

# GAUSSIAN FILTER TO PROCESS TRACER BREAKTHROUGH CURVES

## ANALIZA SLEDILNIH KRIVULJ Z GAUSOVIMI FILTRI

Guangquan LI<sup>1,\*</sup>, Hong LIU<sup>2</sup> & Simeng YANG<sup>1</sup>

### Abstract

UDC 556.34:004.93

*Guangquan Li, Hong Liu & Simeng Yangan: Gaussian filter to process tracer breakthrough curves*

Breakthrough curves in hydrogeology are similar to seismograms in containing a variety of undesired noises and regular interferences characterized with high frequency. In this paper, Gaussian filter for processing seismic waves is used to retain low-frequency trend of breakthrough curves and remove away high-frequency fluctuations. At first, the mathematical fundamental of the filter is introduced. Then the filter is applied to process four breakthrough curves measured in laboratory experiments, in which Gaussian parameter is set to be 0.2 and 0.5. Finally, a breakthrough curve in field test is processed with different Gaussian parameters. The results demonstrate how the parameter controls the cutting-off frequency and the filter is well controllable and very efficient in acquiring the primary trend of the curves.

**Key words:** Gaussian filter; convolution; breakthrough curves, cutting-off frequency, noises.

### Izveček

UDK 556.34:004.93

*Guangquan Li, Hong Liu & Simeng Yangan: Analiza sledilnih krivulj z Gaussovimi filtri*

Sledilne krivulje (krivulja časovne odvisnosti koncentracije povrnjenega sledila) v hidrogeologiji so podobno kot seizmogrami v geofiziki obremenjene z nezaželenimi visokofrekvenčnimi šumi in interferencami. V tem članku uporabimo Gaussov filter, primarno namenjen obdelavi seizmičnih podatkov, za odstanitev visokofrekvenčnih шумov iz sledilnih krivulj. Najprej predstavimo matematične osnove filtriranja, potem Gaussov filter z parametrom 0,2 in 0,5 uporabimo na štirih sledilnih krivuljah, dobljenih v laboratorijskih pogojih. Na koncu z različnimi Gaussovimi parametri obravnavamo sledilno krivuljo, dobljeno pri sledenju v naravi. Z rezultati prikažemo vpliv izbranih parametrov na mejno frekvenco ter prilagodljivost, učinkovitost in uporabnost filtra za izluščenje primarnih značilnosti sledilnih krivulj.

**Ključne besede:** Gaussov filter, konvolucija, sledilne krivulje, mejne frekvence, šum.

<sup>1</sup> Department of Geophysics, Yunnan University, 2 North Green Lake Rd., Kunming, Yunnan 650091, P. R. China, e-mails: guangquan\_li@yahoo.com, yangsimeng@mail.ynu.edu.cn

<sup>2</sup> International Joint Research Center for Karstology, Yunnan University, 5 Xueyun Rd., Kunming, Yunnan 650223, P. R. China, e-mail: hongliu@ynu.edu.cn

\* corresponding author

## INTRODUCTION

Natural or artificial contaminants such as nitrate or phthalate are a persistent threat and danger to the quality of environment and the health of people. Fast turbulent flow in karstic conduits can speed up transmitting rates and can enlarge scopes of contaminants, thus causing exceptional hazards to people who drink untreated groundwater. In the recent years, the National Institute of Environmental Health Sciences at NIH of the United States established a superfund to support such research programs in the way of multi-disciplinary. In China, there are increased preterm birthrate and cancer rate related to the heavily polluted groundwater, and more funds are invested on the relevant research.

Breakthrough curves sampled at springs (or other similar locations) are similar to seismic waves monitored at seismic stations, in acquiring time-series data at fixed locations to investigate underground processes or structures. Such data are often embedded with undesired noises and interferences. The noises are often featured with very high frequency due to their random/stochastic feature. Non-random interferences are likely to originate from the heterogeneous geological structures (at small scales) and/or undesired instrumental responses, often presenting as secondary and regular fluctuations in recorded data.

Scientists in geoscience are typically concerned with major physical/chemical mechanisms or dominant processes. Small-scale structures and minor mechanisms are often beyond consideration because they usually don't play the crucial role in producing geological phenomena. A typical way to extract data reflecting the dominant processes is effective removing away noises and interferences by filters (Blinchikoff & Zverev 2001).

While modeling is often forward, i.e., from mechanisms to mathematical formulization to prediction of observed data, filters are reverse in analyzing measured data to infer mechanisms. Regression invariably requires a mathematical formula based upon implicit assumptions. Therefore, regression is not applicable for the systems where such assumptions are not valid. In this sense, regression is a trial or test of our assumptions. In contrast, filtering is a kind of data extraction or optimization, not alternative to regression. Regression analysis could be conducted after filtering.

