Establishment and application of the estimation model for pollutant concentration in agriculture drain

Qiangkun Li1,2, Yawei Hu1,2,3, Qian Jia2 and Changji Song3

1 Xi’an University of Technology, X’ian, Shaanxi, China; 2 Yellow River Institute of Hydraulic Research, YRCC, Zhengzhou, Henan, China. 3 huyawei@126.com

Abstract. It is the key point of quantitative research on agricultural non-point source pollution load, the estimation of pollutant concentration in agricultural drain. In the guidance of uncertainty theory, the synthesis of fertilization and irrigation is used as an impulse input to the farmland, meanwhile, the pollutant concentration in agricultural drain is looked as the response process corresponding to the impulse input. The migration and transformation of pollutant in soil is expressed by Inverse Gaussian Probability Density Function. The law of pollutants migration and transformation in soil at crop different growth periods is reflected by adjusting parameters of Inverse Gaussian Distribution. Based on above, the estimation model for pollutant concentration in agricultural drain at field scale was constructed. Taking the of Qing Tong Xia Irrigation District in Ningxia as an example, the concentration of nitrate nitrogen and total phosphorus in agricultural drain was simulated by this model. The results show that the simulated results accorded with measured data approximately and Nash-Sutcliffe coefficients were 0.972 and 0.964, respectively.

1. Introduction

In recent years, with emerging of the outstanding hydro-environmental problems, as well as some improvement of point source pollution treatment, non-point source pollution, especially the agricultural non-point source pollution caused by a large-scale use of chemical fertilizers and pesticides, has increasingly attracted more and more concern and attention [1]. Agricultural non-point source pollution refers to the process in which a variety of pollutants (salt, nutrients, pesticides, bacteria, etc.) produced from agricultural production activities, through farmland runoff, agricultural drain and underground seepage, in the form of a low concentration and a wide range, spread from pedosphere to hydrosphere. Since 1970s, the treatment practice in non-point source pollution at home and abroad has indicated that quantitative load research on non-point source pollution lays a basis to control, evaluate and manage it [2]. Thus, the drain estimation of the farmland irrigation and prediction of pollutant concentrations have become the two vital aspects in the research. Recent years’ studies have shown that the soil moisture transfer based farmland irrigation and drainage models have made a great progress, such as development and application of SWAP (soil-water-atmosphere-plant), and DRAINMOD models [3-5]. In comparison, because of the complex nature of the solute transport in soil, the model for estimation of concentration of pollutants in agricultural drainage has been developed slowly, which is difficult to adapt to the needs of agricultural non-point source pollution control and management.
Based on the current study and the development of the random theory, using inverse Gaussian distribution as a probability density function, an estimation model for concentration of pollutants in agricultural irrigation and drainage in field scale was built in the paper. By testing and practice, this model has shown better simulation results, with simple structure, and is easy to use.

2. Model construction

2.1. Principle

Soil moisture and contaminant transport in the soil is a very complex process. The joint solution of Richards equation and the convection-dispersion equation (Convection-Dispersion Equation, CDE) can accurately display the major features of the soil column in laboratory and the water movement of soil profile in field position and contaminant transport [6-7]. However, because the soil texture and pore structure in the area coverage have a strong spatial variability and randomness in field scale, it is unlikely to acquire its spatial variability by measurement in practice, that inevitably leads to the hydrodynamic dispersion theory at the micro-level, difficult to be applied to macroscopic dispersion process at field scale [8]. As early as 1980s, dozens of scholars believed that the pores of the soil are very complex. Solute transport within the soil presents unstable performance, on which stochastic process approach is appropriate to be used for the study on the quantitative characteristics [9-10]. The Transfer Function Model (TFM) theory put forward by Jury et. al (1982) is representative. In the model, Jury characterized output of the solute as a function of input flux, while the solute in the soil which is occurring in the dynamic processes, is expressed by probability density function [11-12].

