic-biophysical model simulates a rich set of agent-level decisions, made in response to the Pareto optimal policy and corresponding environmental outcomes that can be used to evaluate tradeoffs at the individual level. We examine these decisions for a set of grass seed farms situated in the Calapooia River watershed, a predominantly agricultural watershed in Oregon's Willamette Valley. We also make use of detailed microlevel farm production data, which further enhances the evaluation of individual tradeoffs between farm profit maximization and watershed Nitrogen loading.

To value these tradeoffs, we jointly model the profit-maximizing crop production and Nitrogen loading levels, simulated by the economic-biophysical model, as outputs in a production process using the directional output distance function (Chambers et al., 1996). In economic production theory, the directional output distance function is dual to the revenue function, which we exploit to derive shadow price estimates for Nitrogen loading in the basin (Ball et al., 2004; Färe et al., 2005; 2006).

We find that the tradeoff between farm profit and Nitrogen loading varies greatly across farmers in the watershed, most likely due to differences in soil quality and location in the basin's hydrologic network. In practice, managers could use this information to target incentives for fertilizer reduction or reduced nitrate runoff, such as easement payments or funding for best management practices, to farms that have a lower opportunity cost of reducing eventual Nitrogen loading in the basin. Randhir and Shriver (2009) demonstrate the potential gains from using multi-attribute shadow price values to target restoration incentives across a watershed. Moreover, analysis of the tradeoff at the farm level offers a better picture of the distribution of costs across producers in the region. This distribution may be of concern for equity considerations and could affect the feasibility of implementing prospective agri-environmental policies in practice.

2 Multiobjective Optimization Problem

We characterize the joint, and often competing, objectives of farm-level profit maximization and basin-level Nitrogen loading as a multiobjective optimization problem. These objectives are constrained by the farm production technology and by the biophysical processes that determine the fate and transport of Nitrogen through the basin. The solution includes an optimal fertilizer usage and 'green tax' rate targeted to the farm level.

We use a hybrid genetic algorithm (HGA), following Whittaker et al. (2014), to solve over both objectives. We then use the solution values to estimate a frontier for crop production and Nitrogen loading that allows us to measure the economic and environmental tradeoffs for individual farms in the basin.

Because the solution set of optimal tax rates and fertilizer usage depends on the profit maximizing behavior of individual producers, we formulate the multiobjective problem as a bilevel optimization (Bard, 1998). A bilevel optimization nests one optimization inside of another, so that the solution to the outer non-nested optimization, typically referred to as the upper level, depends on the solution to the inner nested optimization, typically referred to as the lower level (Sinha et al., 2013). In our case, the joint maximization of total profit and minimization of basin-level Nitrogen constitutes the upper level while producer level profit maximization makes up the lower level.

For tax rate t and fertilizer input x_N, we represent the nested nature of this problem in general form, following Sinha et al. (2013) as

$$
\begin{aligned}
&\max_{t,\, x_N} \ \mathrm{F}(t,\, x_N) = [\pi(t,\, x_N),\, -N(x_N)] \qquad\qquad (1)\\
&\text{s.t.} \ \ x_N \in \underset{x_N}{\operatorname{argmax}} \begin{cases} \pi(t,\, x_N) = py(x_N) - w_N x_N - t w_N x_N \\ \pi(t,\, x_N) \geq 0 \end{cases}\\
&\qquad x_N \geq 0,\, t \geq 0,
\end{aligned}
$$

where N denotes basin-level Nitrogen loading, π represents farm profit, and w_N is the market price for fertilizer.

We note several important points underlying this general representation. First, the optimal tax rates and fertilizer usage for total profit and Nitrogen loading at the upper level depend on how individual producers respond to the tax, in terms of fertilizer use, at the lower level. The profit-maximizing fertilizer usage, in turn, depends on the production technology. Second, total Nitrogen loading at the upper level also depends on individual fertilizer usage in response to the tax at the lower level, as well as the spatial distribution of fertilizer usage by producers in the watershed. The spatial dynamics of fertilizer usage and Nitrogen loading are governed by biophysical processes in the basin. Third, the nested nature of this problem, coupled with multiple production inputs and many profit-maximizing producers, makes the solution to (1) complex. We employ a hybrid genetic algorithm to iteratively optimize the lower and upper levels of our problem. We explain the production technology specification, biophysical model and genetic algorithm solution method in more detail below.

2.1 Profit maximization at the farm level

In our economic model, each producer chooses inputs and outputs to maximize profit subject to the production technology and the fertilizer tax rate policy. We use nonparametric linear programming methods known as data envelopment analysis (DEA) (Charnes et al., 1978) to estimate the production technology and to simulate the profit maximization decision of each farm. In the DEA representation of the production technology, each of the K producers uses inputs $x = (x_1, ..., x_N)$ to produce outputs $y = (y_1, ..., y_M)$. The production technology T is defined as $T = \{(x,\, y) : x \,\text{can produce}\, y\}$.

The corresponding DEA representation of the technology is

$$
\begin{aligned}
T = \{(x,\, y) : y_m \leq & \sum_{k=1}^{K} z^k y_m^k, && m = 1, ..., M, \\
x_n \geq & \sum_{k=1}^{K} z^k x_n^k, && n = 1, ..., N, \\
& \sum_{k=1}^{K} z^k \leq 1, \\
& z^k \geq 0\},
\end{aligned} \tag{2}
$$

where the variables z^k, known as intensity variables in this framework, are constrained to allow for non-increasing returns to scale. Given input prices $w = (w_1, ..., w_N)$ and output prices $p = (p_1, ..., p_M)$, the profit of the k^{th} farm is computed as the solution to

$$
\begin{aligned}
\pi^k(p,\, w) = \max & \sum_{m=1}^{M} p_m y_m^k - \sum_{n=1}^{N} w_n x_n^k, && && (3)\\
\text{s.t.} \quad & \sum_{k=1}^{K} z^k y_m^k \geq y_m, && m = 1, ..., M, \\
& \sum_{k=1}^{K} z^k x_n^k \leq x_n, && n = 1, ..., N \\
& \sum_{k=1}^{K} z^k \leq 1, \\
& z^k \geq 0, && k = 1, ..., K.
\end{aligned}
$$

Figure 1 illustrates profit maximization for a DEA representation of a single input / single output production technology with three observations, a, b and c. These frontier observations also lie on the profit lines, π_1^*, π_2^* and π_3^*, which represent maximum profit levels for input and output prices $\{(p_1, w_1), (p_2, w_2), (p_3, w_3)\}$.

