

Economic-environmental trade-offs and the conservativeness of the upper partial moment

Nicolette Matthews¹  · Bennie Grové¹ 

Published online: 17 December 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Accounting for the stochastic nature of environmental outcomes when quantifying economic environmental trade-offs with mathematical programming models requires the use of probabilistic programming approaches like the upper partial moment (UPM) method. Application of the UPM model may result in overregulation and losses in farm profit because the probabilistic constraint is satisfied at a higher level than the specified compliance probability, resulting in conservative responses from polluters. The main objective of this article was to present the upper frequency method as an alternative to enforce a probabilistic constraint with a close bound to the actual compliance probability. The UFM uses binary variables in a linear programming framework to enforce the probability bound on an empirically distributed outcome variable. Results showed that the UPM model was very conservative in the estimation of the upper probability bound, which resulted in an overestimation of abatement costs and an underestimation of the average amount of pollution above the environmental goal. Inconsistencies also exist between the ranking of alternatives when comparing the UPM and UFM methods. The UFM is general enough to ensure that the technique can be applied to any problem where the researcher is concerned with the risk of exceeding a specified target level.

Keywords Conservativeness · Economic-environmental trade-offs · Environmental risk · Nitrogen losses · Safety-first · Upper partial moment · Upper frequency method

JEL Classification Q52 · Q15 · Q56 · C61

1 Introduction

Weersink et al. (2002) argue that optimal resource allocation is important not only because of its effects on farm income but also because of its environmental impact. Non-point source (NPS) pollution stemming from agricultural practices is seen as a major cause of the remaining water-quality problems in developed and developing countries (Shortle et al. 1998; Rossouw and Görgens 2005; Ranga Prabodanie et al. 2010; Li et al. 2014a, b). Consequently, there is increased pressure on agriculture to use resources optimally in order to reduce the negative environmental effect caused by agricultural practices (Shortle et al. 2001). In the absence of a market for reduced environmental emissions, the information generated with trade-off analysis will be critical for informed policy decision making, as it allows policy makers and the public to assess whether a given improvement in environmental quality is worth the sacrifice in agricultural production (Stoorvogel et al. 2004).

Generating economic-environmental trade-off curves is a complicated endeavor and requires quantifying the inter-relationships between sustainability indicators implied by the underlying biophysical processes and producers' economic behavior (Ranga Prabodanie et al. 2010). Alternative abatement strategies and/or policy instruments are compared on the basis of the alternative that achieves an environmental goal with the least impact on the economic indicator. A complicating factor is that environmental

✉ Nicolette Matthews
MatthewsN@ufs.ac.za

Bennie Grové
groveb@ufs.ac.za

¹ University of the Free State, Bloemfontein, South Africa

emissions are inherently stochastic as a result of a variety of environmental conditions (Horan 2001; Kampas and White 2004; Kataria et al. 2010). Consequently, pollution-control strategies should be aimed at improving the distribution of outcomes rather than some scalar value (McSweeney and Shortle 1990). By implication, these control strategies will achieve environmental goals with only a measure of certainty.

A modeling alternative to incorporate the variability of environmental outcomes while quantifying economic-environmental trade-offs is chance-constrained programming (CCP) (Li et al. 2014b; Kataria et al. 2010; Kampas and White 2003). The application of CCP requires the specification of a functional form for the distribution of the environmental variable (Qiu et al. 2001). Various researchers have shown that the distributional assumptions employed in CCP models have a significant impact on the estimated trade-offs (Zhu et al. 1994; Qiu et al. 2001; Kampas and White 2003; Kataria et al. 2010) and may not hold for all situations as a result of the site-specific nature of agricultural NPS pollution (Wang et al. 2016; Qiu et al. 2001). To overcome the problem, techniques like the Environmental Target-MOTAD model (Teague et al. 1995) were developed to estimate economic-environmental trade-offs while making use of empirical distributions. Qiu et al. (1998) scrutinized the use of the Environmental Target-MOTAD model and argued that it would be difficult to apply because the scientific basis for the selection of a reasonable environmental risk level is weak. As an alternative, these researchers developed the upper partial moment (UPM) stochastic inequality that provides a stronger scientific basis for modeling economic-environmental trade-offs because the environmental risk level is given by the compliance probability.

A potential problem with the application of the UPM model (Qiu et al. 2001) in enforcing a probabilistic constraint is the fact that the actual compliance probability is larger than the specified compliance probability. Even though specified compliance levels may be equal across alternatives, the actual compliance and the optimal management responses may differ significantly between alternatives. These differences raise questions about the fairness with which alternatives are compared. Some researchers (Atwood et al. 1988; Qiu et al. 2001) have raised concerns about the conservativeness¹ of the UPM, although neither of these researchers has investigated the severity of the conservativeness.

¹ Conservativeness relates to a difference between the specified and actual compliance probabilities. Larger differences give rise to higher levels of conservativeness. Reducing the level of conservativeness implies tighter upper probability bounds.

The main objective of the article was to present an alternative method to enforce a probabilistic constraint with a probability bound close to the actual compliance probability, which will result in a less biased comparison between alternatives. The method is applied to demonstrate that the UPM model is very conservative in the estimation of the upper probability bound, which results in an over-estimation of abatement costs and an underestimation of the average amount of pollution above the environmental goal.

The newly developed upper frequency method (UFM) counts the number of states with deviations above the environmental goal in an effort to ensure that the deviations above the goal do not exceed the number of deviations allowed by the model. Like the UPM, the UFM uses an empirical distribution of the environmental outcome to enforce the probabilistic constraint, which overcomes the need to specify the statistical distribution of the outcome variable. The generality of the method makes it applicable to any situation where the risk of exceeding a specified target level is of concern.

2 Conservativeness of the upper partial moment

Safety-first rules are concerned with the probability of a variable falling above or below a critical or target level. Probabilistic safety-first constraints can be imposed using different chance-constraint bounds such as the distribution-free Chebyshev stochastic inequality. Imposing the probabilistic constraints through the use of Chebyshev's inequality generates strongly conservative probability bounds (Atwood et al. 1988). Realizing the need for a tighter probability bound Berck and Hihn (1982) introduced a semi-variance inequality that is able to generate a tighter upper probability bound compared to the Chebyshev. The semi-variance inequality follows Markowitz (1970) in that the mean-semivariance is a more attractive measure of risk than the mean-variance approach of the Chebyshev. Atwood (1985) extended Berck and Hihn's (1982) semi-variance inequality with a more general lower partial moment stochastic inequality to enforce constraints with a smaller upper probability limit than the Chebyshev and the semi-variance inequality. Although the probability bound of the UPM method is tighter than the Chebyshev inequality, the bound is still conservative (Atwood et al. 1988; Qiu et al. 2001).

