EGU General Assembly 2020

Online | 4–8 May 2020

The Information Entropy Prisms on Riverine Water Quality Evolution

Tianrui Pang[1,2], Jiping Jiang[1,2,*,#], Yi Zheng[1], Sivakumar Bellie[3], Tong Zheng[2]

1.Southern University of Science and Technology , Shenzhen, China
2.Harbin Institute of Technology, Harbin, China
3.Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai, India

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Email: jiangjp@sustech.edu.cn

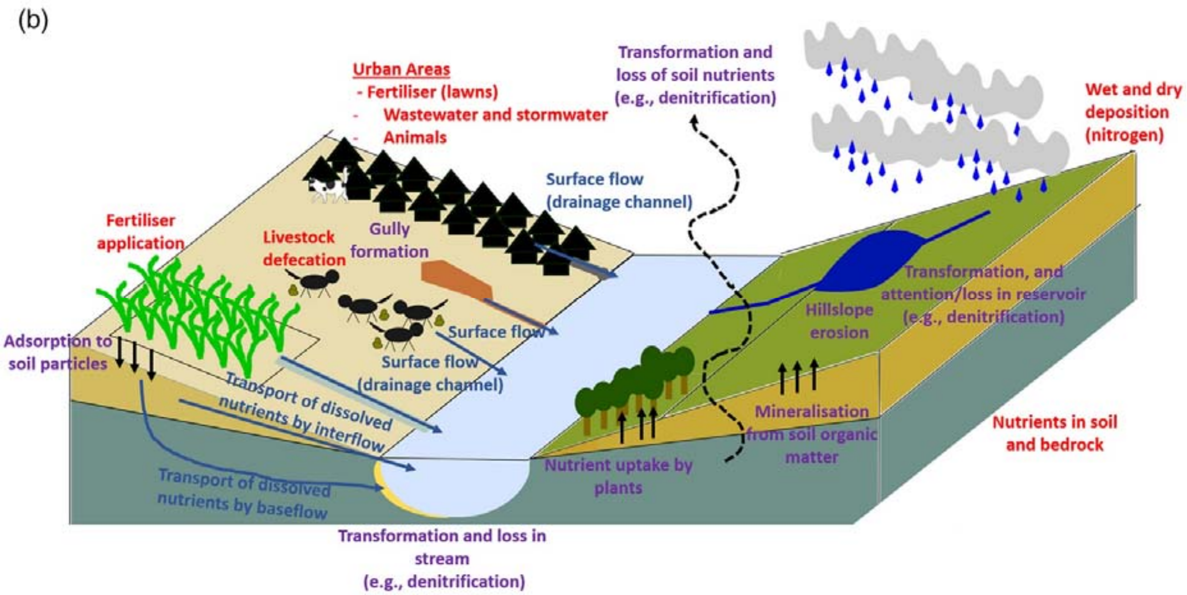
哈爾濱工業大學 Harbin Institute of Technology

Contents

1. Introduction

River water quality system

- River water quality is an open, dynamic, complex, and non-linear system.
- The complexity of river water quality system is reflected in two aspects: one is the complexity of pollutant components, the other is the complexity of pollutant migration and transformation process.



(Lintern et al. 2017)

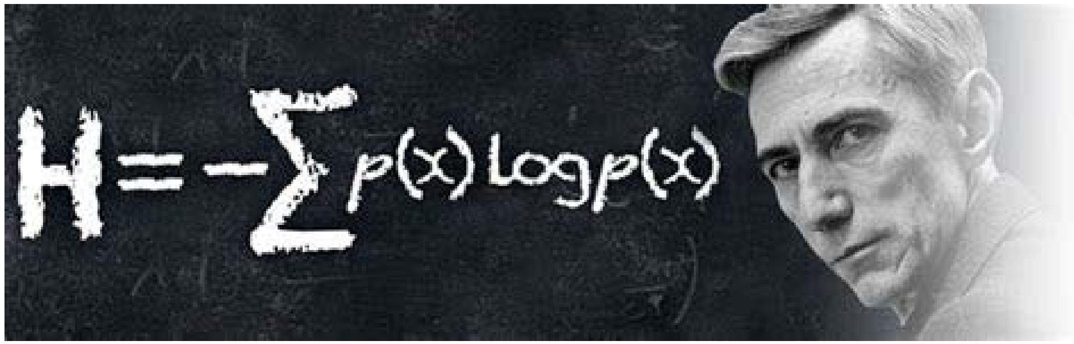
- Mass flow (oxygen, nutrients, suspend solid) , energy flow (e.g. temperature) in surface water has been extensively investigated. However few studies focus on the information flow.
- Information entropy theory has been largely applied in hydrological modeling and engineering optimization.

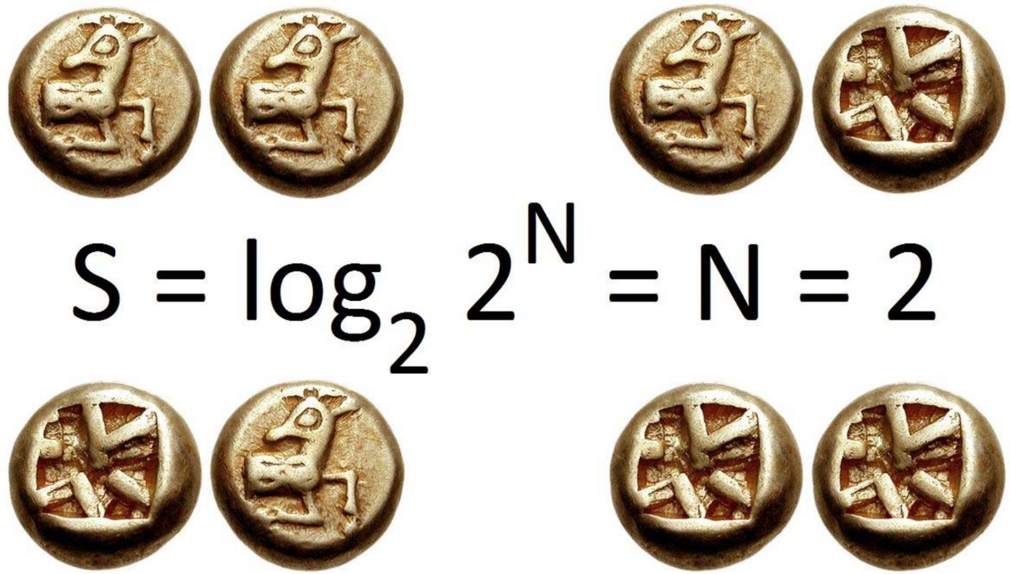
1. Introduction

Information entropy theory

- Entropy is a parameter that describes the disorder of an object.
- Shannon(1948) described information as ‘a decrease of uncertainty’, and gave the definition of information entropy.
- A random variable X follows the distribution $P(X=x_i)=p_i$, i=1,2,…,n. Then the uncertainty or information entropy of X is defined as