To simulate each producer's response to a 'green' tax policy, we add a targeted proportional tax to the profit maximization model in (3). The objective function under the targeted tax, t^k, on Nitrogen fertilizer, the N^{th} input, is

$$
\pi^k(p,\, w) = \max \sum_{m=1}^{M} p_m y_m^k - \sum_{n=1}^{N-1} w_n x_n^k - t^k w_N x_N^k, \tag{4}
$$

subject to the technology representation in (2) and (3). Here the tax rate for each farm, t^k, is multiplied by the quantity and price of the Nth input,

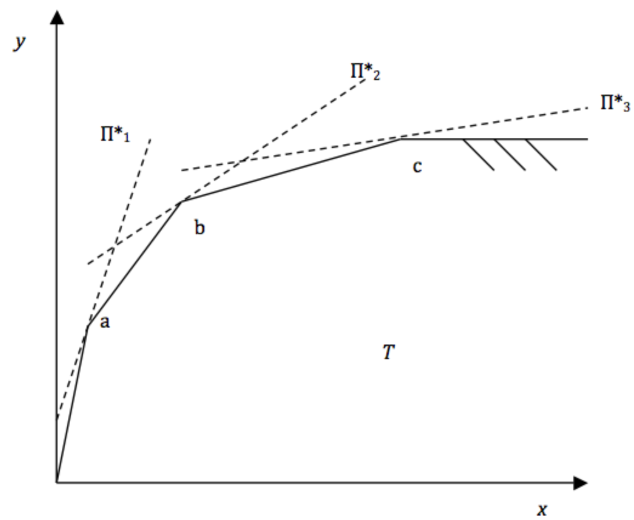

Figure 1: DEA Profit Maximization for Three Different Price Ratios

Nitrogen fertilizer. Note that a tax value of $t = 1$ is equivalent to having no tax on fertilizer and that a given policy consists of K different tax rates for each of the K farms.

2.2 Nitrogen loading in the basin

The environmental objective in this case is to minimize Nitrogen loading in the basin resulting from profit-maximizing fertilizer use. We use the Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998) to specify the environmental objective. SWAT is a biophysical model that can be used to simulate the effects of agricultural production processes at the river basin scale (Arnold et al., 2012). The model divides the entire watershed into subbasins, where each subbasin is further divided into hydrological response units (HRUs), which represent unique combinations of topography, land use and soil properties. Farm-level production decisions in each of the HRUs can then be included to model the spatial distribution of Nitrogen loadings throughout the watershed.

We use the digital elevation model ArcSWAT, which adds a GIS interface to SWAT, to input and designate land use, soil, weather, groundwater, water use management, pond and stream water quality data. SWAT simulates hydrology, soil erosion, plant growth, as well as multiple fate and transport processes, including that of Nitrogen. This framework is specifically designed to simulate the environmental effects of agricultural production practices, thus providing a method to test the effectiveness of agri-environmental policy (Arnold et al., 2012). SWAT is widely used and numerous studies apply it specifically to agri-environmental policy analysis (Bekele and Nicklow, 2005; Richardson et al., 2008; Rabotyagov et al, 2010b)

2.3 The Hybrid Genetic Algorithm

We use a genetic algorithm to solve the multiobjective optimization problem for the case of a targeted environmental policy, in this case a proportional Nitrogen fertilizer tax. This problem is computationally intensive, but relatively easy to implement with parallel execution (Whittaker et al., 2009).

A genetic algorithm (GA) is an iterative algorithm based on retention of the best or ‘fittest’ members of a population until a stopping condition is satisfied (Goldberg, 1989). In an optimization application, the GA consists of an initial randomly generated population that is evaluated for fitness using an objective function, a test for convergence, and application of the GA operations of selection, crossover and mutation. These elements are followed iteratively until an optimum is obtained.

Although GAs generally find promising solution regions quickly, convergence to an optimum can be much slower. In response, a hybrid genetic algorithm (HGA) model adds a local search method to speed convergence (Sinha and Goldberg, 2003). Figure 2 illustrates the HGA used to solve our maximum-profit and minimum-Nitrogen loading problem.

We use the non-dominated sort genetic algorithm (NSGA-II) (Deb et al., 2002) to assign a fitness value to each individual in the GA population, based on the evaluation of the individual for each objective. The result is an estimate of the Pareto optimal set of our objectives, farm profit and Nitrogen loading, at convergence. In our case, a linear program for the DEA model is solved in the evaluation step, which limits the space that is searched by the GA. The DEA results are then passed to NSGA-II, which finds the set of values available across the Pareto optimal frontier. It is important to note that this HGA uses information from both the economic and environmental models used in the integrated simulation of the tax policy during the optimization. Whittaker et al. (2014) provide more computational detail on implementing the HGA.

3 Evaluating the Individual Tradeoffs

The HGA is specified to maximize total basin-wide profit while also minimizing total basin-wide Nitrogen loadings. However, individual tax rates are applied to each farm. Therefore, for this targeted tax policy, it is also