Based on the above conclusions, the pollutant concentration estimate in agricultural drain, can be summarized as shown in Figure 1. \( C_{in}(t - t') \) is to simulate injection of a pollutant (nutrient) concentration of the time \( t' \) before the moment of \( t \), which equals the “composition” of every farming fertilization and closely following irrigation (precipitation) process. In addition, the impact of such factors as fertilizer intensity and amount of irrigation can be reflected. Compared with longer growing season of crops, each process of fertilizing can be regarded as a pulse input to field. "mixed reactor" stands for migration and transformation process of the pollutants in the soil with moisture as well as crop root absorption in different periods, groundwater and other hydraulic characteristics which affect that process, which is adaptable to both conservative and non-conservative pollutants. Discharged pollutant concentration to the outside (drain ditch) by "Mixed reactor" can be referred to as a response process of a pulse input expressed by the transfer function of \( f(t) \), being the integrated function of a "mixed reactor" . Its essence is a response process in that a unit pulse is input to a "mixed reactor," then pollutant concentration emits to the outside, that is called the concentration of unit hydrograph. \( C_{out}(t) \) stands for the response to the process of pollutants concentration in agricultural drain where \( C_{in}(t - t') \) is the injected pulse concentration into the soil.

![Figure 1. Schematic diagram for estimation principle of pollutant concentration in agricultural drain.](image-url)
Therefore, the concentration of pollutants in agricultural drain can be expressed by the convolution integral as:

$$C_{	ext{out}}(t) = \int_0^\infty C_{\text{in}}(t-t')f(t')dt'$$ (1)

The symbol meaning is the same as the above, the unit of the concentration is mg/L.

2.2. Determination of \( f(t) \)

The process of pollutants in the soil with moisture migration and transformation is affected by many factors, the mechanism is very complex, and there is also no law to follow for its output process. Accordingly, it is requested the transfer function \( f(t) \) to have a very good flexibility and adaptability, while the relevant parameters in \( f(t) \) should easily determined due to limited availability of the agricultural drain monitoring data of the concentration of pollutants currently. After comprehensive comparisons, Inverse Gaussian Distribution is selected as the transfer function in Equation. (1) for the study.

Inverse Gaussian distribution as a probability density function was proposed by Tweedie in 1945, with a number of similar features to Gaussian distribution. Gaussian distribution describes a distribution of the distance in a fixed point in the Brownian motion, while the inverse Gaussian distribution is a distribution of the time required to reach a fixed distance, which has a variety of expression forms, one of its common expressions is:

$$f(t) = \frac{\lambda}{\sqrt{2\pi\mu^3}} \exp\left[-\frac{(t-\mu)^2}{2\mu^2\lambda}\right] \quad (t > 0)$$ (2)

In the above, equation the value intervals of parameters \( \mu \) and \( \lambda \) are taken as \((0, \infty)\), both of them have the same dimension of \( t \), which is the dimensionless of independent variables. Inverse Gaussian distribution is a curve with a single peak positive side, the area surrounded by the curve is a unit area, and the shape of the curve depends on the values of \( \mu \) and \( \lambda \). With the change of \( \mu \) and \( \lambda \), inverse Gaussian distribution can be transformed into a wide range of curves of good flexibility and adaptability. In addition, the inverse Gaussian distribution also has a clear physical meaning and the concept of density function itself is a pulse solution of convection-diffusion equation, therefore, using inverse Gaussian distribution as response process of an input pulse field has an adequate theoretical basis.

Inverse Gaussian distribution is provided with many good statistical properties, which have been widely applied to in many sectors. In terms of hydrology and sediment calculations, Moore (1984) once had utilized inverse Gaussian distribution function to continuously simulate sediment transport process, from which satisfactory results had been drawn [13]. Gydesen() had used a continuous function of inverse Gaussian distribution in the study of the leaching of chemical substances in soil, to fit experimental data for indoor soil column, producing good results. In the non-point source pollution research, Li Huai-en(1996) et al. based on inverse Gaussian instantaneous unit hydrograph models, established a set of non-point source pollution mathematical models of mechanism nature that includes watershed runoff model, convergence model, pollutant yield model and non-point source pollution transport model completely. Its application results have shown better practicability [14].

According to the nature of mathematical statistics, the initial value of parameters \( \mu \) and \( \lambda \) in inverse Gaussian distribution can be respectively calculated by the moment method and the maximum likelihood estimation method [15]. As shown in Equations (3) and (4), the initial parameter values after estimated, can be further optimized using the recommended way in reference [16] to obtain the optimal parameters.