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i$$





1. Introduction

Information entropy theory

Joint entropy

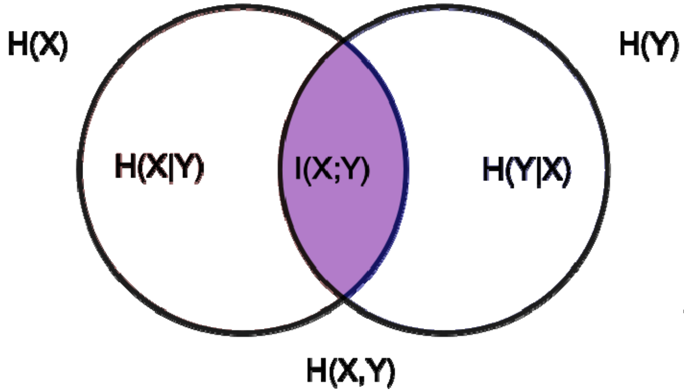
The joint entropy is based on the joint distribution of two random variables X and Y, indicating the total information content contained both in X and Y.

$$H(X,Y) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p_{ij}\log p_{ij}$$

Conditional entropy

The conditional entropy measures the information contained in X after the signal Y has been received.

$$H(X|Y) = -\sum_{x_i,y_j} p(x_i|y_j)\log p(x_i|y_j)$$
$$= H(X,Y) - H(Y)$$



Mutual entropy

The mutual entropy measures the redundant or mutual information between X and Y.

$$I(X;Y) = H(X) - H(X|Y)$$
$$= H(X) + H(Y) - H(X,Y)$$

Information transport index (ITI)

The information transport index is defined after normalizing transinformation.

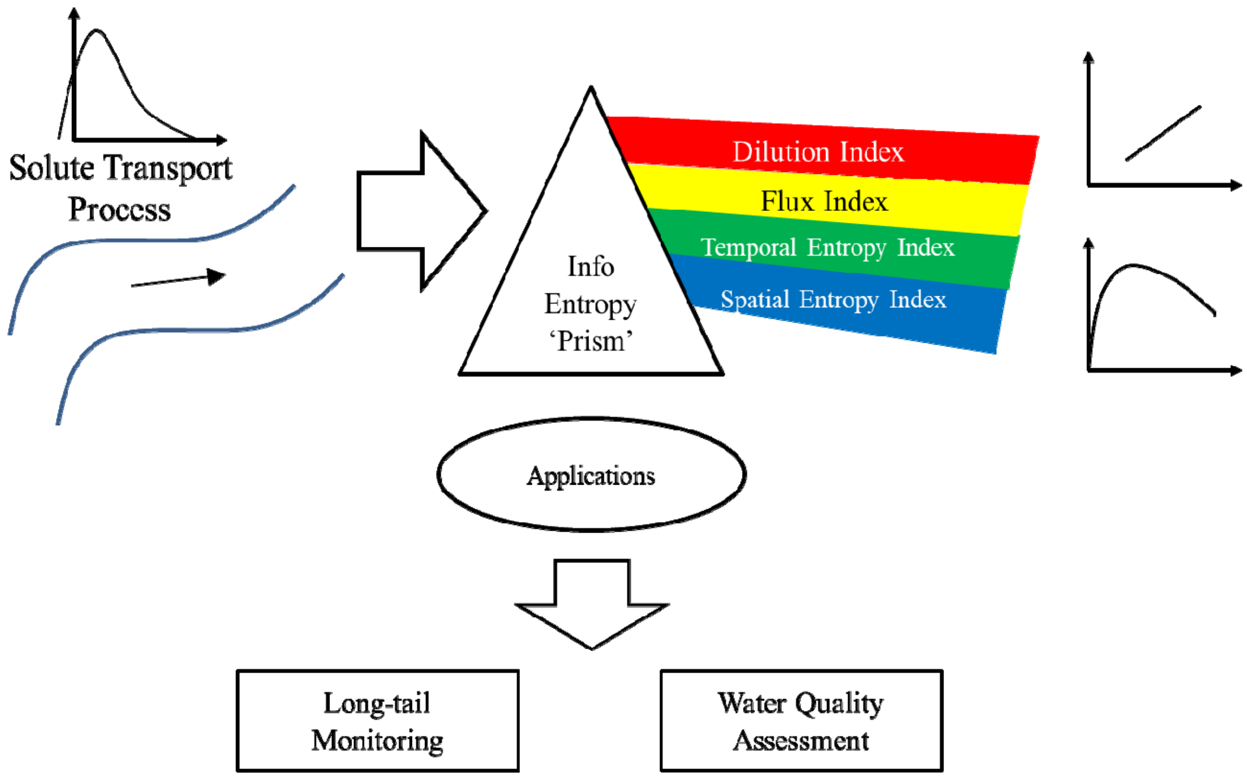
$$ITI = \frac{I(X;Y)}{H(X,Y)}$$

1. Introduction

A comprehensive information entropy based analysis framework : combined optical system with Optical Sources-Filters-Prisms-Images.

These components defined the probability space, statistical sampling method and system boundary of the statistical information entropy here.

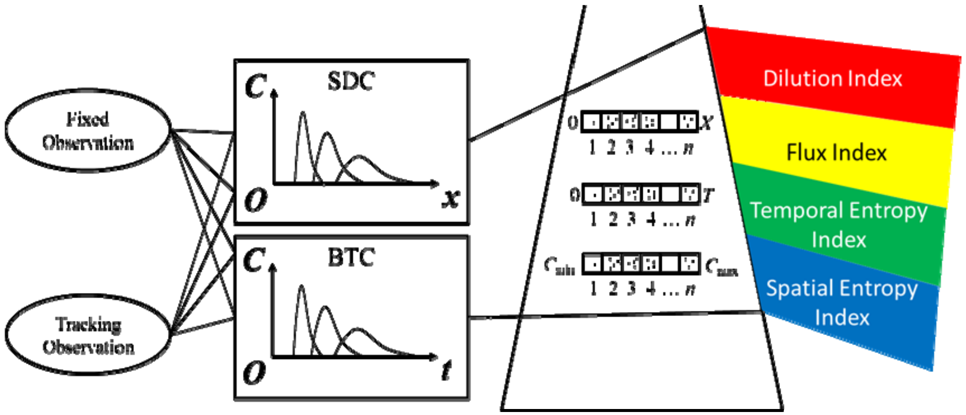
It is expected to uncover the new insights on the river water quality process (i.e. solute transport process including mixing, diffusion, reaction etc.) from the information flow perspective.