The probabilistic constraint of achieving a specified environmental goal is defined as follows using the UPM²:

$$Pr[x \geq t + p\theta(t)] \leq \theta(t)/(g - t) \leq (1/p) \quad (1)$$

² See Appendix 1 for the derivation of the inequality.

where x is the pollution variable, t is a reference pollution level, g is the environmental goal, $\theta(t)$ is the UPM measured as absolute deviation above t , and $p = \left(\frac{1}{1-cp}\right)$ and cp are the compliance probability.

Figure 1 is used to explain the application of Eq. 1 and the origins of the overestimation of the actual compliance probability when using the UPM to enforce the probabilistic constraint. The stylized example that was developed portrays a situation where the environmental goal, g , must be maintained at least 75% of the time. The dotted line represents the cumulative probability distribution of x . Enforcing the probabilistic constraint within an optimization framework requires that a reference pollution level, t , be determined during the optimization so that the UPM, $\theta(t)$, expressed as a portion of the difference between g and t , is equal to $1 - cp$. Graphically the difference between g and t is represented by the summation of the areas labeled from 1 to 4, which are equal in size. The shaded area indicating $\theta(t)$ extends beyond g . However, the area of the shaded triangle that goes beyond g is exactly the same size as the area of block 1 that is not shaded. Therefore, $\theta(t)$ is equivalent to the area of block 1. Thus, $\frac{\theta(t)}{g-t}$ is 25%, even though some pollution levels above g are possible. Specifying a value of $p = 4$ will ensure that the proportion is 25% because $t + p\theta(t) = g$. As a result, t will be achieved with the specified cp while g will be achieved with a higher cp , which gives rise to the overestimation of the actual compliance probability when using the UPM inequality to enforce probabilistic constraints.

The only known input parameters to the optimization problem are g , cp , and, therefore, p . The distribution of nitrate losses is conditional on the choice of production practices that will maximize producers' profit margin, given that nitrate losses are no more than g , $1 - cp$ percent of the time. The choice of t and therefore the size of $\theta(t)$ are significantly affected by the endogenously determined distribution of nitrate losses. Thus, there is no chance of

predicting the actual probability that g will be achieved, apart from knowing the bound will be tighter than cp with which t is satisfied.

From an environmental point of view, a tighter probability bound is beneficial. However, from a polluter's point of view, a tighter bound implies overregulation, which may cause considerable loss of profits. The only way to compare alternatives for reducing environmental pollution correctly is to compare alternatives with methods that will generate small differences between specified- and actual cp .

The dashed line represents the distribution of nitrate losses that will achieve g at the given cp . Such an environmental outcome could be achieved by determining states of nature with deviations above g and then restricting the number of states to 25% of the number of total states of nature. Teague et al. (1995) have demonstrated that states with deviations above g could easily be identified using an Environmental Target-MOTAD framework.

Several indicators could be used to determine the conservativeness of the UPM. The most obvious indicator is to compare the specified compliance probability that is used in the UPM to the actual compliance probability as an indicator of the conservativeness of the compliance probability estimate. The UFM allows for at least two new measures to determine the conservativeness of the UPM. Firstly, the difference between the average pollution levels above the environmental goal for the UPM and UFM³ could be compared for obtaining an indication of the environmental impact. Secondly, the cost to the polluter could be estimated by comparing the objective function values of the UPM and UFM to determine the impact on the polluter.

3 Data and procedures

3.1 Data simulation

Crop growth modeling provides a powerful means of generating yield response and environmental indicators for alternative management practices when field measurements are lacking (Weersink et al. 2004; Samarawickrema and Belcher 2005). Quasi-experimental data on yield response and nitrate losses were simulated with a mechanistic, generic crop growth model originally developed for irrigation scheduling (Annandale et al. 1999). The Soil Water Balance (SWB) model was extended by Van der Laan (2009) through the addition of nitrogen and phosphorus simulation routines and algorithms to simulate above-

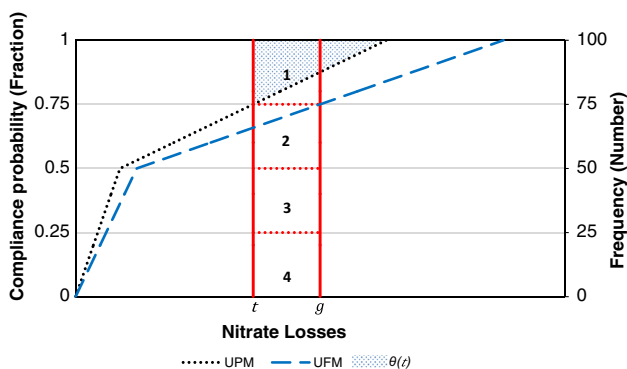


Fig. 1 A stylized graphical illustration of the upper partial moment (UPM) and the upper frequency method (UFM)

³ For the UFM the average pollution levels above the environmental goal are the α -percentile conditional value at risk used in the finance literature.

ground nitrogen mass, grain nitrogen mass, soil water content and the fate of nitrogen. Van der Laan (2009) tested and validated SWB using historical datasets collected in the Netherlands, Kenya and South Africa.

The SWB model was used to simulate crop production and an environmental indicator consisting of nitrate losses (runoff and leaching) for the production of late monoculture maize (planting date 15 December) under irrigation on two soil types at Glen, South Africa. Maize production was simulated for a sandy clay loam (SCL) and sandy clay (SC) soil using 19 years of weather data while assuming an initial soil nitrogen level of 33 kg. Nine levels of fertilizer could be applied in either a single or a split application. When using split applications two-thirds of the desired nitrogen level were applied on the day of planting, while the remaining third was applied seven weeks later. Only applications above 70 kg/ha were applied in a split application.

3.2 Quantifying environmental risk

Unique production conditions during a specific production year cause nitrate loss response to increasing levels of fertilizer application rates to be different between production years. As a result the procedure that is adopted in this research deviates from the norm where a single response function is fitted using all the data points and risk is characterized as deviations from the fitted response function. Instead, our methodology estimates a response function for each production year. Any unexplained variability not captured by the year-specific response function is treated as the risk of not being able to predict nitrate loss as a function of nitrate application rates within a specific year exactly. Using all the year-specific stochastic nitrate loss response functions simultaneously will characterize the risk of not knowing which year will occur, as well as the risk of not being able to exactly predict nitrate loss in the circumstances that the resulting year is known. The benefit of estimating year-specific response functions is that the procedure automatically models the heteroscedasticity of nitrate losses embedded in the data.