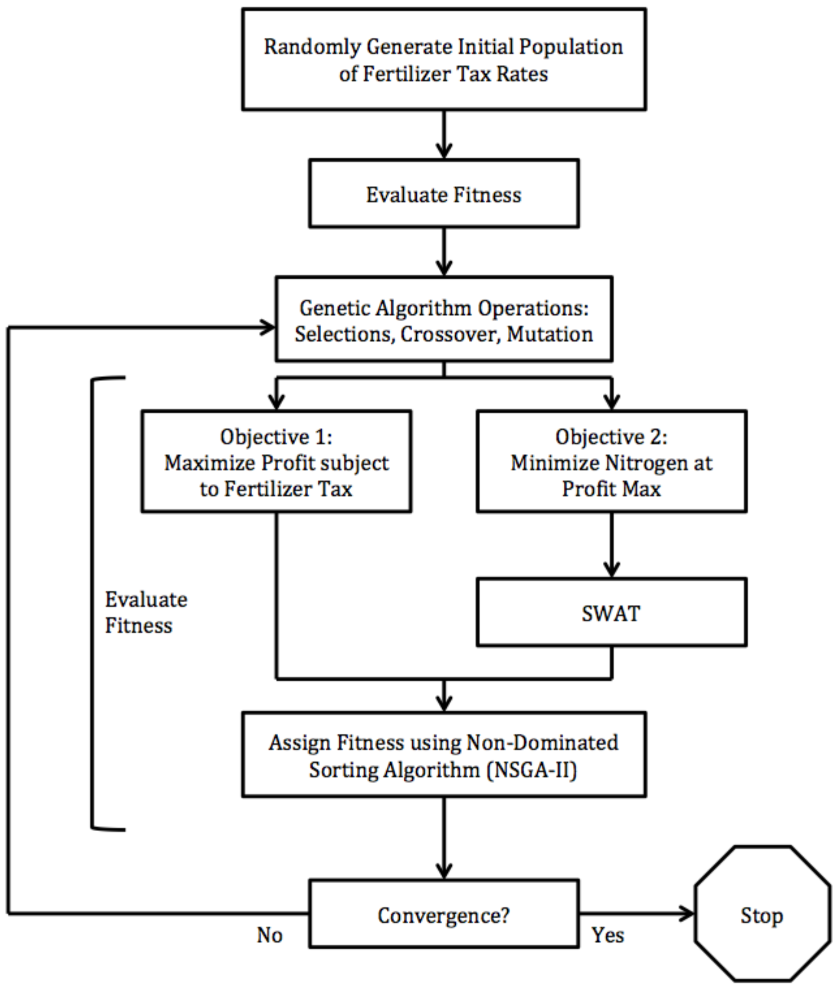

Figure 2: The hybrid genetic algorithm

important to understand the tradeoffs that exist for individual producers. To evaluate the tradeoff between Nitrogen loading and crop production at the farm level, we first calculate each farm's share of total basin Nitrogen loading as a function of their fertilizer application rate and HRU location. We then use a directional distance function approach to model individual Nitrogen loading as an undesirable output, produced jointly with the desirable output, crop production.

3.1 The underlying theory

We let $P(x)$ denote the feasible output set for the vector of farm outputs $y = (y_1, ..., y_M)$ and undesirable outputs $u = (u_1, ..., u_J)$ given inputs $x = (x_1, ..., x_N)$, so that

$$P(x) = \{(y,\, u) : x \, \text{can produce} \, (y,\, u)\} \, . \tag{5}$$

In this case, y represents each farm's crop production output, u its Nitrogen loading and x the vector of inputs, including acreage, labor, equipment and fertilizer.

We make the standard assumption that $P(x)$ is compact and convex, acknowledging that output is scarce and thus, tradeoffs exist at the frontier. We also assume that good and bad outputs are weakly disposable, which allows for their proportional scaling up or down over $P(x)$, meaning that for $(y,\, u) \in P(x) \, \text{and} \, 0 \leq \theta \leq 1, (\theta y, \theta u) \in P(x)$. We relax the usual assumption of null jointness, that if $(y,\, u) \in P(x) \, \text{and} \, u = 0$, then $y = 0$, due to its violation in practice by one of the farms in our study. Given these assumptions, we use the directional output distance function to represent the feasible output set (Chambers et al., 1996), as well as individual measures of performance. Figure 4 illustrates the feasible output set for the joint production of good and bad output and the directional output distance function, defined as

$$\vec{D_O}(x,\, y,\, u; g_y,\, g_u) = max\left\{\beta : [(y + \beta g_y,\, u - \beta g_u)] \in P(x)\right\}, \qquad (6)$$

where ($g_y \in \Re^M_+$, $g_u \in \Re^J_+$) is a directional vector that specifies the simultaneous expansion of desirable output and contraction of undesirable output This model measures each observation's distance, in a particular direction, to the production frontier. Thus, for observations on the frontier, $\vec{D_O}(x,\, y,\, u; g_y,\, g_u) = 0$, and for any observation below the frontier, $\vec{D_O}(x,\, y,\, u; g_y,\, g_u) > 0$. Individual performance deteriorates with distance to the frontier, so that the directional output distance value can be interpreted as a measure of inefficiency for each observation.

The directional output distance function can be used to account for the undesirable nature of some outputs of a production process, in this case Nitrogen loading, by specifying a negative direction for those outputs (Chung et al., 1997). This enables the simultaneous expansion of desirable output and contraction of undesirable output in the measurement of performance. The properties of the directional output distance function follow from the assumptions made to characterize $P(x)$, and include Representation, Monotonicity and Translation. Chambers et al. (1996) prove these properties for the input oriented case and we outline their use for estimation purposes in

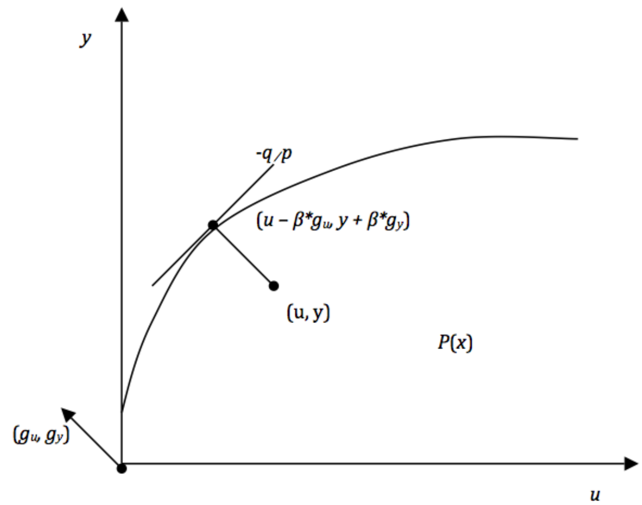

Figure 3: The Directional Output Distance Function for Desirable Output, y, and Undesirable Output, u

the next section.

We use this model to construct the feasible output set for crop production and Nitrogen loading, which allows us to measure the physical tradeoffs for individual producers in the watershed. Given the market value of grass seed, it is also possible to value these tradeoffs in monetary terms (Färe et al., 2001; 2005; 2006) by exploiting the duality that exists between the directional output distance function and the revenue function,

$$R(x,\, p,\, q) = \max_{y,\, u} \left\{ py - pu : (y,\, u) \in P(x) \right\}, \tag{7}$$

where $p = (p_1, \, ..., \, p_M) \in \Re_+^M$ is the vector of output prices corresponding to y and $q = (q_1, \, ..., \, q_J) \in \Re_+^J$ is the vector of output prices corresponding to u. By definition,

$$R(x,\, p,\, q) \geq py - qu, \forall (y,\, u) \in P(x), \tag{8}$$

and this, along with the definition of the directional output distance function from (6) and the representation property imply

$$\begin{aligned} R(x,\, p,\, q) &\geq (p,\, q)(y + \vec{D_O}(x,\, y,\, u; g_y,\, g_u)g_y, u - \vec{D_O}(x,\, y,\, u; g_y,\, g_u)g_u) \\ &\geq (py - qu) + \vec{D_O}(x,\, y,\, u; g_y,\, g_u)pg_y + \vec{D_O}(x,\, y,\, u; g_y,\, g_u)qg_u. \end{aligned} \tag{9}$$