2. Methodology

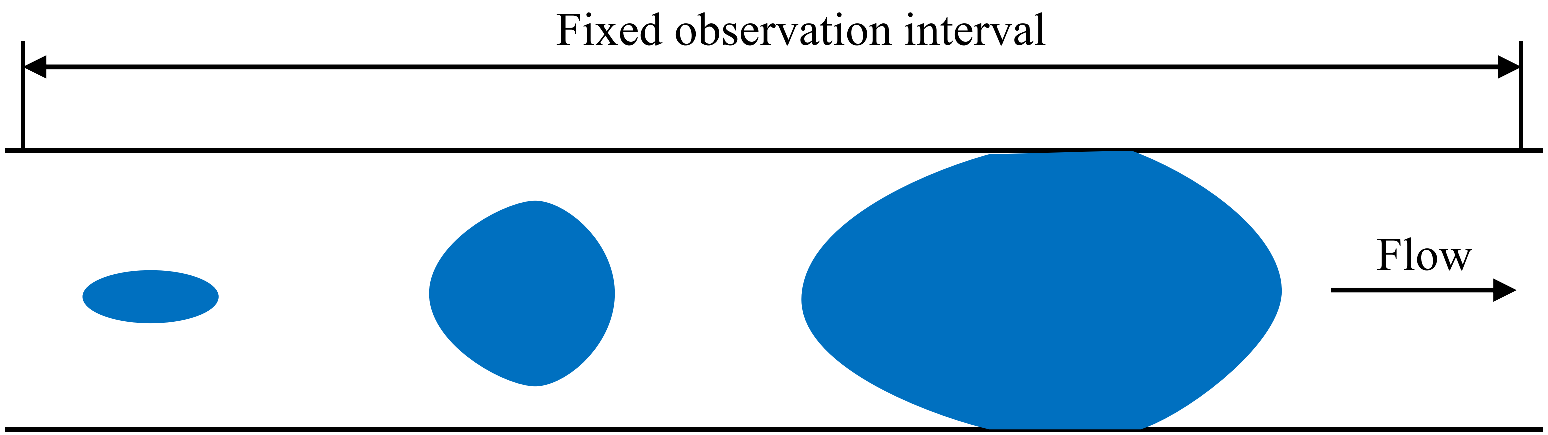
Four types of information flow-index

A comprehensive information entropy based analysis framework were proposed, which works like a combined optical system with Optical Sources-Filters-Prisms-Images. We established four basic probability space, leading to four basic information entropy indexes: Dilution index (E), Flux index (F), Spatial entropy index (Gx) , and Temporal entropy index (Gt).



Information entropy index	Expression	Expression of $P_{\cdot,k}$	Notes
Dilution index (E)	$E(t)=-\sum_{k=1}^{n}P_{E,k}(t)\ln P_{E,k}(t)$	$P_{E,k}(t)=\frac{C(x_k,t)V_k}{\sum_{i=1}^{n}C(x_i,t)V_i}$	SDCs
Flux index (F)	$F(x)=-\sum_{k=1}^{n}P_{F,k}(x)\ln P_{F,k}(x)$	$P_{F,k}(x)=\frac{C(x,t_k)Q_k}{\sum_{i=1}^{n}C(x,t_i)Q_i}$	BTCs
Spatial entropy index (Gx)	$G_x(t)=-\sum_{k=1}^{n}P_{G_x,k}(t)\ln P_{G_x,k}(t)$	$P_{G_x,k}(t)=\frac{\int_{C_{k-1}<C(x,t)<C_k}\mathrm{d}x}{\int_{C_{\min}<C(x,t)<C_{\max}}\mathrm{d}x}$	
Temporal entropy index (Gt)	$G_t(x)=-\sum_{k=1}^{n}P_{G_t,k}(x)\ln P_{G_t,k}(x)$	$P_{G_t,k}(x)=\frac{\int_{C_{k-1}<C(x,t)<C_k}\mathrm{d}t}{\int_{C_{\min}<C(x,t)<C_{\max}}\mathrm{d}t}$	

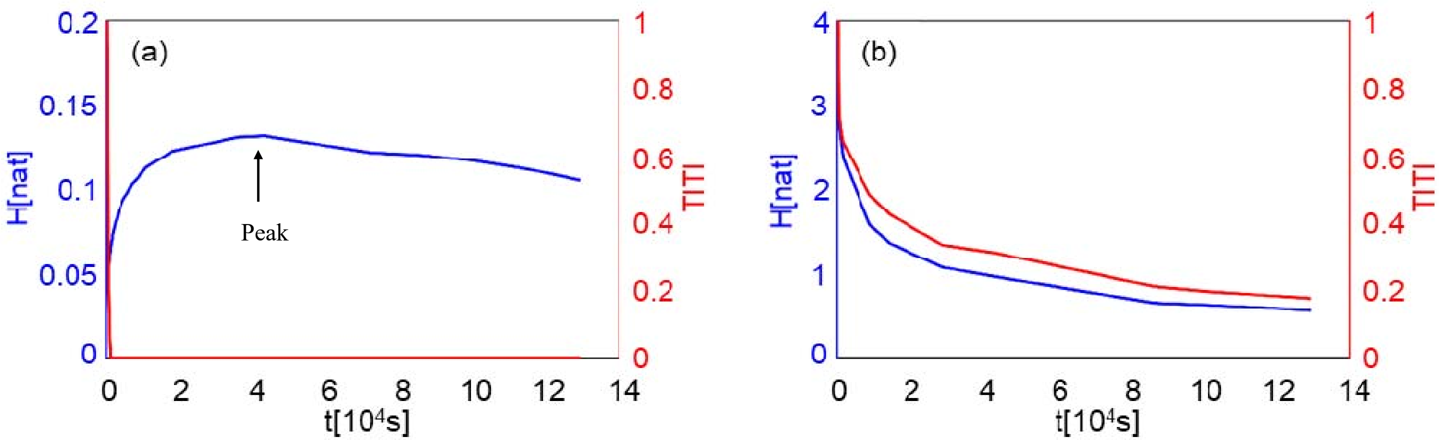
	Fixed observation based on the whole stream segment	Tracking observation based on pollutant plume
Cross section breakthrough curve	Spatial changes under fixed system	Spatial changes under dynamic system
Spatial distribution curve	Temporal changes under fixed system	Temporal changes under dynamic system



3.2 Temporal and Spatial Features

The maximum value of information entropy in fixed observation is 0.13nat, which occurs at t=40000s, about 30% in the first part of fixed observation interval. The information transfer index declined rapidly from SITI=1 to nearly 0 at t=2000 s, and then stabilized.

Under dynamic observation, the information entropy decreases with the increase of time, and there is no expected maximum point. After t = 40000s, the information entropy decreases to 1 nat, and then the decrease is slowed down. The trend of information transfer index and information entropy decrease is approximately the same.