Next, the procedure that was used to construct the empirical distribution of the environmental risk indicator is discussed in more detail. According to Richardson et al. (2000), the first step is to determine the non-random (predictable) component using regression analysis. The following equation was estimated for each production year using ordinary least squares (OLS):

$$\hat{E}_s(N_f) = e_{1s} + e_{2s}N_f + e_{3s}N_f^2 + \tau_{sf} \quad (2)$$

where $\hat{E}_s(N_f)$ represents the predicted nitrate losses in production year s as a function of the simulated nitrogen application rates (N_f) (kg/ha), e_{is} is the i th estimated

coefficient for the nitrate loss function in year s , and τ_{sf} is the estimation error for the regression of year s given nitrogen application rate f . In total 19 different regression equations were estimated using the nitrate losses simulated for nine distinct fertilizer application rates ($N_f = 20, 45, 70, 95, 120, 145, 170, 195, 220$). The random component associated with nitrate loss response in each year is represented by the regression residual, which was calculated as:

$$\tau_{sf} = E_{sf} - \hat{E}_s(N_f) \quad (3)$$

where E_{sf} represents simulated nitrate losses in year s for nitrogen application rate f . The empirical outcomes that characterize the variability of nitrate losses for any given level of nitrogen fertilizer application rate are calculated by combining the predictable and random components as follows:

$$\tilde{E}_{sf}(N) = \hat{E}_s(N) + \tau_{sf} \quad (4)$$

where $\tilde{E}_{sf}(N)$ is the empirically distributed nitrate losses as a function of nitrogen application rate. Important to note is that $\tilde{E}_{sf}(N)$ is a continuous function that is not restricted to the nine levels of N used during the simulation process. Equation (4) shows that the empirical distribution of nitrate loss is represented by outcomes for every production year (s) and the error associated with every simulated fertilizer application rate (f). Therefore, 171 ($s \times f$) outcomes characterize the risk of nitrate losses.

The nitrate loss response functions estimated using Eq 2 are presented in Appendix 2. Results for the response functions show that the nitrate losses are unique in every production year. During production year, S12, no relationship could be identified between nitrate losses and fertilizer use. Investigation of the data showed that no nitrate losses were simulated for the production year in question since no losses occurred as a result of a very dry production year. The bulk of the estimations explain a great deal of the variation in the simulated data with a good R^2 . However, not all of the estimations show a high R^2 , indicating that not all of the variation in nitrate losses is due to the amount of nitrogen fertilizer applied. A detailed discussion of the estimated response functions is available in Matthews (2014).

3.3 Gross margin estimation

In our application, modeling economic-environmental trade-offs requires a continuous function that relates average gross margins to any nitrogen application level. The use of continuous response functions overcomes the problem of input diversification. Use of discrete activities (non-continuous) for nitrogen application levels, gross

margin and the nitrate loss levels could results in input diversification by the solution procedure, resulting in results that are near impossible to achieve in practice. The procedure that was used to construct the empirical distribution of nitrate losses was used to construct the variation in gross margins as a function of fertilizer application. The gross margin outcomes were then averaged to yield the economic indicator. Specifically, expected gross margins were estimated using the following equation:

$$\overline{GM}(N) = \sum_{sf} p_{sf} (\tilde{Y}_{sf}(N)P_Y - NP_N - \tilde{W}_{sf}(N)P_W - C_a - C_Y \tilde{Y}_{sf}(N)) \tag{5}$$

where $\overline{GM}_s(N)$ is the expected gross margin as a function of applied nitrogen (ZAR/ha).⁴ $\tilde{Y}_{sf}(N)$ is the empirical distribution of crop yield (ton/ha) as a function of applied nitrogen (N), $\tilde{W}_{sf}(N)$ is the empirical distribution of water applications (mm) as a function of applied nitrogen, N is the amount of nitrogen fertilizer (kg/ha) applied. P_Y is the price of maize (ZAR/ton), P_N is the price for nitrogen fertilizer (ZAR/kg), P_W is the cost of applying irrigation water (ZAR/mm). C_A is the area-dependent cultivation cost (ZAR/ha), C_Y is the yield-dependent harvesting cost (ZAR/ton), and p_{sf} is the probability that outcome sf will occur. p_{sf} is equal to $\frac{1}{s \times f}$.

The empirical distributions of crop yield ($\tilde{Y}_{sf}(N)$) and applied irrigation water ($\tilde{W}_{sf}(N)$) were respectively calculated with Eqs 6 to 8 and Eqs 9 to 11.

$$\hat{Y}_s(N_f) = \beta_{1s} + \beta_{2s}N_f + \beta_{3s}N_f^2 + \varepsilon_{sf} \tag{6}$$

$$\varepsilon_{sf} = Y_{sf} - \hat{Y}_s(N_f) \tag{7}$$

$$\tilde{Y}_{sf}(N) = \hat{Y}_s(N) + \varepsilon_{sf} \tag{8}$$

$$\hat{W}_s(N_f) = \omega_{1s} + \omega_{2s}N_f + \omega_{3s}N_f^2 + \mu_{sf} \tag{9}$$

$$\mu_{sf} = W_{sf} - \hat{W}_s(N_f) \tag{10}$$

$$\tilde{W}_{sf}(N) = \hat{W}_s(N) + \mu_{sf} \tag{11}$$

β_{is} and ω_{is} represent the i th OLS-estimated coefficients respectively for the yield response function and the irrigation water response function in the regression for year s , while ε_{sf} and μ_{sf} represent the estimation errors of the yield response and irrigation water response functions respectively.

Account should be taken of the fact that crop yield was only estimated as a function of nitrogen applications and seemingly no relationship exists between water applications and crop yield. No relationship was modeled because

⁴ The exchange rate as on 30 September 2014: 1 ZAR = 0.08861 USD, where ZAR indicates South African Rands.

Table 1 Crop price and input costs for maize production at Glen, South Africa

Maize price	2400 ZAR/ton
Fertilizer (nitrogen)	15.38 ZAR/kg
Water and cost of irrigation	2.25 ZAR/mm
Mechanization cost to apply fertilizer in a split application	504.88 ZAR/ha
Harvesting cost	65.39 ZAR/ha
Transport cost	69 ZAR/ton
Fixed costs (for seed, plant protection, machinery, irrigation equipment, cost of other nutrients applied, etc.)	8723 ZAR/ha

the auto irrigation strategy that was used to determine the timing and number of water applications during the data-simulation process was set up in such a manner that water was never limiting to crop development. Inspection of the simulated data, however, revealed that water applications were lower when crop yield was reduced because of nitrate deficiencies. SWB reduces the leaf area index when nitrate deficiencies occur and consequently crop transpiration was reduced and resulted in less irrigation water being applied. Thus, crop yield was modeled as a function of nitrogen applications because water never limited crop production while changes in water applications were modeled as a function of nitrogen applications because an underdeveloped crop requires less irrigation water.

Production cost data and input prices for 2014 are from Griekwaland-Wes Cooperation (GWK Ltd), South Africa. Table 1 presents the crop price and the input costs used in this paper.

3.4 Economic-environmental compliance models

Data parameters for average gross margins and empirical distributions of nitrate losses are estimated for 220 different fertilizer application rates,⁵ with the use of the procedures outlined above. The generated data parameters are incorporated into an UPM model and an UFM model to estimate the conservativeness of the UPM. Both compliance models include equations that are generic to both compliance models and equations that are specific to the method used to model compliance. The optimization model was developed in GAMS (GAMS Development Corporation 2007a) and solved using the CPLEX solver (GAMS

⁵ Due to slow convergence of the solution procedure data parameters for gross margin and the empirical distribution of nitrate losses were simulated using the estimated response functions. The use of 220 different fertilizer-application rates ensures a smooth approximation of the response functions.