Rearranging terms in (9),

$$\vec{D_O}(x,\, y,\, u; g_y,\, g_u) \leq \frac{R(x,\, p,\, q) - (py - qu)}{(pg_y + qg_u)}. \tag{10}$$

The directional output distance function can then be recovered from the right hand side in (10) as the solution to

$$\vec{D_O}(x,\, y,\, u; g_y,\, g_u) = \underset{p,q}{min} \frac{R(x,\, p,\, q) - (py - qu)}{(pg_y + qg_u)}. \tag{11}$$

The vector of shadow prices is derived by applying the envelope theorem to (11), so that

$$\nabla_u \vec{D_O}(x,\, y,\, u; g_y,\, g_u) = \frac{q}{(pg_y + qg_u)} \geq 0, \tag{12}$$

and

$$\nabla_y \vec{D_O}(x,\, y,\, u; g_y,\, g_y) = \frac{-p}{(pg_y + qg_u)} \leq 0. \tag{13}$$

For a single observation, the shadow price ratio is

$$-\frac{q_j}{p_m} = \frac{\partial \vec{D_O}(x,\, y,\, u; g_y,\, g_u)/\partial u_j}{\partial \vec{D_O}(x,\, y,\, u; g_y,\, g_u)/\partial y_m}, \forall m \in M \,\text{and}\, \forall j \in J. \tag{14}$$

The shadow price ratio values the tradeoff in relative terms between the desirable and undesirable output. If at least one of the outputs y_m in $P(x)$ is marketed, in this case crop production, the shadow price of the nonmarketed undesirable output, in this case Nitrogen loading, can be recovered in absolute terms as

$$q_j = -p_m \frac{\partial \vec{D_O}(x,\, y,\, u; g_y,\, g_u)/\partial u_j}{\partial \vec{D_O}(x,\, y,\, u; g_y,\, g_u)/\partial y_m}, \forall m \in M \,\text{and}\, \forall j \in J. \tag{15}$$

We note that in this application, the desirable output, crop production, is measured in terms of total sales, so that a unit of output is \$1.00. This normalizes the price of output, p_m, to equal \$1.00 as well.

3.2 Estimating the tradeoffs in practice

To compute the marginal effects and shadow prices of each output in practice requires parameterization of the output frontier. In choosing a functional form for that parameterization, we are guided by the properties of the directional output distance function. Only two forms are known to satisfy the translation property, and of these, only the quadratic form contains the first order parameters necessary to compute marginal effects (Färe and Lundberg, 2006). More recently, Färe et al. (2010) use Monte Carlo simulations to demonstrate the ability in practice of the quadratic directional output distance function to characterize the output set. The quadratic (also as in Aigner and Chu, 1968) directional output distance function (Färe et al., 2001; 2005; 2006) is estimated as

$$
\begin{aligned}
\vec{D}_O(x, y, u; g_y, g_u) =& \alpha_0 + \sum_{n=1}^{N}\alpha_n x_n + \sum_{m=1}^{M}\beta_m y_m + \sum_{j=1}^{J}\gamma_j u_j \\
&+ \frac{1}{2}\sum_{n=1}^{N}\sum_{n'=1}^{N}\alpha_{nn'}x_n x_{n'} + \frac{1}{2}\sum_{m=1}^{M}\sum_{m'=1}^{M}\beta_{mm'}y_m y_{m'} \\
&+ \frac{1}{2}\sum_{j=1}^{J}\sum_{j'=1}^{J}\gamma_{jj'}u_j u_{j'} + \sum_{n=1}^{N}\sum_{m=1}^{M}\delta_{nm}x_n y_m \\
&+ \sum_{n=1}^{N}\sum_{j=1}^{J}\nu_{nj}x_n u_j + \sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{mj}y_m u_j.
\end{aligned}
\tag{16}
$$

We estimate the quadratic directional output distance function as a constrained linear programming problem, choosing the parameters to minimize each observation's distance to the frontier. The solution to this problem, the optimal parameter values and $\vec{D}_O^{\,k}$ minimize

$$
\sum_{k=1}^{K}\vec{D}_O^{\,k}(x^k, y^k, u^k; g_y, g_u) \tag{17}
$$

subject to

i. Representation

$$
\vec{D}_O^{\,k}(x^k, y^k, u^k; g_y, g_u) \geq 0, k = 1, ..., K,
$$

ii. Monotonicity

$$\frac{\partial \vec{D_O}^k(x^k,\, y^k,\, u^k; g_y,\, g_u)}{\partial y_m^k} \leq 0, m = 1, ..., M, k = 1, ..., K,$$

$$\frac{\partial \vec{D_O}^k(x^k,\, y^k,\, u^k; g_y,\, g_u)}{\partial u_j^k} \geq 0, j = 1, ..., J, k = 1, ..., K,$$

$$\frac{\partial \vec{D_O}^k(x^k,\, y^k,\, u^k; g_y,\, g_u)}{\partial x_n^k} \geq 0, n = 1, ..., N, k = 1, ..., K,$$

iii. Translation

$$\vec{D_O}(x,\, y + \alpha g_y; g_y) = \vec{D_O}(x,\, y; g_y) - \alpha$$

The constraints ensure that the quadratic form satisfies the properties of the directional output distance function. The first constraint satisfies the representation property by requiring all observations to either lie on or below the output frontier. The second constraint states that an increase in any desirable output, a decrease in any undesirable output, or a decrease in any input can only reduce an observation's distance to the output frontier, which guarantees monotonicity for both inputs and outputs. The third constraint imposes the translation property, restricting the parameters so that additional output in the g_{y_m} direction reduces an observation's distance to the frontier by an equal amount. A reduction of undesirable output in the g_{u_j} direction decreases an observation's distance to the frontier by an equal amount. The final constraint adds the symmetry condition for cross-input and cross-output effects.