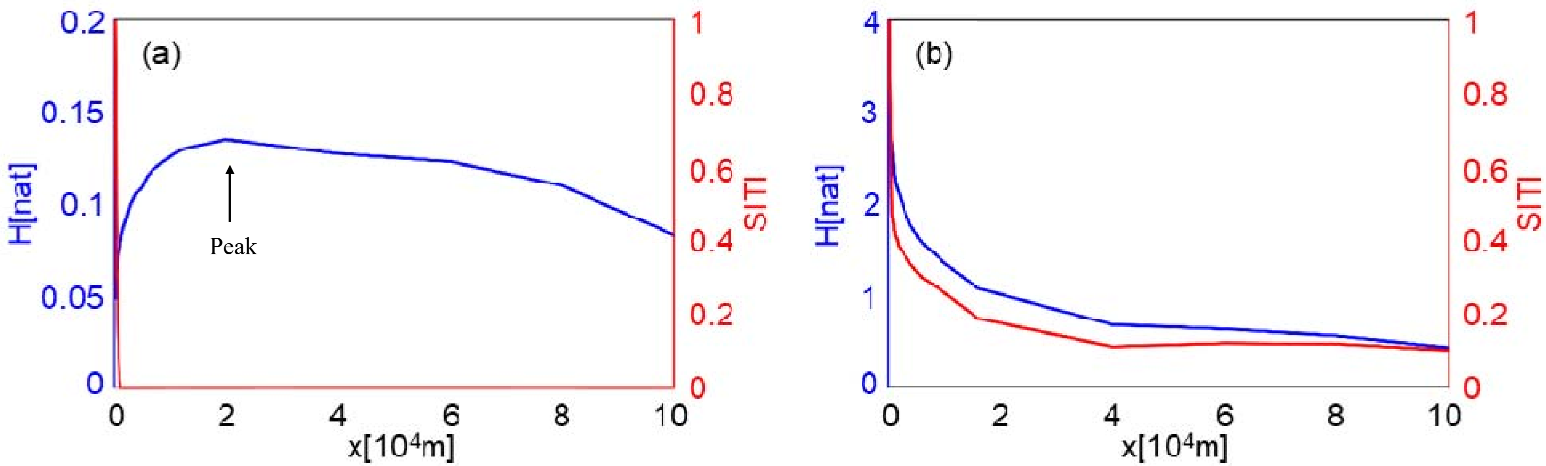


Temporal changes of entropy and ITI under fixed system (a) and dynamic system (b)

3.2 Temporal and Spatial Features

The spatial variation of information entropy and information transfer index is similar to that of time variation. The peak value of information entropy appears at x=20000m under fixed observation, about 20% before fixed observation interval, and the maximum value of information entropy is 0.13nat. The information transfer index dropped to nearly 0 at x=2000m.

Under dynamic observation, the information entropy decreases to 1nat after x = 20000m, and then the decrease slows down. The trend of information transfer index and information entropy decrease is approximately the same.



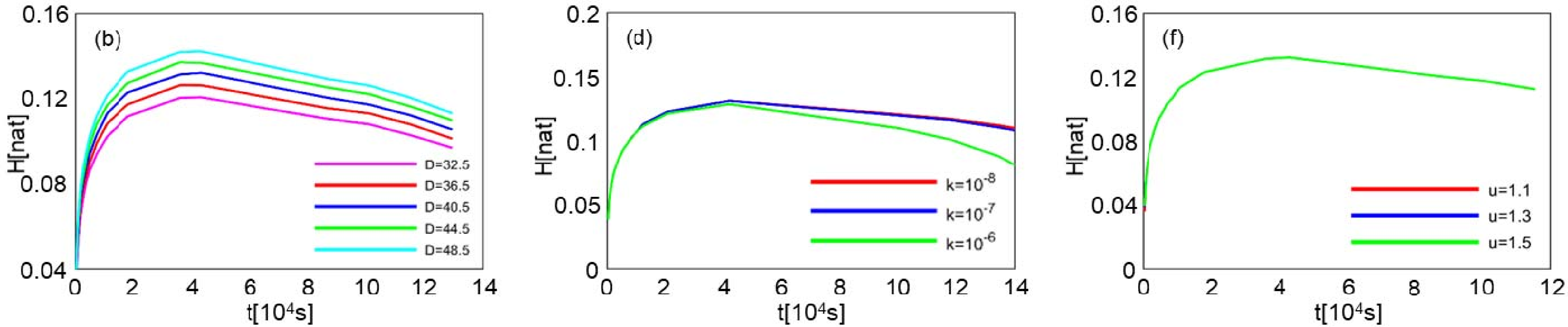
Spatial changes of entropy and ITI under fixed system (a) and dynamic system (b)

3.3 Local sensitivity analysis

Temporal variation of information entropy under different water quality model parameters (local sensitivity analysis)

The information entropy is a function of paclet number variant $\frac{\sqrt{D_x t}}{X}$ and m (only correspond with k) in terms of the analytic solution

$$H = -\left(1 - \frac{4\sqrt{-D_x t \ln m\lambda_1}}{X}\right)\ln\left(1 - \frac{4\sqrt{-D_x t \ln m\lambda_1}}{X}\right) + a(m)\frac{\sqrt{D_x t}}{X} + b(m)\frac{\sqrt{D_x t}}{X}\ln\frac{\sqrt{D_x t}}{X}$$

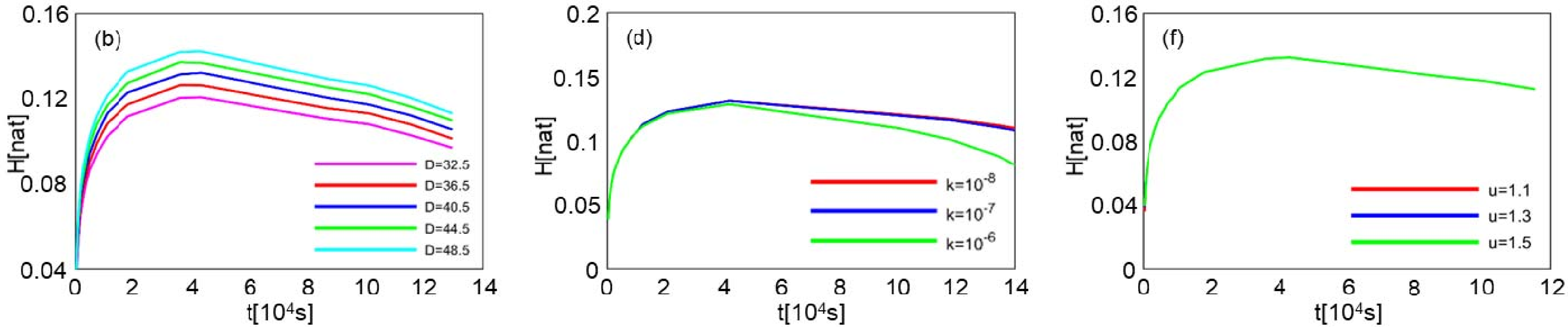


Temporal changes of entropy under different dispersion coefficient(b), degradation coefficient(d) and velocity (f)

3.3 Local sensitivity analysis

$$H = -\left(1 - \frac{4\sqrt{-D_x t \ln m\lambda_1}}{X}\right) \ln\left(1 - \frac{4\sqrt{-D_x t \ln m\lambda_1}}{X}\right) + a(m)\frac{\sqrt{D_x t}}{X} + b(m)\frac{\sqrt{D_x t}}{X} \ln\frac{\sqrt{D_x t}}{X}$$