Development Corporation 2007b). Next, the generic model will be discussed followed by the specific equations necessary to model compliance with the UPM model and the UFM model.

3.4.1 Generic model

The generic model specification includes the objective function as well as constraints to limit intensive and extensive margin responses. The following equations are generic to both compliance models:

$$\text{Maximise } TGM = (\overline{GM}(N)) * HA \quad (12)$$

s.t.

$$N \leq 220 \quad (13)$$

$$HA \leq 1 \quad (14)$$

where TGM is the total gross margin as a function of applied nitrogen (ZAR) and the area cultivated, HA (measured in ha). The area cultivated can be interpreted as the absolute area cultivated or as a fraction of the area available for cultivation.

The decision variables are the fertilizer application rate and the irrigated area that will maximize the total gross margin. Fertilizer applications were limited to a maximum of 220 kg/ha while the area planted was constrained to be no more than one hectare.

3.4.2 Environmental compliance with the upper partial moment (UPM)

The compliance models require additional equations to model compliance with the user-specified environmental goal of 28 kg of nitrate. The generic model was used to determine baseline levels of nitrate losses for production on all soil types and using both fertilizer application methods. The assumption was made that policy makers would want to reduce the probability of an average amount of nitrate loss. Therefore, the nitrate losses for all four alternatives were averaged to determine a homogenous nitrate loss goal of 28 kg.

The equations that are added to the generic model to complete the UPM model are given below:

$$t - (\tilde{E}_{sf}(N))HA - d_{sf} \geq 0 \quad (15)$$

$$\sum_{sf} p_{sf} d_{sf} - \theta(t) = 0 \quad (16)$$

$$t + p^* \theta(t) \leq g \quad (17)$$

t is the endogenously determined reference level for the environmental variable with d_{sf} being the deviation of

pollution emissions above the pollution reference level t for outcome sf and g – the environmental goal set by the environmental regulator. $\theta(t)$, where $\theta(t) = \theta(1, t) = \rho(1, t)$, represents the endogenously determined environmental risk level or the expected deviation above the reference level t . Furthermore, $p^* [p^* = (\frac{1}{1-cp})]$ is the inverse of one minus the compliance probability with respect to g .

As mentioned earlier in this article the probabilistic constraint of the UPM in Eq. 16 is enforced by choosing a reference pollution level, t , so that the UPM, $\theta(t)$, expressed as a portion of the difference between g and t , is equal to the acceptable probability $(1 - cp)$ of the pollution level being greater than the goal. The deviation of pollution emissions (d_{sf}) above the endogenously determined reference pollution level (t) is estimated with Eq. 15. These deviations are multiplied by their occurrence probability to estimate the UPM, $\theta(t)$, as absolute deviations from the reference pollution level.

3.4.3 Environmental compliance with the upper frequency method (UFM)

The UFM of enforcing probabilistic environmental compliance is based on the premise that any compliance probability can be expressed for the discrete case as the frequency with which a goal may be exceeded. Restricting the number of states in which the environmental goal might be exceeded guarantees compliance. The UFM utilizes the Environmental Target-MOTAD model specification to identify states of nature in which the environmental goal is exceeded and uses binary variables to restrict the number of times the goal is exceeded. The following equations were used to ensure compliance:

$$g - (\tilde{E}_{sf}(N))HA - d_{sf} \geq 0 \quad (18)$$

$$-lB_{sf} + d_{sf} \leq 0 \quad (19)$$

$$\sum_{sf} B_{sf} \leq uf \quad (20)$$

where B_{sf} is a binary variable indicating whether the environmental goal is exceeded by outcome sf , while uf is the upper frequency indicating the number of times a goal might be exceeded to enforce compliance, and l is a large number that is used to give permission for outcome sf to exceed the goal, given that B_{sf} has a value of one.

Absolute deviations (d_{sf}) are estimated in Eq. 18 as the deviation in nitrate loss ($\tilde{E}_{sf}(N)$) from the environmental goal (g). Equation 18 is the same as for the UPM (Eq. 15), with the exception that the deviations are calculated from g and not t as in the UPM. The UFM, therefore, overcomes

the conservativeness of the UPM in maintaining the true environmental goal and not an endogenously determined reference pollution level that is dependent on the distribution of the environmental variable. Equation 19 uses a binary variable to identify whether a specific outcome exceeds the environmental goal. Every time $\tilde{E}_{sf}(N)$ exceeds g , B_{sf} takes a value of one. The B_{sf} s are counted to determine the frequency with which the environmental goal is exceeded. The probabilistic constraint is enforced by Eq. 20, which restricts the number of times g is exceeded to uf . The value of uf is calculated as $(1 - cp)sf$, where sf is the total number of outcomes. The choice of uf is an integer value that corresponds with a value closest to the estimated discrete compliance probability without exceeding the compliance probability. Therefore, the UFM can also be conservative in the estimation of the trade-offs if the number of discrete states is small. However, the UFM will never be as conservative as the UPM.

4 Results

4.1 UPM economic-environmental trade-offs

The UPM-generated economic-environmental trade-offs of maintaining a nitrate loss goal of 28 kg at increasing levels of compliance for two soils (SCL and SC) and two fertilizer application methods (Single and Split) are shown in the lower section of Fig. 2.

The UPM trade-off curves show that the gross margins for the SCL soils are consistently higher when compared to SC soil and that a single fertilizer application is preferred to a split application when a specific soil is being considered. Total gross margins are decreasing at an increasing rate with increasing levels of specified compliance probability with

the exception of increases in the cp beyond 90% for the SCL soil. The reduction in total gross margins from the lowest to the highest specified cp is on average 48% for the SCL soil and 63% for the SC soil, with little difference between fertilizer-application methods for a specific soil type.

4.2 Compliance probability conservativeness

The UPM method is said to be conservative with respect to the actual compliance that is achieved with the modeling procedure, while probabilistic constraints are being enforced. The compliance probability conservativeness is evaluated by comparing the specified compliance probability with the actual probability with which the environmental goal is achieved. The actual compliance probability of the UPM model is computed *ex-post* to the optimization, using the optimized distribution of the environmental variable. The comparison between specified and actual compliance is shown in Fig. 3. A 45° line is also shown to indicate perfect correspondence between the specified- and the actual compliance probabilities.