4 Empirical Application

We use the HGA to maximize profit and minimize Nitrogen loading for a set of 87 real grass seed farms in the Calapooia river watershed, a tributary of the Willamette river basin west of the Cascades Mountain range in Oregon. These farms are situated in the lower portion of the watershed, which has a drainage area of 682 km^2. The environmental effects of agricultural land use in the Lower Calapooia have been previously studied as part of the USDA Conservation Effects Assessment Project (CEAP) (Confessor and Whittaker, 2007; Mueller-Warrant et al., 2012). The vast majority of the watershed is used for agricultural crop production (83 %) with most of this in grass seed farming. This is followed by hay/ pasture/ range areas (12 %). Wetlands, water bodies and urban areas comprise the remaining area. A recent National Water Quality Assessment of the watershed identifies nitrate Nitrogen as a particular concern, due to the increasing trend of stream and groundwater concentrations in excess of human health and aquatic life standards (Mueller-Warrant et al., 2012; Dubrovsky et al., 2010). Recent sampling confirms that these Nitrogen concentrations vary greatly across the basin, even for areas with over 90% of land in agriculture (Mueller-Warrant et al., 2012), making this a particularly interesting case to consider for policy targeting.

4.1 The SWAT application

We use the SWAT model to divide the study area into 381 subbasins and 533 HRUs. We calibrated the SWAT model with daily streamflow data at the basin outlet Albany, OR, obtained from the U.S. Geological Survey

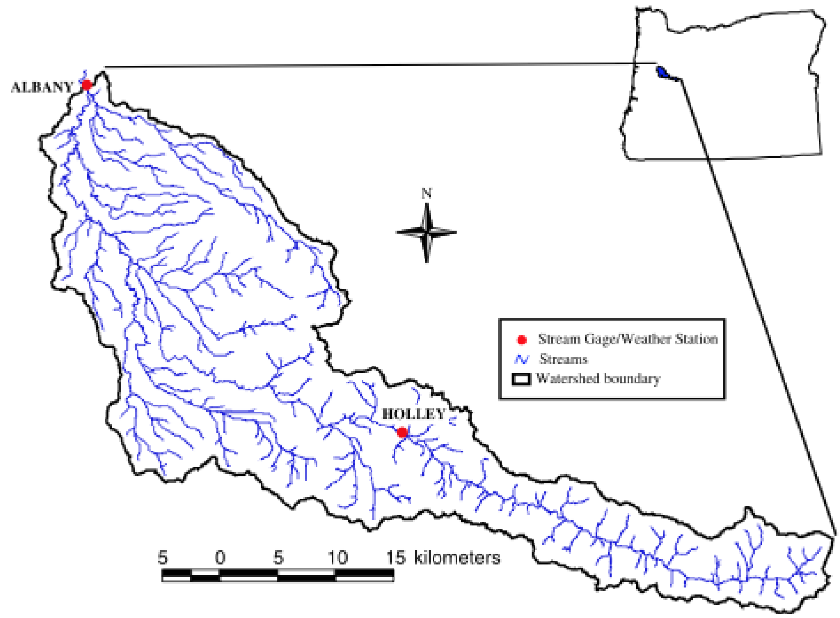

Figure source: Confessor and Whittaker (2007)

Figure 4: The SWAT delineation of the Lower Calapooia watershed

(USGS) National Water Information System (NWIS) website. We used soil data from the SSURGO state soil geographic database for Oregon, obtained from the U.S. Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS), land use data from the USGS National Water Quality Assessment (NAWQA) program, and climate data from the Oregon Climatic Service (OCS). We calibrated the model using the automatic calibration method described in Confessor and Whittaker (2007) and Whittaker et al. (2010). Figure 4 depicts the Calapooia watershed stream system.

4.2 The Pareto optimal tax policy

Due to USDA confidentiality restrictions, agricultural policy studies commonly model the decisions of a representative farm, and are applied to aggregated production data. The USDA National Agricultural Statistics Service (NASS) granted us access to detailed farm-level records from the 2002 Census of Agriculture with the confidentiality restriction that the data could only be accessed from NASS computers.

The HGA requires parallel computation, and could not be run using available NASS computing capability. To maintain the confidentiality of individual producers, we constructed a synthetic data set from the original records for application of the economic model. Fully synthetic data sets are constructed by multiple imputation (Rubin, 1993) of all observations for all variables in the data set, and are generally considered protection against disclosure of confidential data. Bayesian networks provide a useful method for imputation and creation of synthetic data sets, particularly in high dimensions (Thibaudeau and Winkler, 2002; Di Zio et al., 2004).

The estimated Bayesian network satisfies the confidentiality restrictions and can be copied to non-secured computers. We construct the synthetic microdata for use in the DEA profit maximization model using constrained draws from the Bayesian network. Our constructed synthetic microdata has the same statistical properties as the original census records and protects the confidentiality of the individual producers. The synthetic data were also shown to generate the same results for the DEA profit maximization model, which can be run in isolation using NASS computers, as the original

Table 1: Descriptive Statistics for the Calapooia Synthetic Microdata

87 Obs.*	Mean	Min	Max	Std. Dev
Crop Sales ($)	731,800.63	7,744.39	3,404,889.01	591,995.20
Acres	1,715.48	27.54	6,972.44	1,370.35
Labor	112,772.43	241.37	484,628.39	101,673.74
Fertilizer	92,911.99	6,524.38	342,890.26	72,011.17
Seed	16,903.00	4.58	104,308.10	21,278.89
Chemicals	60,243.97	27.21	565,094.90	90,974.07
Fuel	25,720.85	283.84	169,372.22	27,747.60
Utilities	13,392.67	0.00	82,088.85	14,950.13
Maintenance	43,410.06	21.20	159,912.51	37,177.90
Other Expenses	204,159.40	8,411.60	716,892.97	154,316.76

*Note, all input data with the exception of acreage is listed in expenditure form.

census records. Table 1 provides descriptive statistics for the input and output data listed in expenditure and revenue form, with the exception of acreage. According to NASS records, Nitrogen fertilizer sold for $191/ton in 2002, which implies that farms in our sample applied roughly 486.5 tons of fertilizer on average.

For the targeted tax policy HGA, we set up a population of 200 individuals (the number of cluster nodes). Each individual genome consists of 87 targeted tax rates, one for each farm in the watershed. The tax rate values range from 1 to 10, so that the optimal tax payments could range from 0 to up to 9 times the total fertilizer expenditure for a given farm. The HGA runs and tests the fitness of different individuals for their ability to simultaneously optimize both environmental and economic objectives. After several thousand generations, only the fittest solutions are retained and the resulting solutions approximate the Pareto optimal frontier.