Moment method of estimation:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i \quad \lambda = \frac{\bar{x} \cdot \mu^2}{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$ (3)
Maximum likelihood estimation method:

\[
\mu = \frac{\sum_{i=1}^{n} x_i}{n}, \quad \lambda = \frac{\mu^2}{\left( \frac{1}{n} \sum_{i=1}^{n} x_i \right)^2} - \frac{1}{n} \sum_{i=1}^{n} x_i
\]

In the above equations, \( n \) is the number of tests for monitoring; \( x_i \) for the \( i \) time test data; \( \bar{x} \) is the mean value of the test data; other symbolic meaning is as similar as the above.

To sum up, Equations (1) and (2) are a complete expression indicating pollutants concentration estimation model for farmland drain proposed in the paper and Equations. (3) and (4) can be used for estimating initial parameters of the model. The optimized parameters obtained are substituted into Equation. (2) to calculate unit line of concentration. Then the pulse concentration \( C_{m}(t-t') \) is injected into the soil, and Equation. (1) is used for calculating the concentration of pollutants in the irrigation drain continually.

3. Model utilization

3.1. Test survey

The test area is located in the Xiyu Village, east of Wanghong Town, Yongning County, Qingtongxia Irrigation District, Ningxia Autonomous Region, belonging to arid zone, with long term average precipitation of 170 ~ 210 mm, of which that from July to September accounts for about 70% of the total. The annual evaporation is about 1 000 ~ 1 500 mm. The silty loam soil in the area is mainly cultivated in paddy-upland rotating, rice grows in paddy field, while wheat, corn and so on are planted in dry fields, rotating every other year. The strip land in the area is about 600 m in length, the field is drained by ditches, that is laid out in parallel with equal spacing of 100 m and depth of 80 ~ 100 cm, and Each ditch controls an area of about 6 hm². Agricultural non-point source pollution in the area includes a number of pilot projects. This study intends to take two drain ditches of paddy fields as the research object, in which No.1 ditch is the trial ditch, providing test data for initial value for the parameter estimation, and No.2 ditch is analog groove, its monitoring data is to verify the accuracy of the model. The two drains lie approximately 500 m away.

The test was conducted in May ~ September 2008. Rice was planted at the end of early May, irrigation started on May 8, water sampling was taken since May 9 at the end of agricultural ditch, once every five days until the completion of rice field drainage process. Water samples are extracted from 1 / 2 water depths by the stream central line method, each time is taken 1000 ml, analytical indexes include ammonium nitrogen (NH₄⁺-N), nitrate nitrogen (NO₃-N), total nitrogen (TN), total phosphorus (TP) and so on. In accordance with Water and Wastewater Monitoring Analysis Methods (fourth edition), ammonium nitrogen uses Nessler's reagent colorimetry, nitrate nitrogen uses ion chromatography, total nitrogen uses K₂S₂O₈ oxidization-ultraviolet spectrometry, phosphorus uses molybdenum-antimony anti-spectrophotometric method.

As the previous research [17] indicates that nitrate nitrogen is the main reason of nitrate loss in drainage, the study takes nitrate nitrogen and total phosphorus as simulative object.

3.2. Parameters calculation

In order to meet rice-growing needs, except for fertilizing base fertilizer, top dressing nitrogen and phosphate fertilizer have to be applied at turning green - tillering stage (from the end of May to early June) and top dressing nitrogen fertilizer has also to be added at jointing - heading stage (from the end of July to the beginning of June). Subject to some influence from crops root absorption as well as groundwater levels changing, the migration and transformation process of nitrogen, phosphorus and other pollutants in the soil is different in different growth periods, leading to variation of the pollutant concentrations in farmland drainage. Thus, nitrate nitrogen, in the light of fertilizing process, is
simulated in three stages, i.e. seedling, turn green - tillering and jointing - heading; while total phosphorus is divided into two stages: seedling and late stage.

According to nitrate and total phosphorus concentration monitored in varying stages of field drainage of No.1 drainage ditch, the initial values $\mu$ and $\lambda$ are estimated by using Equation (3) i.e. the moment method, and optimal parameters are obtained by using accelerated optimization method as shown in Table 1.