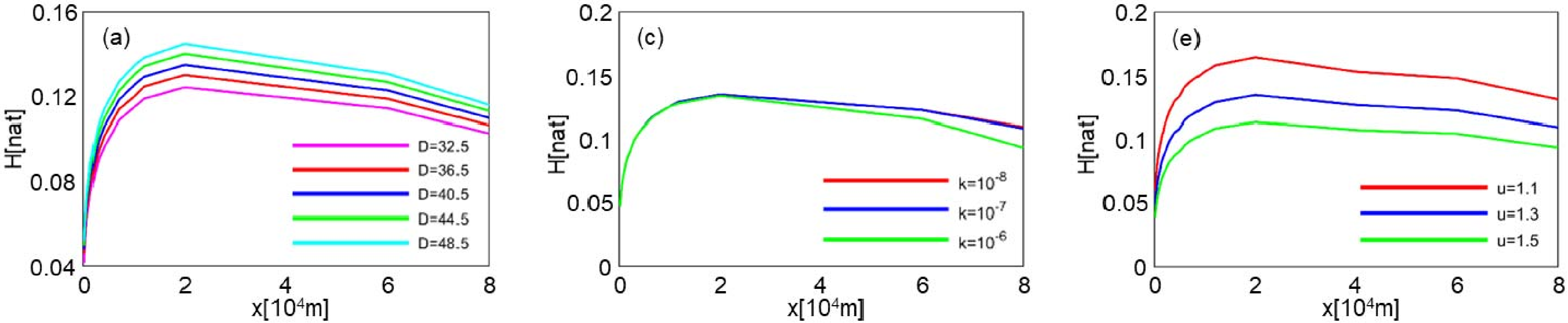
- H varies proportionally with Dx. Specifically, when D_x increases or decreases by 10%, H increases or decreases by 4%. Moreover, the change ratio of information entropy value H to parameter D_x is relatively constant, which is not obvious with the change of monitoring time t.
- The changes of k affects m, and with the change of m, the coefficient a(m) and b(m) will also change, so the influence of curve shape is greater than that of D_x. As m is affected by the monitoring time t at the same time, the greater t is, the more sensitive the information entropy H is to the change of parameter k.
- u_x does not appear in the information entropy equation. So the change of u_x has no effect on the information entropy, and the value of information entropy is insensitive to the change of u_x.



Temporal changes of entropy under different dispersion coefficient(b), degradation coefficient(d) and velocity (f)

3.3 Local sensitivity analysis

For the spatial change of information entropy, the sensitivity of information entropy H to diffusion coefficient D_x and attenuation coefficient k is the same as that of time change dimension of information entropy. The only difference is that the migration velocity u_x will affect the spatial change of information entropy. Because the change of migration velocity means that the time required to pass through the same distance will also change, and the diffusion and attenuation effects will increase with the increase of diffusion time, so the spatial change of information entropy will increase with the decrease of migration velocity u_x, and the larger the migration distance, the more sensitive the information entropy is to u_x.

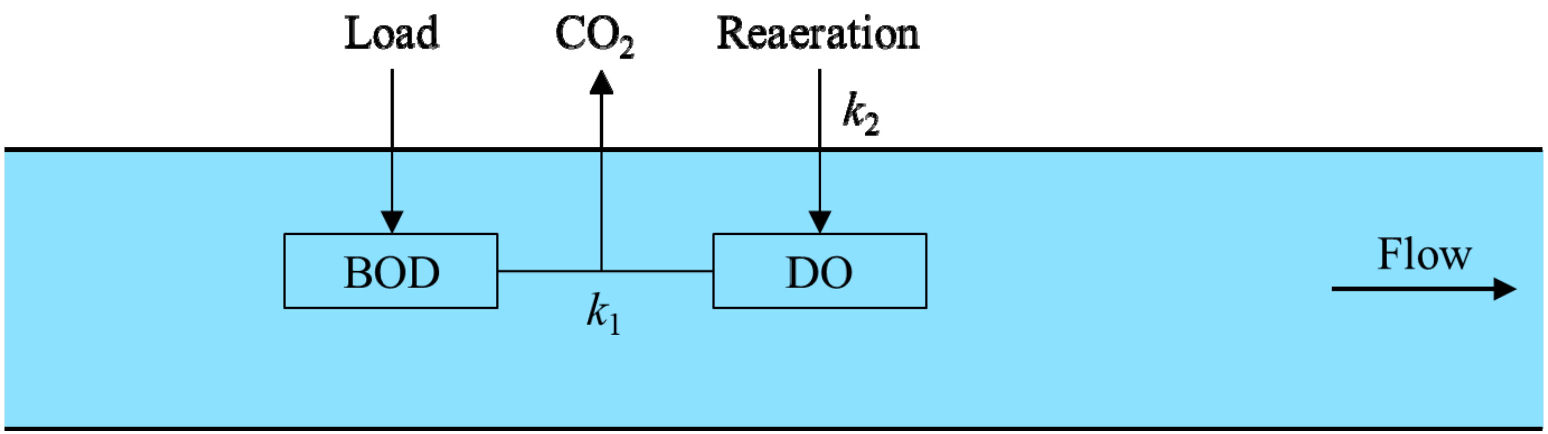


Temporal changes of entropy under different dispersion coefficient(a), degradation coefficient(c) and velocity (e)

4. Spatial and temporal changes of entropy in complex water quality process

BOD-DO migration and transformation process

- When the BOD pollution load enters the river, besides diffusion and convection, it will be degraded to CO_2 by dissolved oxygen in the water, resulting in the decrease of BOD and DO concentration in the river. At the same time, the oxygen in the atmosphere will dissolve again into the river water to restore DO to the level before the pollution enters the river. This process is called reaeration.



4. Spatial and temporal changes of entropy in complex water quality process

BOD-DO process: Study area

- The research background of this example is the Guangming Section of Maozhou River in Shenzhen (from Guangming WWTP to Yanchuan), with dispersion coefficient D_x=22.5m^2/s, average flow velocity u_x=0.2m/s, average river area A=30m^2, instantaneous BOD discharge M=10000kg. Taking the space step of one-dimensional river solute transport model as x=50m, 200 river sections were divided, the total length X=10km; the time step was taken as t=120s, and 400 time nodes, the total time T=48000s=13.33h. The degradation coefficient of BOD is k_1=0.18d^{-1}, the reaeration coefficient k_2=2d^{-1}, and the saturated dissolved oxygen concentration C_s=7mg/L.