Figure 3 reflects huge discrepancies between specified- and actual compliance probabilities. In all cases, the actual compliance level is much higher than the specified compliance level, especially at low levels of specified compliance. Furthermore, the actual cp achieved on the SC soil is higher when compared to the SCL soil for a specific level of compliance. At the lowest level of cp , the difference is 24.6 percentage points for the SCL soil and about 28.4 percentage points for the SC soil. For increasing levels of specified compliance, there is very little change in the actual compliance probabilities to a point where the actual probabilities increase to the highest level of specified compliance. At the highest level of cp , the differences in compliance probabilities decrease to 2.4 percentage points

Fig. 2 Gross margins (GM measured in ZAR) for the upper partial moment (UPM) and the upper frequency method (UFM) at increased specified compliance probability levels for two soils (SCL and SC) and two fertilizer application methods (single and split)

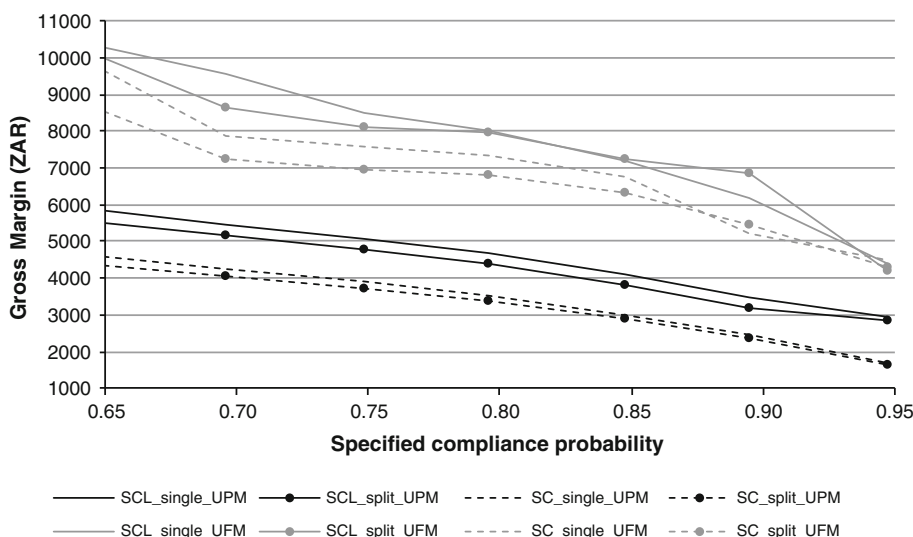
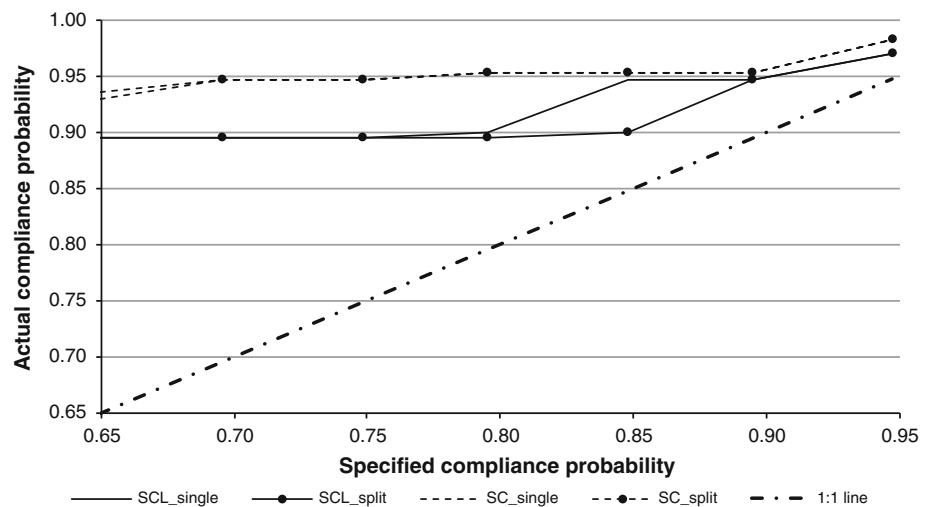


Fig. 3 Actual compliance for increased specified compliance probability levels for the upper partial moment (UPM) for two soils (SCL and SC) and two fertilizer application methods (single and split)



and 3.5 percentage points respectively for the SCL and SC soils.

Users of the UPM may justify the use of the method by arguing that the difference between the specified- and the actual compliance levels becomes very small at high levels of specified compliance and, therefore, the UPM can be used if the specified cp is high. Cognizance should be taken of the method used to enforce compliance using the UPM method. With the UPM model, the intensive and extensive margin responses for achieving the environmental goal are optimized in such a way that the pollution reference level (t) is achieved with the specified cp . The UFM estimates non-compliance directly from the environmental goal (g), which will result in significant changes in the intensive and extensive margin responses and affect the total gross margin and the resulting distribution of nitrate emissions. Evaluating the conservativeness of the UPM in terms of cp alone does not provide any indication of the impact of the conservative estimates of the UPM on the economic indicator or nitrate losses to the environment. The specified compliance of the UPM model was incorporated into the UFM by expressing the cp as the number of observations with which the goal may be exceeded. Consequently, the specified and actual compliance levels are the same for the UFM model results. Thus, comparing the results of the UPM with the UFM allows for better evaluation of the conservativeness of the UPM because the impact on the gross margins of the polluter and the environmental consequences are considered.

4.3 Economic indicator conservativeness

The upper section of Fig. 2 shows the economic-environmental trade-offs generated with the UFM model. The specified compliance of the UPM model was incorporated

into the UFM by expressing the cp as the number of observations with which the goal may be exceeded. Consequently, the specified and actual compliance levels are the same for the UFM model results.

The optimized gross margins of the UFM model are much higher in comparison with those of the UPM model. The difference in the optimized gross margins between the two compliance models measures the impact of the conservativeness of the UPM on the polluters' profitability. From the graph, it is clear that the underestimation of gross margin is not constant across the range of specified compliance probabilities since the trade-off curves of the UFM cross each other, which is not the case with the UPM model. Consequently, choices between different fertilizer application methods on a specific soil type for increasing levels of environmental compliance with the UFM model are not as consistent as with the UPM. However, the SCL is still the preferred soil type. At lower levels of specified compliance, a single fertilizer application is preferred, while split applications are preferred at higher levels of specified compliance. Important to note is that the gross margins tend to converge to a gross margin of R4 342 at the highest level of specified compliance.

Even though the differences in gross margins between the two model specifications are reduced for all strategies with increasing levels of compliance, the differences remain large. The average gross margin differences between the compliance models at the highest cp are R1 372 and R2 737 respectively for an SCL soil type and an SC soil type with respective fertilizer application strategies inducing differences of R78 and R122 respectively for SCL soil and SC soil. On average these differences respectively constitute a 32 and 62% underestimation of gross margins with the UPM model for the SCL and SC soil.

4.4 Environmental conservativeness

The average nitrate losses above the environmental goal are calculated for each model specification and are compared to identify the impact on the environment when using the UFM model with its close bound to the actual *cp*. The model comparisons are shown in Fig. 4.