Figure 5 depicts the Pareto optimal frontier for Nitrogen Loading and

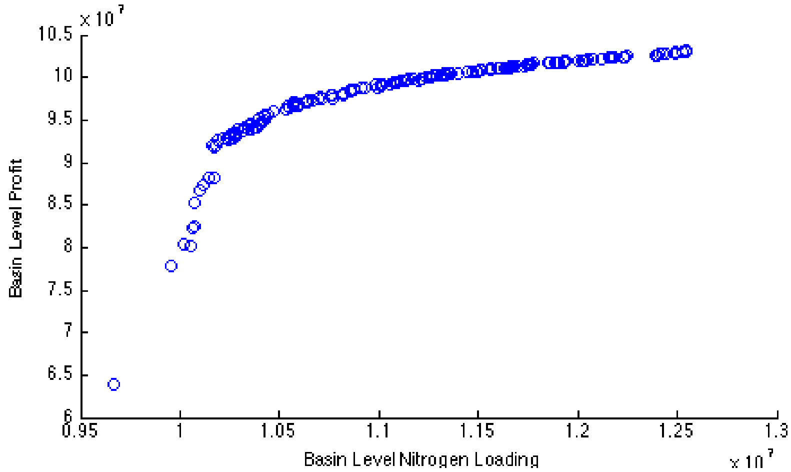

Figure 5: The Pareto optimal frontier for the targeted tax policy

Profit at the basin level, summing over all 87 farms for the 200 individual candidate solutions to the targeted tax policy HGA.

4.3 Tradeoff results

The targeted tax policy HGA generates the Pareto optimal tax rate and corresponding profit-maximizing production decisions and Nitrogen loading for each of the 87 farms in each of the 200 individual cluster nodes in the parallel computation. This yields a data set of 17,400 simulated observations. For this second stage of analysis, where the profit-maximizing input and output quantities have already been chosen, we also combine some of the inputs to reduce the number of parameters that must be estimated.

For computational purposes, we convert each observation's input and output level to a mean-weighted amount. Weighting each input and output by its respective sample mean insures independence of unit of measurement (Shephard, 1970) and corrects for differences in scale.

Table 2: Calapooia Simulated Microdata and Distance Results

17,400 Obs.	Mean	Min	Max	Std. Dev
Acres	1,715.47	27.54	6,972.40	1,362.48
Labor Expenditures	112,772.43	241.37	484,628.39	101,090.62
Other Expenditures	363,829.96	19,981.10	977,014.99	232,010.46
Fertilizer (tons)	253.83	0	8,928.30	374.89
Crop Sales	734,068.28	0.47	2,834,500	573,605.17
Nitrogen Loading (lbs)	128,978.06	0	3,462,393.05	162,294.89
Distance	0.62	0.00	5.20	0.38
Tax rate	2.03	1	10	1.90
q elasticity	0.98	0.00	3.01	0.42
q price	5.58	0.00	17.11	2.39

Thus, the distance value for a hypothetical observation at the mean can be interpreted as the percent increase in desirable output y_m^k and decrease in undesirable output u_j^k required to reach the corresponding point (y_m^{k*}, u_j^{k*}) on the output frontier. The marginal effects of each output can then be interpreted as percent changes in inefficiency, so that the shadow price ratio provides a measure of the elasticity of the tradeoff between crop sales and Nitrogen loading for each producer. The simulated microdata, Pareto optimal tax rates, directional output distance function results and Nitrogen loading shadow price ratios are summarized in Table 2.

The average Pareto optimal tax rate from the HGA is 2.03, or 2.03 times the price of fertilizer. The market price of fertilizer in this study is $191 per ton, making the average Pareto optimal fertilizer cost equal to roughly $380 per ton. Profit-maximizing fertilizer application decreases substantially, falling from an average of 486.6 tons per farm to an average of 253.8 tons per farm under the tax policy. While crop sales decrease for more than half of the farms in our sample under the tax policy, average crop sales increase slightly, from roughly \$732,000 to \$734,000 per farm. This is due to a shift in optimal production intensities under the tax policy and our estimate of the production technology for the basin.

The distance value of 0.62 suggests that on average, producers in the basin could increase their crop sales and decrease their Nitrogen loading by 62 percent from mean levels, based on the production levels of other farms in the basin. For a hypothetical observation at the mean, this corresponds to a feasible reduction of roughly 80,000 lbs. of Nitrogen loading and an increase of roughly \$450,000 in crop sales. We caution that differences in location within the basin stream system, as well as unobserved differences in soil quality may be driving these relatively high estimated inefficiencies.

Along the frontier, the tradeoff between crop sales and Nitrogen loading, measured in elasticity form, is close to one on average. This implies that on average, a one percent reduction in Nitrogen loading (from mean levels) corresponds to a one percent reduction in crop sales (from mean levels). To convert this value to monetary terms,

$$q = -p \frac{\partial \vec{D_O}(x,\, y,\, u); g_y,\, g_u)/\partial u}{\partial \vec{D_O}(x,\, y,\, u); g_y,\, g_u)/\partial y} \frac{\bar{y}}{\bar{u}}. \tag{18}$$

The desirable output, grass seed sales, is measured in dollars, so that the price for an additional dollar of grass seed sales, p, is normalized to equal \$1.00. Thus, the average estimate for the shadow price of Nitrogen loading, q, in monetary terms is \$ 5.58 per lb., and q ranges from 0.00 to \$17.11 per lb. across individual producers. These values should be interpreted with caution, particularly given that they are derived from simulated outcomes.

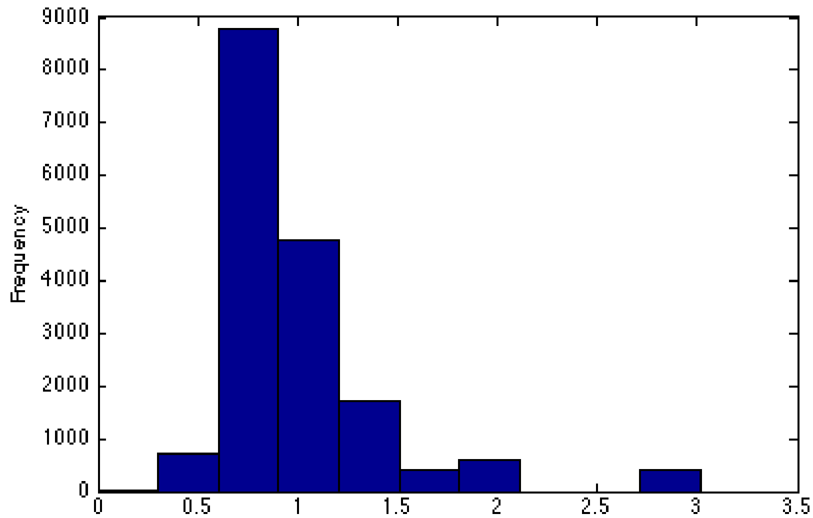

Figure 6: Distribution of Nitrogen Loading Shadow Price Elasticities

They do however shed some light on the possible range of values for Nitrogen loading in the basin, as well as how these values vary across the farms.