Map of the study area

4. Spatial and temporal changes of entropy in complex water quality process

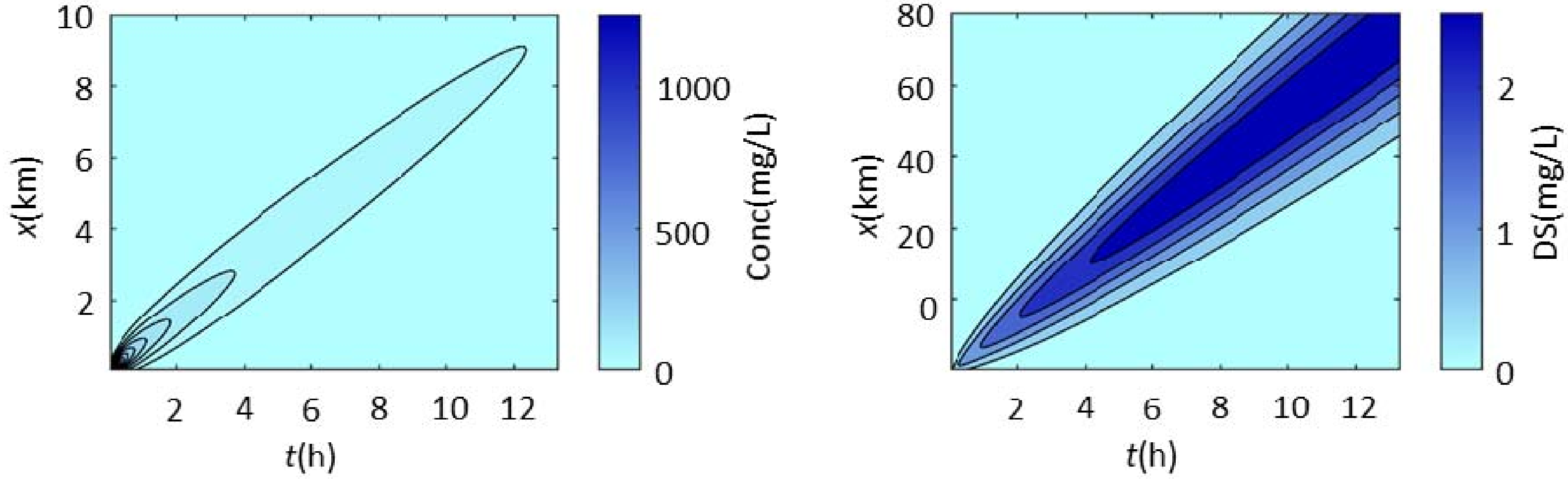
BOD-DO process: Information entropy calculation

- In order to facilitate the division of concentration intervals in the calculation of information entropy, the relationship

$$DS = C_s - DO$$

  is used to convert DO into DS (oxygen deficit) value.

- The calculated BOD and DS concentration values are plotted as contours, as shown in the figure



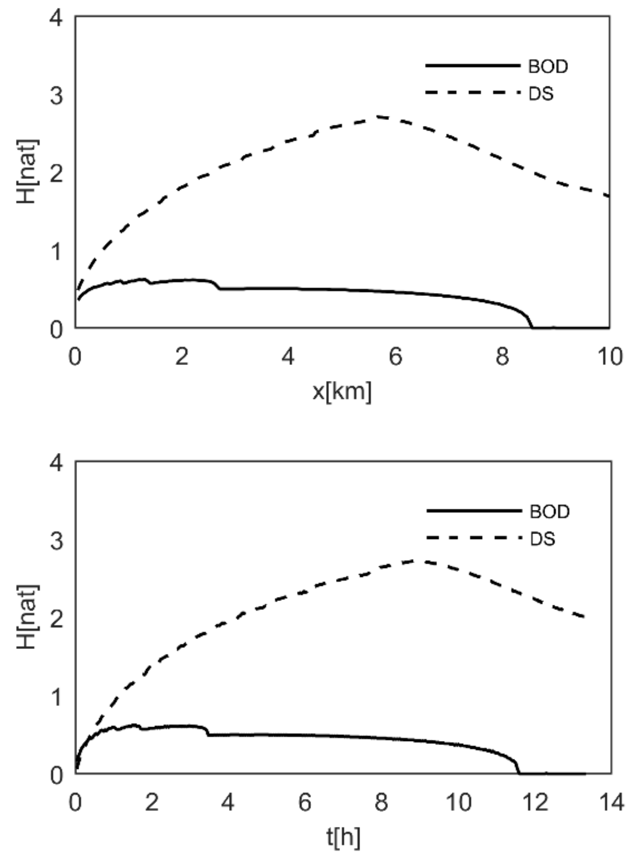
Concentrations of BOD（left）and DS（right）

4. Spatial and temporal changes of entropy in complex water quality process

Spatial and temporal changes of BOD-DO information entropy under fixed observation

- The change of BOD's information entropy with time and space is similar to that of one-component system. The maximum value of space information entropy appears at x=1.5km, the maximum value of space information entropy is 0.62nat, the maximum value of time information entropy appears at t=1.8h and the maximum value of time information entropy is 0.63nat.

- The peak values of spatial and temporal information entropy of DS occur at x=6km and t=8.8h respectively, about 60% of the observed area. The maximum values of spatial and temporal information entropy are 2.7nat and 2.8nat, respectively. Moreover, the information entropy value of DS is obviously larger than that of BOD in both time and space



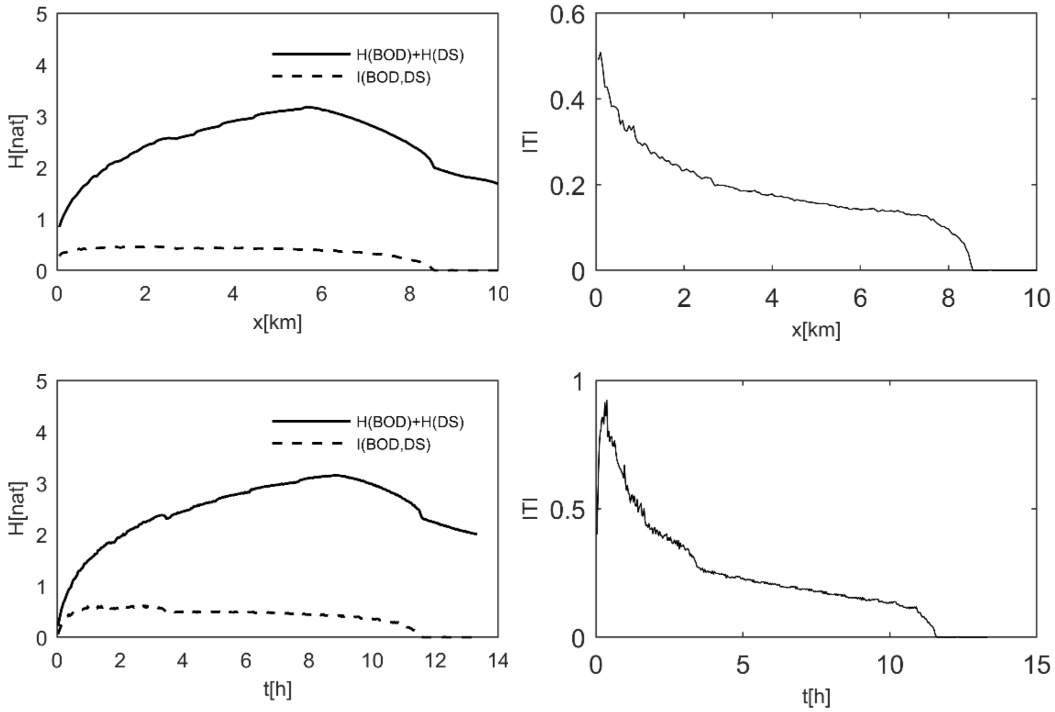
Spatial (up) and temporal (down) changes of BOD-DO information entropy

4. Spatial and temporal changes of entropy in complex water quality process

Mutual information and ITI between BOD and DO under fixed observation

- The mutual information between BOD and DS is almost the same as the information entropy of BOD, which indicates that there is a great information overlap between the two pollutant parameters. The information provided by DS almost "completely covers" the information provided by BOD. It also shows that the change of BOD is closely related to the change of DS.