Figure 4 shows that soil-fertilizer application-method combinations with lower profitability consistently generated the highest average nitrate losses above the goal of 28 kg when considering the UPM model. The magnitude of the losses decreases to almost zero for all the strategies when the specified compliance probability is increased to 94.7%. Of the two soils, the SC soil realized the higher average nitrate losses. Fertilizer-application method does not greatly influence the magnitude of the losses. The results of the UFM model are not as clear cut as for the UPM. However, the observation was made that soil-fertilizer application-method combinations with lower profitability generate the highest average pollution level above the environmental goal. The magnitude of the average nitrate losses also decreases with increasing compliance probability. However, the average losses for the UFM do not converge to almost zero, as is the case with the UPM model. Instead, the average nitrate losses to the environment are about 2.5 kg for the SC soil and respectively 0.98 and 1.59 kg for a single fertilizer application and split fertilizer application on SCL soil. Percentage wise, the average nitrate losses above the environmental goal across all compliance probability levels are respectively 80 and 77% more for the SCL and SC soil types in comparison to the UPM model.

4.5 Changes in the intensive and extensive margin

To ensure compliance with the environmental goal (*g*), both the UPM and UFM models change the intensive and extensive margin. The baseline amount of fertilizer applied (kg/ha) and the area planted (ha), together with the optimal amount for the UPM and UFM, are given in Table 2. The baseline amount of fertilizer applied and the area planted reveal the producers’ production decision for the generic optimization model. The producer is therefore not faced with an environmental constraint and can make production decisions for optimal gross margins without considering his or her environmental impact.

Profit-maximizing nitrogen input levels vary by 7 and 2 kg/ha between fertilizer-application methods on SCL and SC soil types respectively, with no need to comply with an environmental nitrate loss goal. On average the optimal fertilizer application rate on SCL soil is 141 kg/ha, while the application rate on SC soil is 125 kg/ha in the absence of environmental compliance. The UPM results show that the nitrogen fertilizer application rates are reduced more from the optimal level on the SCL soil when compared to the SC soil. However, larger areas are irrigated with the SCL soils irrespective of the fertilizer-application method to comply with the environmental nitrate loss goal. Irrigation areas decrease with increasing environmental compliance probabilities for all soil-crop fertilizer-application method combinations. Interestingly, per hectare fertilizer application rates decrease with increasing compliance probabilities only for the SC soil, as the SCL soil fertilizer application rates are highest at the highest compliance probability.

Fig. 4 Average nitrate losses above the goal (kg) for the upper partial moment (UPM) and the upper frequency method (UFM) at increased specified compliance probability levels for two soils (SCL and SC) and two fertilizer application methods (single and split)

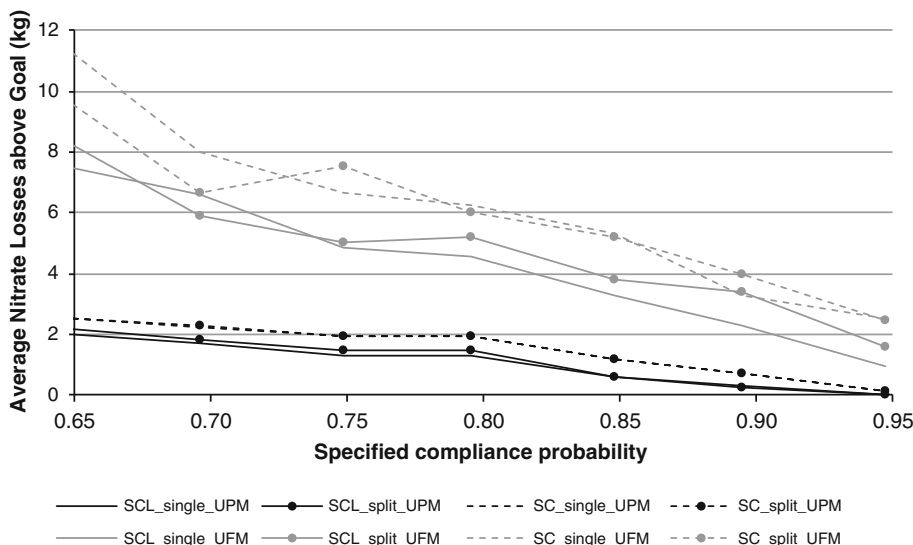


Table 2 Estimated results for extensive (ha) and intensive (kg/ha) margin changes using an upper partial moment (UPM) and upper frequency method (UFM) for a sandy clay loam (SCL) and sandy clay (SC) soil using a single- and split fertilizer application (kg/ha)

	Specified compliance	Sandy clay loam						Sandy clay					
		Sandy clay loam			Sandy clay loam			Sandy clay			Sandy clay		
		GM ^a (R)	Nitrogen (kg/ha)	Hectare (ha)	GM (R)	Nitrogen (kg/ha)	Hectare (ha)	GM ^a (R)	Nitrogen (kg/ha)	Hectare (ha)	GM (R)	Nitrogen (kg/ha)	Hectare (ha)
Single Split		10,014	137	1.00			9586	124	1.00				
		9524	144	1.00			9175	126	1.00				
		Upper partial moment (UPM)			Upper frequency method (UFM)			Upper partial moment (UPM)			Upper frequency method (UFM)		
Single	0.649	5823	105	0.499	10,260	106	4570	117	0.401	9646	145	0.852	
	0.696	5475	104	0.469	9542	120	4268	115	0.374	7893	164	0.709	
	0.749	5057	101	0.436	8508	125	3907	111	0.343	7579	125	0.664	
	0.795	4660	100	0.402	8025	111	3529	108	0.310	7336	125	0.643	
	0.848	4092	99	0.354	7196	85	3017	107	0.265	6752	120	0.592	
	0.895	3482	100	0.300	6190	85	2473	105	0.218	5219	124	0.457	
	0.947	2966	129	0.249	4377	58	1697	96	0.150	4495	134	0.395	
	0.649	5502	103	0.495	10,006	116	4338	122	0.394	8536	136	0.776	
	0.696	5161	102	0.465	8658	123	4068	119	0.369	7222	106	0.659	
	0.749	4791	101	0.433	8123	123	3739	116	0.340	6958	151	0.637	
Split	0.795	4388	97	0.399	7979	132	3388	112	0.308	6830	146	0.623	
	0.848	3831	91	0.352	7266	91	2907	108	0.265	6323	146	0.577	
	0.895	3183	91	0.293	6839	91	2384	101	0.218	5450	146	0.497	
	0.947	2868	138	0.251	4201	13	1619	92	0.149	4295	123	0.390	

^a GM is the estimated gross margin

Vast differences are observed when comparing the UPM-model- and UFM-model results. The areas irrigated are much higher for the UFM model, which is the main reason for the higher total gross margins optimized with the model. Fertilizer-application rates are also higher for the UFM model when considering the SC soil, with the exception of the 106 kg/ha applied in a single application at a compliance probability of 0.696. Fertilizer-application rates on the SCL soil are higher than the UPM model at low levels (0.795 and below) of compliance and lower at high levels (0.848 and above) of compliance, regardless of fertilizer-application method.