Figure 6 illustrates the distribution of estimated tradeoff elasticities. Elasticity values lie between 0.5 and 1.5 for the majority of observations in our sample. For relatively inelastic observations, a one percent reduction in Nitrogen loading corresponds to more than a one percent reduction in crop sales. The opportunity cost of reductions to Nitrogen loading is greatest for these farms under the tax policy. Several factors could explain a more inelastic tradeoff. These farms may be situated on more productive land in the basin, on land where applied fertilizer is less apt to run off due to gradient conditions, or they may also be located at a point in the stream network where runoff has less of an effect on basin-level Nitrogen loading.

All of these are important to consider from a policy perspective. For instance, it may be desirable to target incentives for additional best management practices, such as buffer strips, to these farms given that their reductions in Nitrogen loading that rely on decreased fertilizer usage alone are so costly. At the other end of the distribution, for farms where the tradeoff is relatively elastic, a one percent reduction in Nitrogen loading corresponds to much less than a one percent reduction in sales. These farms may be situated on less productive land, on land where applied fertilizer is more likely to run off, or at a point in the stream network where runoff has more of an effect on Nitrogen loading in the basin. From a policy perspective, it may be desirable to target additional incentives for land retirement to these farms.

5 Conclusion

The tradeoff between agricultural production and water quality is widely acknowledged. The effectiveness of any policy incentive to address this problem depends not only on how farmers respond, but also on the physical relationship between their production activities and the surrounding watershed. Recent computational advances allow for simultaneous consideration of both questions, enabling more complete policy analysis. In this study, we take just such an integrated approach by developing a hybrid genetic algorithm to solve for an optimal tax policy that jointly maximizes agricultural profit and minimizes basin-level Nitrogen loading. Our framework advances the integrated economic and biophysical literature by incorporating realistic models of both farm production and the basin hydrology, by more freely optimizing over both objectives, and by fully endogenizing economic cost without imposing an *a priori* production technology.

We then use our framework to better understand the tradeoffs that result at the farm level under the prospective tax policy. Working with a set of grass seed farms from Oregon's Calapooia River watershed, we estimate an average shadow price of \$5.58/lb. for Nitrogen loading, providing information on the cost to farmers of decreasing current loadings in the basin. We also find that this tradeoff varies across farms, from relatively elastic for some to relatively inelastic for others. The distribution of tradeoff values likely depends on several factors, including differences in soil productivity, topography, and location in the basin's hydrological network. This suggests the need for more adaptive management policies in conjunction with the fertilizer tax, such as incentives for the use of best management practices on more productive working land and taking some marginal, or critically-located land out of production altogether. The distribution of tradeoff values would also likely affect the feasibility of implementing these policies in practice. For instance, a policy that concentrates Nitrogen reduction costs among producers in one are of the basin may be less feasible than one that would spread costs more evenly across the watershed. Individual tradeoff values could be used to assess the distributional implications of prospective agri-environmental policies.

It is important to also note the limitations of this study. Perhaps most importantly, we focus on a single fertilizer reduction policy. A more realistic analysis would consider a range of policies to address Nitrogen loading, including best management practices and land retirement. Allowing for more policy options would likely lower the overall cost of Nitrogen reduction. Our framework does not preclude multiple policies. One can add multiple policy incentives to the profit maximization problem and use SWAT to model their physical effects. Here we focus on the overall framework to endogenize policy response, and leave the question of multiple policies for a separate application. We also do not attempt to estimate the causal determinants of tradeoff differences across farms. Likely determinants include on-farm practices, topographical characteristics and location in the basin system. A better understanding of how these factors affect tradeoff differences would be useful for targeting policies in practice.

While our application focuses on a small agricultural watershed in the Pacific Northwest, this framework could be adapted to analyze environmental tradeoffs for larger and more policy-relevant watersheds, in both the U.S. and internationally. It is also possible to expand the analysis to include additional environmental objectives, such as biodiversity measures or water flow by using the HGA approach. This framework could also be adapted to model changing environmental tradeoffs over time, in response to a variety of factors, including efficiency and technology change, prospective agri-environmental policies, changing development patterns, commodity price changes, or projected climate change.

Acknowledgments

We wish to thank Sarah Stafford, Amy Henderson, and participants in the 2014 Asia Pacific Productivity Conference for providing useful comments. All data used, including the HGA results used to calculate the individual

tradeoffs, are available from the authors upon request.

References

Aigner, D.J., and S.F. Chu, 1968. On estimating the industry production function. American Economic Review 58, 826-838.

Arabi, M., Rao, S.G., Mohamed, M.H., 2006. Cost-effective allocation of watershed management practices using a genetic algorithm. Water Resources Research 42, W104529.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment Part 1: Model development. Journal of the American Water Resources Association 34(1), 73-89.

Arnold, J.G., Moriasi, D. N., Gassman, P. W., et al., 2012. SWAT: Model Use, Calibration, and Validation. Transactions of the ASABE 55(4), 1491-1508.

Ball, V. E., Lovell, C. A. K., Luu, H., Nehring, R., 2004. Incorporating environmental impacts in the measurement of agricultural productivity growth. Journal of Agricultural and Resource Economics, 29(3) 436-460.

Bard, J., 1998. Practical Bilevel Optimization: Algorithms and Applications. Kluwer, The Netherlands.

Bekele, E.G., Nicklow, J.W., 2005. Multiobjective management of ecosystem services by integrative watershed modeling and evolutionary algorithms. Water Resources Research 41, W10406.

Chambers, R. G., Chung, Y., Färe, R., 1996. Benefit And Distance Functions. Journal of Economic Theory 70, 407-419.

Charnes, A., Cooper, W. W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2, 429-444.

Chung, Y.H., Färe, R., Grosskopf, S., 1997. Productivity and undesirable outputs: a directional distance function approach. Journal of Environmental Management 51, 229-240.