The base of filters is the categorization of signal into two levels in the time domain: the quick change (associated to small spatial length) and the slow change (associated to large spatial length) (Nixon & Aguado 2008; Yu *et*

*al.* 2016). In pattern recognition and feature extraction, low-pass filters (which only allow the passing of low frequency components, but do not allow the passing of high frequencies) such as Gaussian filter are appropriate for the recognition of pattern as a whole. Wavelets analysis (e.g., Erlebacher *et al.* 1996) is essentially a moving-window technique designed for the extraction of local features in time or space.

A primitive filtering technology is a simple superposition of many similar datasets, such as those measured in repeated laboratory experiments under controlled/consistent conditions. However, that way often is not applicable to breakthrough curves because such datasets acquired arduously on field are often very limited in quantity and vary appreciably in time and space.

Moving average filters or Savitzky-Golay filters (Savitzky & Golay 1964) are often used to perform data averaging, remove noises, clean up signals and discern major patterns. This type of filters is also called least squares (Hamming 1983) or digital smoothing polynomial (Ziegler 1981). The principle is replacing the original value at a point with a weighted average on those values at the point and its neighboring points. Those methods preserve the basic pattern of the curves, but often drastically reduce peak amplitude.

Gaussian filter is designed for filtering away normally distributed noises (Ross 1996) and smoothening measured signals. However, some interferences are not noises, but regular and arise from secondary mechanisms or instrumental erroneous responses. Gaussian filter is capable of partially suppressing such regular interferences. The key parameter in Gaussian filter is Gaussian parameter. If the parameter is infinite large, the filter won't have any filtering effects. If the parameter is zero, the signals after the filter will vanish (becomes the average everywhere). As such, Gaussian parameter ranges from 0 to  $+\infty$ .

Gaussian filter has been widely used in seismology and in this paper, we are motivated to apply the filter to processing tracer breakthrough curves. The target is investigating how Gaussian parameter affects the filtering effect by controlling the cutting-off frequency. At first, breakthrough curves of artificial tracer, i.e., sodium chloride in laboratory experiments by Li (2004) is processed. Then a breakthrough curve of Uranine measured in a field test is processed to see the effect of Gaussian parameter.

METHODOLOGY

There are two approaches to apply Gaussian filter to signal processing. One is directly in the time domain involving convolution (Fig. 1a), while the other is in the frequency domain (Fig. 1b). The latter requires Fourier Transform to convert temporal signals to their spectra before the filter, and the inverse Fourier Transform to convert the spectra to signals after the filter. This approach is used in this study.

The low-pass Gaussian filter has the mathematical expression in the frequency domain as follows (Langston 1977, 1979).

$$G(\omega) = \exp\left[-\left(\frac{\omega}{2\alpha}\right)^2\right] \tag{1}$$

where  $\omega$  is dimensionless angular frequency and  $\alpha$  is Gaussian parameter. Fig. 2 illustrates what Gaussian filter looks like and according to Equation (1), parameter  $\alpha$  is a dilatational factor of  $\omega$  which controls the cutting-off frequency. For this reason, the three curves in Fig. 2 have similarity between each other, so that there is no concern with the curves with  $\omega > 1$ . For  $\alpha = 1$ ,

we have  $G(\omega > 0.4) < 0.2$  and the cutting-off frequency is  $\omega > 0.4$  in this case. We shall further demonstrate the cutting-off frequency with a specific example at the end of Discussion Section.

Using the inverse Fourier Transform (Oppenheim & Schaffer 2009), it is not difficult to get the following Gaussian function in the time domain from the above filter in the frequency domain.

$$g(t) = \sqrt{2}\alpha \exp\left[-(\alpha t)^2\right] \tag{2}$$

The difference of Gaussian filter from Savitzky-Golay filters (Savitzky & Golay 1964) mentioned in Introduction Section may be explored by Fig. 1a, in which the convolution means the weighted average over the whole (original) signal sequence. In contrast, a Savitzky-Golay filter usually means a weighted average over the neighborhood of the point under processing.

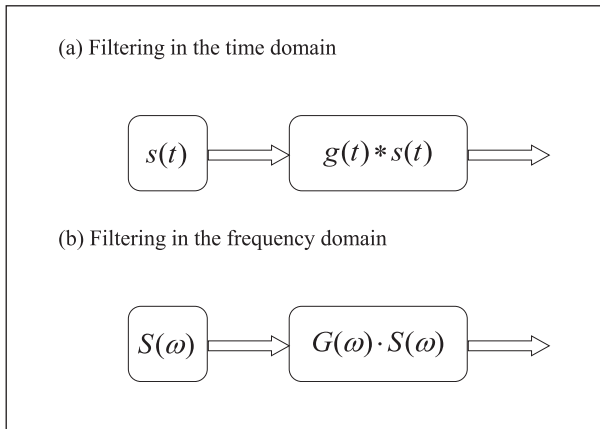


Fig. 1: Gaussian filter in the time domain  $g(t)$  (a) and in the frequency domain  $G(\omega)$  (b).