5 Conclusions

The main conclusion from this research is that the UPM method of enforcing probabilistic constraints is very conservative as is evident from the comparison with the newly developed UFM. The UPM method underestimates the actual probability with which the environmental goal is achieved as indicated in the objective function value of the model and the degree (average pollution above the goal) by which the environmental goal is exceeded. Even more important is the fact that the intensive- and extensive- margin responses necessary to satisfy the probabilistic constraint in the UPM model are much different from the optimal response optimized with the UFM model. Thus, use of the UPM may lead to the misidentification of appropriate management practices to combat pollution. The UFM generates solutions that are close to the probability bound and the responses seem more realistic when compared to those from the UPM.

The UFM is easy to use and requires no assumptions regarding the distribution of the environmental variable as the empirical data is used. The UFM behaved well during the optimization process and is much less conservative in the estimation of the trade-offs due to the probability limit that is closer to the actual probability limit displayed by the data. Although the UFM provides a stricter probability bound than the UPM there are some concerns regarding the application of the UFM. The UFM ensures compliance by ensuring that the number of deviations above the goal does not exceed the number of deviations allowed; for this reason, a fairly large number of observations are necessary to ensure probability limits close to the actual probability. Further research is necessary to determine the sensitivity of the UFM to sample size and mining of statistical outliers.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1

The upper partial moment (UPM) is defined as:

$$\rho(\alpha, t) = \int_t^{+\infty} (x - t)^\alpha f(x) dx \tag{21}$$

where α is constant greater than zero, t is a reference pollution level, x is the pollution variable and $f(x)dx$ is the probability density function.

Assume that $\theta(\alpha, t) \equiv [\rho(\alpha, t)]^{1/\alpha} > 0$ and $g \equiv t + p\theta(\alpha, t)$ for $p \geq 0$, by implication $p = \frac{g-t}{\theta(\alpha, t)}$ and $(g - t) = p\theta(\alpha, t)$.

Using the definitions above the, integral in Eq. 21 can then be expressed as the sum of two integrals with $t \leq G$

$$\begin{aligned} \rho(\alpha, t) &= \int_t^{+\infty} (x - t)^\alpha f(x) dx \\ &= \int_t^g (x - t)^\alpha f(x) dx \\ &\quad + \int_g^{+\infty} (x - t)^\alpha f(x) dx \geq \int_g^{+\infty} (x - t)^\alpha f(x) dx \end{aligned} \tag{22}$$

Which implies that:

$$\rho(\alpha, t) \geq \int_g^{+\infty} (g - t)^\alpha f(x) dx \tag{23}$$

The integral $\int_g^{+\infty} f(x) dx$ is the probability that x is larger than g . Rearranging Eq. 23 generates:

$$[g - t]^\alpha Pr[x \geq g] \leq \rho(\alpha, t) = [\theta(\alpha, t)]^\alpha \tag{24}$$

Assume that $g = t + p\theta(\alpha, t) > t$ then the stochastic inequality is:

$$Pr[x \geq t + p\theta(\alpha, t)] \leq [\theta(\alpha, t)/(g - t)]^\alpha = (1/p)^\alpha \tag{25}$$

If $\alpha = 1$, $\theta(\alpha, t) = \theta(t)$ and $q > 0$ then the linear UPM, $\theta(t)$ can be estimated as a set of linear constraints:

$$t + q\theta(t) \leq g \Rightarrow (g - t) \geq q\theta(t) \Rightarrow \frac{\theta(t)}{g - t} \leq \frac{1}{q} \tag{26}$$

which implies:

$$Pr[x \geq t + p\theta(t)] \leq \theta(t)/(g - t) \leq (1/q) \tag{27}$$

Appendix 2

See Table 3.

Table 3 Coefficients estimated for nitrate response function for nitrogen applications (kg/ha) in state of nature s on a SCL and SC soil with single and split nitrogen application

	SCL							
	Single				Split			
	Intercept	N	NN	R ²	Intercept	N	NN	R ²
S1	109.182***	0.025***	3.63E−05***	0.998	109.680***	0.008**	4.71E−05***	0.995
S2	42.506**	−0.139	6.26E−04	0.043	35.218***	0.007	2.06E−04	0.674
S3	2.882	−0.006	−1.21E−06	0.739	3.207***	−0.012***	3.18E−05*	0.877
S4	3.731	0.002	−6.74E−06	0.319	3.791***	−0.001	7.14E−06	0.42
S5	10.093***	0.013***	−4.20E−05**	0.185	10.063***	0.017	−6.76E−05	0.198
S6	18.207***	0.111***	−1.89E−04**	0.977	18.266***	0.107***	−1.47E−04	0.979
S7	10.027***	−0.018***	5.31E−05**	0.883	9.664***	−0.007	1.07E−05	0.634
S8	50.572***	0.22***	2.25E−04***	0.999	46.363***	0.389***	−1.41E−04	0.999
S9	18.287***	−0.033**	1.16E−04**	0.692	18.306***	−0.033***	1.11E−04**	0.725
S10	9.588***	−0.042**	1.42E−04*	0.595	9.602***	−0.039*	1.24E−04	0.596
S11	40.109***	−0.019	2.27E−04*	0.851	40.907***	−0.031	1.77E−04	0.463
S12	0.000	0.000	0.00E+00	0.000	0.000	0.000	0.00E+00	0.000
S13	3.669***	−0.004***	1.14E−05***	0.873	3.670***	−0.003***	1.02E−05***	0.904
S14	28.348***	0.026***	8.31E−05***	1.000	28.490***	0.022***	4.52E−05***	1.000
S15	8.616***	−0.032*	8.44E−05	0.667	7.720***	0.006	−8.15E−05	0.605
S16	35.490***	0.036	1.12E−04	0.898	37.571***	−0.005	1.41E−04	0.482
S17	20.267***	0.027	−5.01E−05	0.328	19.710***	0.022	−2.25E−05	0.338
S18	35.198***	0.114***	−4.57E−05	0.983	36.533***	0.055	1.71E−04	0.973
S19	8.727***	−0.010**	3.25E−05*	0.618	8.727***	−0.010**	3.25E−05*	0.618