Confessor, R.B., Whittaker, G., 2007. Automatic Calibration of Hydrologic Models With Multiobjective Evolutionary Algorithm and Pareto Optimization. Journal of the American Water Resources Association 43, 981 - 989.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6, 182.

Di Zio, M., Scanu, M., Coppola, L., Luzi, O., Ponti, A., 2004. Bayesian networks for imputation. Journal of the Royal Statistical Society. Series A: Statistics in Society 167, 309.

Dubrovsky, N.M., K.R. Burow, G.M. Clark, J.M. Gronberg, P.A. Hamilton, K.J. Hitt, D.K. Mueller, M.D. Munn, B.T. Nolan, L.J. Puckett, M.G. Rupert, T.M. Short, N.E. Spahr, L.A. Sprague, and W.G. Wilber, 2010. The Quality of our Nation's Waters–Nutrients in the Nation's Streams and Groundwater, 1992-2004. US Geological Survey Circular 1350. Reston, BA: US Geological Survey.

Färe, R., Grosskopf, S., Weber, W., 2001. Shadow prices of Missouri public conservation land. Public Finance Review 29(6), 444-460.

Färe, R., Grosskopf, S., Weber, W., 2006. Shadow prices and pollution costs in U.S. agriculture. Ecological Economics 56, 89-103.

Färe, R., Grosskopf, S., Noh, D., Weber, W., 2005. Characteristics of a polluting technology: theory and practice. Journal of Econometrics 126, 469-492.

Färe, R., Lundberg, A., 2006. Parameterizing the shortage function. mimeo. Department of Economics, Oregon State University.

Färe, R., Martins-Filho, C., Vardanyan, M., 2010. On functional form representation of multi-output production technologies. Journal of Productivity Analysis 33, 81-96.

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman, Inc., New York.

Hillyer, C., Bolte, J., Van Evert, F., Lamaker, A., 2003. The ModCom modular simulation system. European Journal of Agronomy 18, 333-343.

Jha, M., Schilling, K.E., Gassman, P.W., Wolter, C.F., 2010. Targeting land use change for nitrate-Nitrogen load reductions in an agricultural watershed. Journal of Soil and Water Conservation, in press.

Liu, J., Peng, C., Dang, Q., Apps, M., Jiang, H., 2002. A component object model strategy for reusing ecosystem models. Computers and Electronics in Agriculture 35, 17-33.

Moore, R.V., Tindall, C.I., 2005. An overview of the open modelling interface and environment (the open MI). Environmental Science and Policy 8, 279-286.

Mueller-Warrant, G.W., Griffith, S.M., Whittaker, G.W., et al., 2012. Impact of land use patterns and agricultural practices on water quality in the Calapooia River Basin of Western Oregon. Journal of Soil and Water Conservation 67(3), 183-201.

Rabotyagov, S., Jha, M., Campbell, T., 2010a. Searching for Efficiency: Least cost nonpoint source pollution control with multiple pollutants, practices, and targets. Journal of Natural and Environmental Sciences 1, 75-90.

Rabotyagov, S., Jha, M., Campbell, T., 2010b. Nonpoint-source pollution reduction for an Iowa watershed: An application of evolutionary algorithms. Canadian Journal of Agricultural Economics 58, 411-431.

Rabotyagov, S., Jha, M., Campbell, T., 2010c. Impact of crop rotations on optimal selection of conservation practices for water quality protection. Journal of Soil and Water Conservation 65, 369-380.

Randhir, T.O., Shriver, D.M., 2009. Multiattribute optimization of restoration options: Designing incentives for watershed management. Water Resources Research 45, W03405.

Richardson, C.W., Bucks, D.A., Sadler, E.J., 2008. The Conservation Effects Assessment Project benchmark watersheds: Synthesis of preliminary findings. Journal of Soil and Water Conservation 63(6), 590-604.

Rosenthal. R.A., 2008. GAMS: A User's Guide. GAMS Development Corporation, Washington, DC, USA.

Schönert, M., Schauppenlehner, T., Schmid, E., Muhar, A., 2011. Integration of bio-physical and economics models to analyze management intensity and landscape structure effects at farm and landscape level. Agricultural Systems 104, 122 - 134.

Secchi, S., Babcock, B.A., 2007. Impact of high crop prices on environmental quality: A case of Iowa and the Conservation Reserve Program. Working Paper, Iowa State University Center for Agricultural and Rural Development.

Shephard, R., 1970. Theory of Cost and Production Functions. Princeton University Press, Princeton.

Sinha, A., Goldberg, D.E., 2003. A survey of hybrid genetic and evolutionary algorithms. Illinois Genetic Algorithms Laboratory (ILLiGAL).

Sinha, A., Malo, P., Frantsev, A., Deb, K., 2013. Multi-objective Stackelberg game between a regulating authority and a mining company: A case study in environmental economics. Evolutionary Computation (CEC) IEEE Congress, 478-485.

Thibaudeau, Y., Winkler, W.E., 2002. Bayesian networks representations, generalized imputation, and synthetic micro-data satisfying analytic constraints. Technical Report RRS2002/9.

Uthes, S., Matzdorf, B., Müller, K., Kaechele, H., 2010. Spatial targeting of agri-environmental measures: Cost-effectiveness and distributional consequences. Environmental Management 46, 494-509.

Volk, M., Hirschfeld, J., Dehnhardt, A., Schmidt, G., Bohn, C., Liersch, S., Gassman, P.W., 2008. Integrated ecological-economic modelling of water pollution abatement management options in the Upper Ems

River Basin. Ecological Economics 66, 66-76.

Whittaker, G., 2005. Application of SWAT in the evaluation of salmon habitat remediation policy. Hydrological Processes 19, 839-848.

Whittaker, G., Confesor Jr., R., Griffith, S.M., Färe, R., Grosskopf, S., Steiner, J.J., Mueller-Warrant, G.W., Banowetz, G.M., 2009. A hybrid genetic algorithm for multiobjective problems with activity analysis-based local search. European Journal of Operational Research 193, 195 - 203.

Whittaker, G., Confessor, R., Luzio, M.D., Arnold, J.G., 2010. Detection of Overparameterization and Overfitting in an Automatic Calibration of Swat. Transactions of the ASABE 53, 1487 - 1499.

Whittaker, G., Barnhart, B., Färe, R., Grosskopf, S., Bostian, M., Griffith, S.M., Mueller-Warrent, G. W., 2014. Spatial targeting of agri-environmental policy: inter-model communication during optimization. Working paper, available at SSRN: $http://ssrn.com/abstract = 2497134$.