- The information transfer index (ITI) decreases rapidly from the initial high to near 0.2 at any time, then slows down gradually and finally reaches 0.

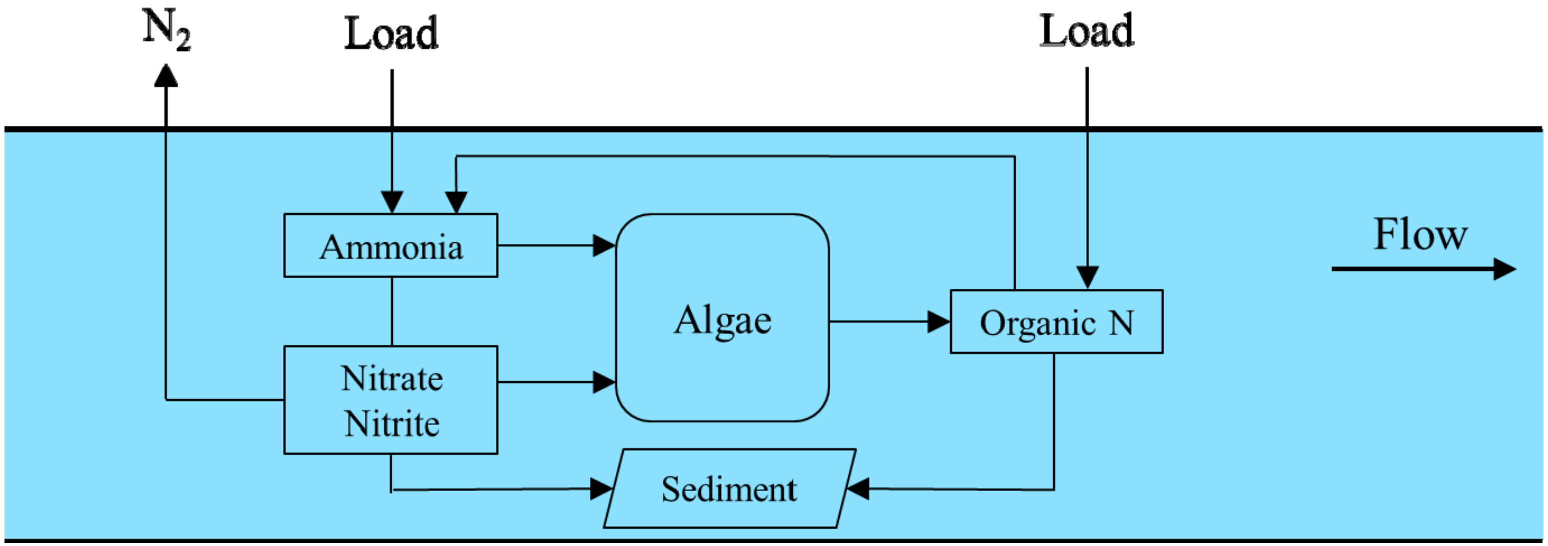


Spatial and temporal changes of mutual information and ITI between BOD and DO

4. Spatial and temporal changes of entropy in complex water quality process

Nitrogen transport and transformation process in rivers

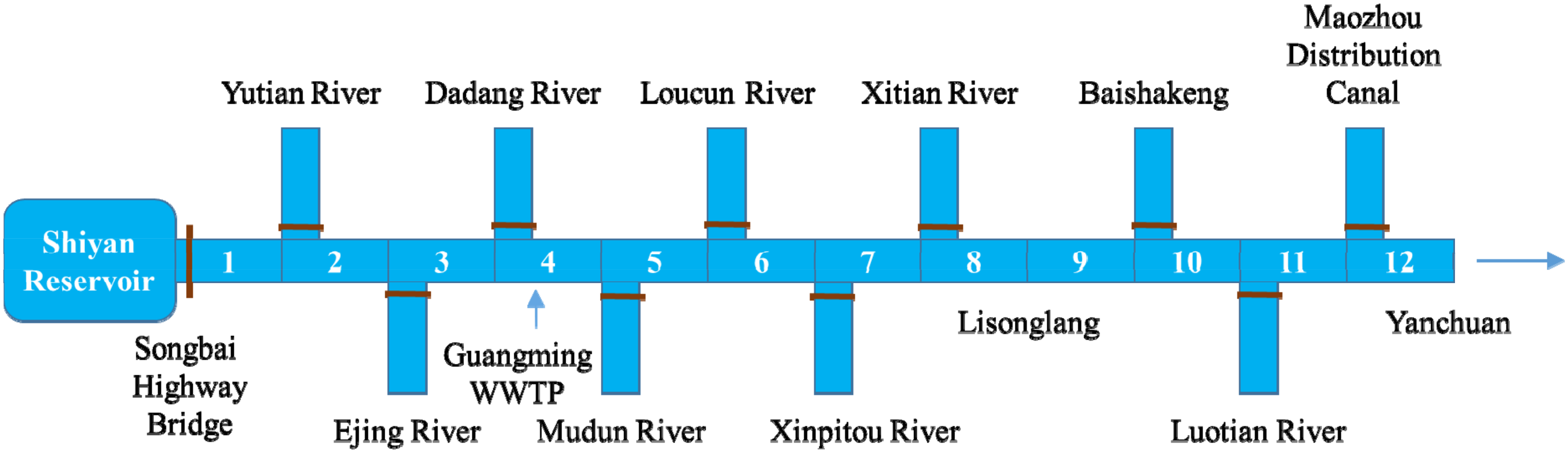
- There are three main forms of nitrogen in rivers: ammonia, nitrate and organic nitrogen. Their transport and transformation processes in rivers are shown in the figure.



4. Spatial and temporal changes of entropy in complex water quality process

Nitrogen process: Study area

- The research background of this example is the middle and upper reaches of Maozhou River in Shenzhen (from Songbai Highway Bridge to Yanchuan).
- Considering the distribution of tributaries, sewage outlets and water quality monitoring stations in the study area, the main stream of the study area is divided into 12 segments, as shown in the figure.