**Table 1. Initial parameters and optimum results.**

| Different growth stages of rice | Initial value | Optimum result | Initial value | Optimum result |
|---------------------------------|---------------|----------------|---------------|----------------|
| (NO$_3^-$-N)                    | $\mu$ | $\lambda$ | $\mu$ | $\lambda$ | $\mu$ | $\lambda$ | $\mu$ | $\lambda$ |
| Seedling                        | 17  | 391 | 25 | 391 | 0.23 | 2.25 | 37 | 1200 |
| Turning green - tillering       | 24  | 93 | 16 | 480 | 4.44 | 105 | 33 | 345 |
| Jointing - heading              | 32  | 371 | 15 | 240 |

3.3. *Unit line of concentration*

After substituting the optimized parameters of Table 1 into Equation (2), the response process of corresponding pulse input $f(t)$ for rice in different growth periods can be figured out, that is the pollutant concentration in the drainage of farmland when the unit concentration of pollutants is injected into the field in different growth periods. To connect $f(t)$ values for different simulated time periods, the changing process of the pollutants concentration in drainage in the whole growth period is created, which is known as concentration unit line. Figures 2 and 3 show respectively the concentration unit line of nitrate nitrogen and total phosphorus.

![Figure 2. Concentration unit line of nitrate nitrogen in agricultural drain.](image-url)
3.4. Simulation of pollutant concentration in the drainage

The survey and simulation of drainage ditch to control fertilizing process of paddy land is shown in Table 2.

The survey data of amount of fertilizer listed in Table 2 are investigated from farmers, of which ammonium bicarbonate includes 17.7% nitrogen, ammonium phosphate has 12% phosphorus, and urea has 45% nitrogen. When paddy fields are irrigated the first time, irrigation quantity of 15 cm is measured at the weir of the irrigated canal intake, followed by water depth of 10 cm that is an average of the measured multi-point depth. Converted concentration is a concentration of volume ratio by nitrogen and phosphorus after dissolved in the corresponding water.

| Stage of fertilization | Date       | Type of fertilization | Amount of fertilization (kg/hm²) | Contented (kg/hm²) | Irrigation quantity (field depth) (cm) | Converted concentration (mg/l) |
|------------------------|------------|-----------------------|----------------------------------|--------------------|----------------------------------------|--------------------------------|
| Base fertilizer        | before planting | ammonium bicarbonate | 750                              | 132.75             | 15                                     | 89                             |
|                        |             | ammonium phosphate   | 125                              | 15                 |                                        | 10                             |
| Top dressing in turning green-tillering | 30-May | urea                  | 225                              | 101.25             |                                        | 101                            |
|                        |             | ammonium phosphate   | 150                              | 18                 | 10                                     | 18                             |
| Top dressing in jointing - heading | 30-Jun | urea                  | 255                              | 114.75             | 10                                     | 115                            |

The converted concentration $C_m(t - t')$ in Table 2, as the injected concentration into paddy fields, is substituted into Equation (1), the process of simulation for concentration of nitrate nitrogen and total phosphorus in agricultural drain by drainage ditch can be extrapolated, see Figures 4 and 5.
Figure 4. Simulation for concentration of nitrate nitrogen in agricultural drain.

Figure 5. Simulation for concentration of total phosphorus in agricultural drain.
is synthetic of the fertilizing intensity and irrigation quantity, therefore, even under different situations of fertilization and irrigation, different $C_{an}(t - t')$ can be calculated, and then the corresponding concentration of pollutants in agricultural drainage can be estimated.

3.5. Simulated results verification

After the nitrate and total phosphorus concentration measured from No.2 ditch were plotted on Figures 4 and 5, we can see that both of them have a good integration. The results of simulation can basically reflect the change of nitrate and total phosphorus concentration in farmland drainage.

Nash-Sutcliffe is adopted to simulate efficiency coefficient (NSC) so that the results of model estimation can be assessed. The formula is:

$$NSC = 1 - \frac{\sum (x_{obs} - x_{calc})^2}{\sum (x_{obs} - \bar{x}_{obs})^2}$$

$x_{obs}$ is the measured value, $x_{calc}$ is the model calculated value, $\bar{x}_{obs}$ is the arithmetic mean of the measured values. When the model calculated values and the actual monitoring values are equal, $NSC = 1$, the best simulation results are achieved; usually the value of $NSC$ is in the range of 0 and 1, the larger $NSC$ is, the better the calculated and observed values are matched.

The simulated efficiency coefficient of the concentration of nitrate NSC equals 0.972, and simulation of the concentration of the total phosphorus NSC is 0.964, that indicates that the results of model are more close to the actual monitoring data, and simulation results are good.