	SC							
	Single				Split			
	Intercept	N	NN	R ²	Intercept	N	NN	R ²
S1	34.390***	0.156	−4.20E−04	0.493	34.289***	0.122	−2.24E−04	0.48
S2	65.939***	−0.228	9.15E−04	0.346	69.795***	−0.255	7.90E−04	0.491
S3	1.115**	0.013	−8.32E−05**	0.821	1.498***	0.000***	6.56E−08	0.99
S4	4.222***	−0.010***	1.91E−05***	0.992	4.145***	−0.006***	5.62E−06	0.971
S5	6.262***	0.026***	−1.09E−04***	0.732	6.406***	0.015*	−4.35E−05	0.658
S6	21.324***	0.021	−2.13E−04	0.822	21.711***	−0.003**	5.74E−06	0.869
S7	7.194***	0.082**	−2.75E−04*	0.593	7.149***	0.084**	−2.88E−04*	0.596
S8	170.531***	0.102***	3.17E−04***	0.997	169.803***	0.143***	2.16E−04**	0.997
S9	9.150***	−0.029	7.18E−05	0.609	8.840***	−0.010	−1.97E−05	0.636
S10	0.610	0.004	2.81E−08	0.175	2.081***	−0.034**	1.50E−04**	0.641
S11	72.166***	−0.027	1.89E−05	0.368	71.015***	−0.061	2.62E−04	0.083
S12	0.000	0.000	0.00E+00	0.000	0.000	0.000	0.00E+00	0.000
S13	2.028***	0.005**	−2.90E−05**	0.812	2.063***	0.002	−7.50E−06	0.427
S14	54.637***	−0.157***	4.80E−04***	0.967	54.119***	−0.115***	2.86E−04***	0.979
S15	5.375***	−0.041***	1.08E−04**	0.927	4.855***	−0.013	−1.65E−05	0.886
S16	50.027***	−0.125	4.92E−04	0.219	43.048***	0.093	−5.63E−04	0.529
S17	13.995***	−0.070***	2.23E−04**	0.793	14.321***	−0.065***	1.93E−04***	0.909
S18	58.010***	−0.269***	8.70E−04**	0.779	56.694***	−0.159	3.35E−04	0.65
S19	0.039	0.006**	−2.01E−05*	0.584	0.037	0.006**	−1.98E−05***	0.595

*** Significant at a 1% level; ** significant at a 5% level; * significant at a 10% level

References

- Annandale JG, Benadé N, Jovanovic NZ, Steyn JM, Du Sautoy N (1999) Facilitating irrigation scheduling by means of the soil water balance model. WRC Report No. 753/1/99. Pretoria, South Africa
- Atwood JA (1985) Demonstration of the use of lower partial moments to improve safety-first probability limits. *Am J Agric Econ* 67:787–793
- Atwood JA, Watts MJ, Helmers GA, Held LJ (1988) Incorporating safety-first constraints in linear programming production models. *West J Agric Econ* 13(1):29–36
- Berck P, Hihn JM (1982) Using the semivariance to estimate safety-first rules. *Am J Agric Econ* 64:298–300
- GAMS Development Corporation (2007a) General algebraic modeling system (GAMS) Release 22.6. GAMS Development Corporation, Washington, DC, USA
- GAMS Development Corporation (2007b) GAMS: the solver manuals. GAMS Development Corporation, Washington, DC
- Horan RD (2001) Cost-effective and stochastic dominance approaches to stochastic pollution control. *Environ Resour Econ* 18:373–389
- Kampas A, White B (2003) Probabilistic programming for nitrate pollution control: comparing different probabilistic constraint approximations. *Eur J Oper Res* 147:217–228
- Kampas A, White B (2004) Administrative costs and instrument costs for stochastic nonpoint source pollutants. *Environ Resour Econ* 27:109–133
- Kataria M, Elofsson K, Hasler B (2010) Distributional assumptions in chance-constrained programming models of stochastic water pollution. *Environ Model Assess* 15:273–281
- Li X, Lu H, He L, Shi B (2014a) An inexact stochastic optimization model for agricultural irrigation management with a case study in China. *Stoch Environ Res Risk Assess* 28:281–295
- Li YP, Zhang N, Huang GH, Liu J (2014b) Coupling fuzzy-chance constrained program with minimax regret analysis for water quality management. *Stoch Environ Res Risk Assess* 28:1769–1784
- Markowitz H (1970) *Portfolio selection*. Yale University Press, New Haven CT
- Matthews N (2014) *Modelling economic-environmental trade-offs of maintaining nitrate pollution standards*. Ph.D. thesis. Department of Agricultural Economics, University of the Free State, Bloemfontein
- McSweeney WT, Shortle JS (1990) Probabilistic cost effectiveness in agricultural nonpoint pollution control. *South J Agric Econ* 22:95–104
- Qiu Z, Prato T, Kaylen M (1998) Watershed-scale economic and environmental tradeoffs incorporating risks: a Target MOTAD Approach. *Agric Resour Econ Rev* 27:231–240
- Qiu Z, Prato T, McCamley F (2001) Evaluating environmental risks using safety-first constraints. *Am J Agric Econ* 83(2):401–413
- Ranga Prabodanie RA, Raffensperger JF, Milke MW (2010) A pollution offset system for trading non-point source water pollution permits. *Environ Resour Econ* 45:499–515
- Richardson JW, Klose SL, Gray AW (2000) An applied procedure for estimating and simulating multivariate empirical (MVE) probability distributions in farm-level risk assessment and policy analysis. *J Agric Appl Econ* 32(2):299–318
- Rossouw JN, Görgens AHM (2005) Knowledge review of modelling non-point source pollution in agriculture from field to catchment scale. WRC Report No. 1467/1/05. Water Research Commission, Pretoria
- Samarawickrema AK, Belcher KW (2005) Net greenhouse gas emissions and the economics of annual crop management systems. *Can J Agric Econ* 53(4):385–401
- Shortle JS, Horan RD, Abler DG (1998) Research issues in nonpoint pollution control. *Environ Resour Econ* 11(3–4):571–585
- Shortle JS, Ribaud M, Horan RD, Blandford D (2001) Reforming agricultural nonpoint pollution policy in an increasingly budget-constrained environment. *Environ Sci Technol* 46:131–1325
- Stoorvogel JJ, Antle JM, Crissman CC, Bowen W (2004) The tradeoff analysis model: integrated bio-physical and economic modelling of agricultural production systems. *Agric Syst* 80:43–66
- Teague ML, Bernardo DJ, Mapp HP (1995) Meeting environmental goals efficiently on a farm-level basis. *Rev Agric Econ* 17:37–50
- Van Der Laan M (2009) *Development, testing and application of a crop nitrogen and phosphorus model to investigate leaching losses at the local scale*. Ph.D. thesis in Agronomy. Faculty of Natural and Agricultural Sciences. University of Pretoria, Pretoria
- Wang X, Yang H, Cai Y, Yu C, Yue W (2016) Identification of optimal strategies for agricultural nonpoint source management in Ulansuhai Nur watershed of Inner Mongolia, China. *Stoch Environ Res Risk Assess* 30:137–153
- Weersink A, Jeffery S, Pannell D (2002) Farm-level modeling for bigger issues. *Rev Agric Econ* 24(1):123–140
- Weersink A, Jeffery S, Pannell D (2004) *Modelling farming system linkages*. In: Filson GC (ed) *Intensive agriculture and sustainability: a farming systems analysis*. UBC Press, Vancouver
- Zhu M, Taylor DB, Sarin SC, Kramer RA (1994) Chance constrained programming models for risk-based economic and policy analysis of soil conservation. *Agric Resour Econ Rev* 23(1):58–65