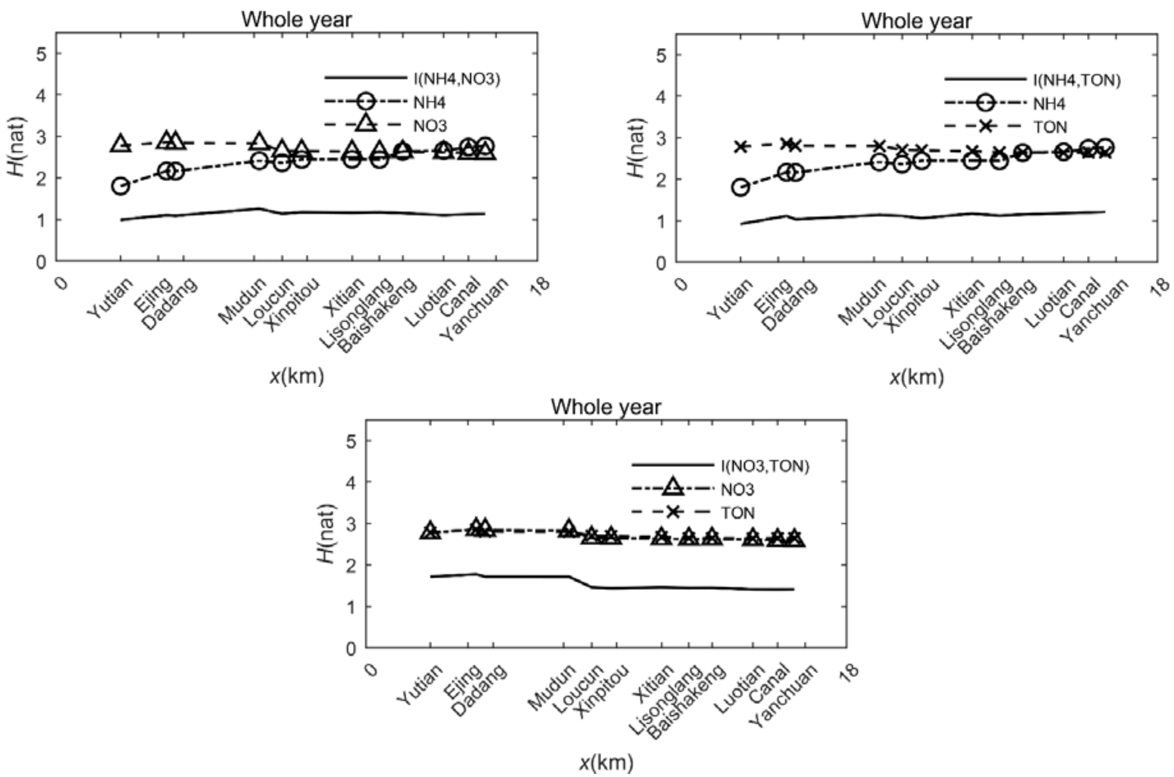
Map of the study area

4. Spatial and temporal changes of entropy in complex water quality process

Nitrogen process: Spatial changes of information entropy and mutual information

- Because the main flow of Maozhou River varies greatly between dry season and flood season, the change of information entropy of nitrogen components is discussed in three cases: whole year, flood season (April to September) and dry season (January to March, October to December).

- Because the model can only simulate nitrogen concentration data of 12 river segments every day in 2017, and there is not enough data to calculate the temporal variation of information entropy, we only discusses the spatial variation of information entropy and mutual information of three nitrogen components.



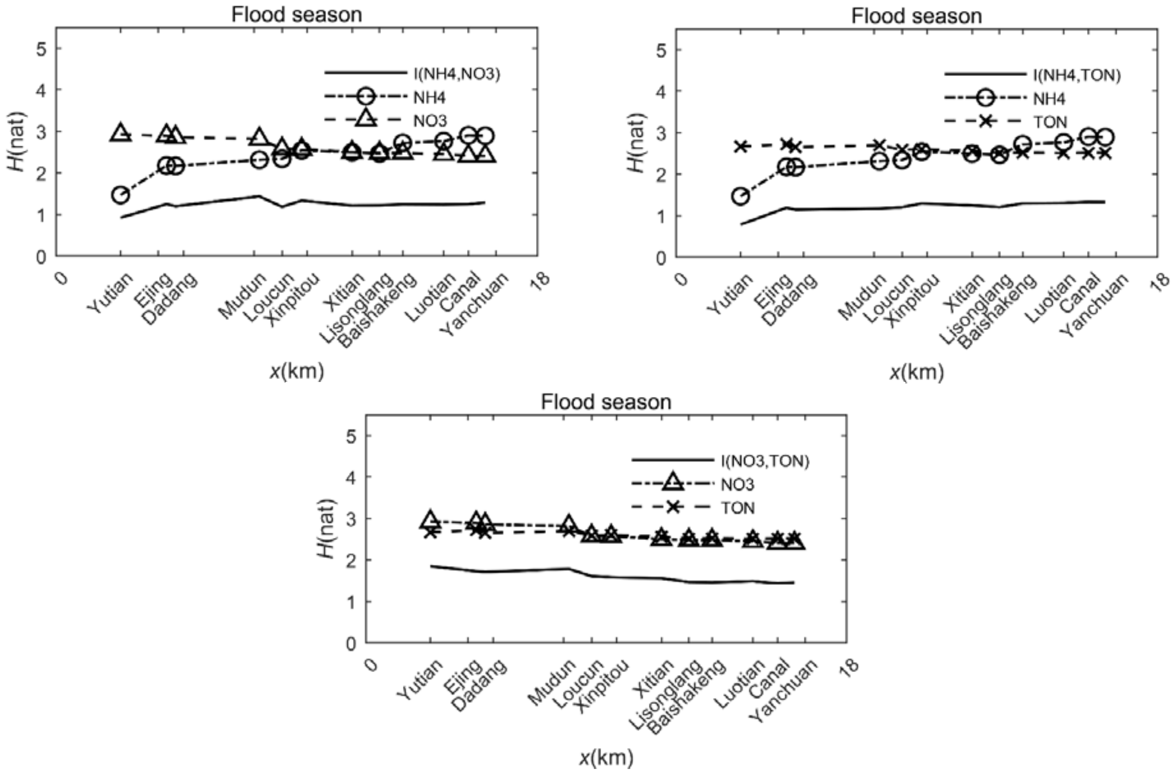
Spatial changes of information entropy and mutual information in nitrogen process: whole year （NH4-ammonia，NO3-nitrate，TON-organic N）

4. Spatial and temporal changes of entropy in complex water quality process

Nitrogen process: Spatial changes of information entropy and mutual information

- From the spatial change of information entropy of nitrogen components, the spatial change of information entropy of ammonia is the most significant. In the case of high flow rate and large input of source and sink, the change of information entropy of ammonia is very obvious, from 1.4 nat to 3 nat. The change of information entropy of nitrate and organic nitrogen is not as obvious as that of ammonia.

- The information entropy of nitrate and organic nitrogen decreased with the increase of migration distance, from the initial 3 nat to 2.5 nat. Especially in the section from Mudun River to Loucun River, because of the existence of Guangming WWTP, the flow in the upstream of the reach varies greatly, and the decrease of information entropy is also the most obvious.



Spatial changes of information entropy and mutual information in nitrogen process: flood season
（NH4-ammonia，NO3-nitrate，TON-organic N）

4. Spatial and temporal changes of entropy in complex water quality process

Nitrogen process: Spatial changes of information entropy and mutual information

- In the process of single component water quality, the most important factor affecting information entropy is convection and diffusion, because there is no pollutant transformation reaction or the effect of pollutant decay is not obvious.

- For complex water quality processes, besides convection and diffusion, the effect of pollutant transformation on information entropy is gradually emerging.

- For nitrogen process, whether ammonia, nitrate or organic nitrogen, it is difficult to observe the information entropy change characteristics similar to the one-component water quality process. It can be seen that the effect of transformation reaction in the process of nitrogen conversion has exceeded the effect of convection and diffusion, and become the dominant factor affecting the change of information entropy.

Management implication includes:

- Risk analysis on toxic pollutant transport;
- Baseline river solutes identification;
- Network design

EGU General Assembly 2020

THANKS!