4. Conclusions

It is the key point of quantitative research on agricultural non-point source pollution load the estimation of pollutant concentration in agricultural drain. In the guidance of uncertainty theory, the synthesis of fertilization and irrigation is used as an impulse input to the farmland, meanwhile, the pollutant concentration in agricultural drain is considered to be the response process corresponding to the impulse input. The migration and transformation of pollutant in soil are expressed by Inverse Gaussian Probability Density Function, the law of pollutants migration and transformation in soil at crop different growth periods is reflected by adjusting parameters of Inverse Gaussian Distribution. Based on above, the estimation model for pollutant concentration in agricultural drain at field scale was constructed. Taking the of Qing Tong Xia Irrigation District in Ningxia as an example, the concentration of nitrate nitrogen and total phosphorus in agricultural drain was simulated by this model. The results show that the simulated results accorded with measured data approximately and Nash-Sutcliffe coefficients were 0.972 and 0.964, respectively.

Acknowledgement

(1)The research is supported by The National Natural Science Foundation of China (NO: 51379085).
(2)Central public welfare scientific research basic scientific research and special funds: Regional diagnosis model and agricultural non-point source pollution control technology research (HKY - JBYW - 2018-04);

References

[1] LI Li-Hua, LI Qiang-Kun 2014 The Progress and Trends of Agricultural Non-point Source Pollution Research Journal of Agricultural Resources and Environment 31 (1) 13-22
[2] LI Qiang-kun, LI Huai-en, Hu Ya-wei and Sun Juan 2008 Estimation of non-point source pollution loading on Tongguan section of the Yellow River Advances in Water Science 19 (4) 460-466
[3] Wang Shao-li, Xu Di and S.O. Prashe 2006 Evaluation of SWAP-MANIMEA model for predicting NO-3-N leaching from hog waste manured plots Journal of Hydraulic Engineering 37 (1) 89-96
[4] Wang Shao-li, Wang Xing-kui and S.O. Prashe 2006 Field application of DRAINMOD model to the simulation of water table, surface runoff and subsurface drainage Transactions of the Chinese Society of Agricultural Engineering 22 (2) 54-58

[5] Luo Wan, Jia Zhong-hua, and R.W. Skaggs 2006 Rice fields drainage process simulation in Yinnan irrigation district with DRAINMOD model Transactions of the Chinese Society of Agricultural Engineering 22 (9) 53-57

[6] Bresler E 1973 Simultaneous transport of solute and water under transient unsaturated flow conditions Water Resources Research 9(4) 975-986

[7] Yang J.Z. 1988 Experimental and numerical studies of solute transport in two-dimensional saturated-unsaturated soil Journal of Hydrology 97(3) 303-322

[8] Yang Jin-zhong, Cai Shu-ying and Ye Zi-tong 1998 A review of stochastic analysis for solute transport at regional scale Advances in Water Science 9 (1) 84-98

[9] Gedeon Dagan and Eshel Bresler 1979 Solute dispersion in unsaturated heterogeneous soil at fields scale Theory Soil Science Society of America Journal 43 461-467

[10] Jury W. A, Stolzy L. H. and Shouse P 1982 A field test of the transfer function model for predicting solute transport Water Resources Research 18 (2) 369-375

[11] Jury W. A, Sposito G. and White R, E 1986 A transfer function model of solute transport through soil, 1 Fundamental Concepts Water Resources Research 22(2) 243-247

[12] White R. E, Dyson J S and Haigh R A 1986 A transfer function model of solute transport through soil, 2 Illustrative applications Water Resources Research 22(2) 248-254

[13] Moore R.J.A 1984 Dynamic Model of Basin Sediment Yield Water Resources Research 20(1) 89-103

[14] Li Huai-en, Shen Jin and Liu Yu-sheng 1997 The establishment and case study of the model for non-point source pollution for watershed Acta Scientiae Circumstantiae 17 (2) 141-146

[15] Jin Xiu-yan 2006 The estimation of Inverse Gaussian Distribution parameters Journal of Changchun Normal University Natural Science 25(4) 25-26

[16] Li Huai-en and Shen Jin 1996 Non-point source pollution mathematical model Xi'an: Northwest Industry University Press 40-68

[17] Li Qiang-kun, Li Huai-en, and Hu Yawei 2008 Nitrogen loss in Qingtongxia irrigation area Journal of Agro-environment Science 27 (2) 683-686