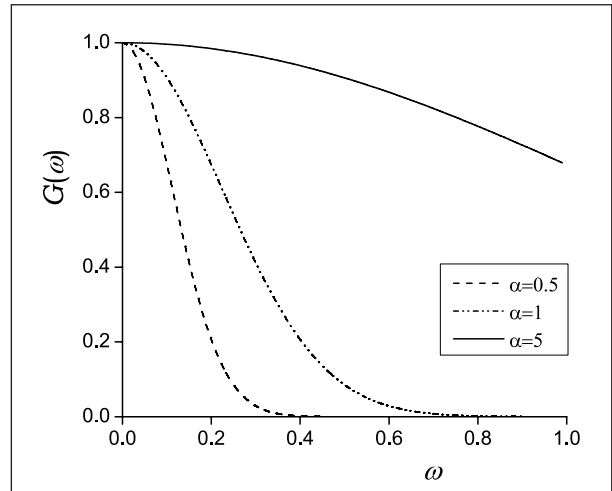


Fig. 2: Gaussian filter in the frequency domain, with different values of Gaussian parameter  $\alpha$ . Note that  $\alpha$  is a dilatational factor of dimensionless angular frequency ( $\omega$ ) which controls the cutting-off frequency.

EXAMPLES

Tracer breakthrough curves are a prevailing way to acquire flow and transport parameters and to predict fate of soluble contaminants. Using sodium chloride as the tracer, Li (2004) conducted a series of laboratory experiments to simulate the interaction of solute between karst conduit and its surrounding matrix. As the first example,

we use Gaussian filter to process breakthrough curves measured in a set of the experiments.

In the experiments (Li *et al.* 2008), the conduit-matrix system was flushed/cleaned by distilled water (Fig. 3a). Then salty water with step-like function passed through the conduit, when part of it is sequestered into

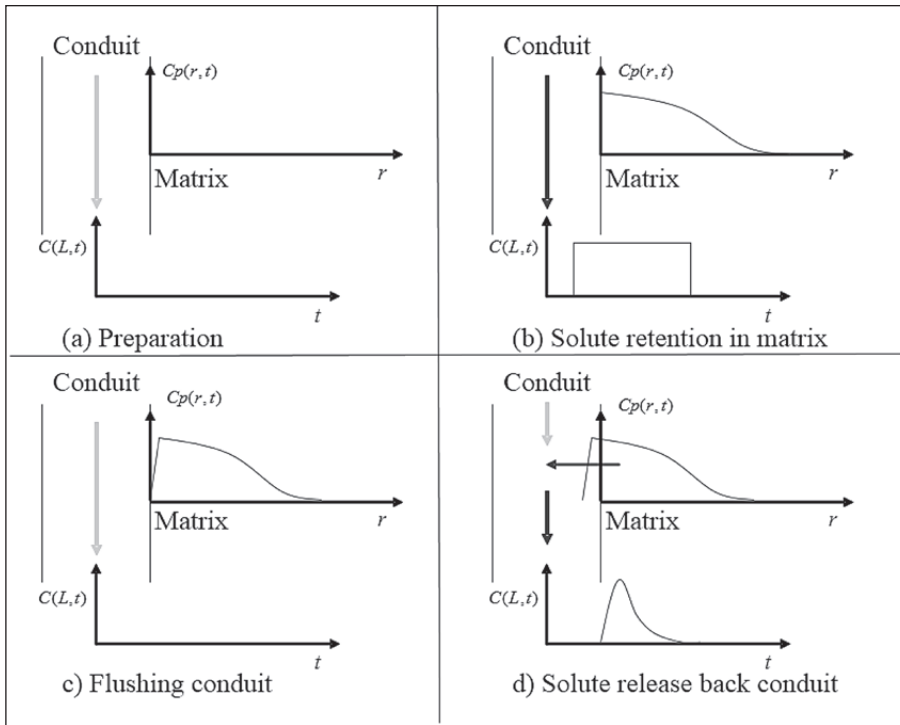


Fig. 3: Schematic solute distribution in the porous medium,  $C_p(r, t)$ , and the breakthrough curves at the conduit exit,  $C(L, t)$  in the laboratory experiments: (a) the conduit and medium are initially clean; (b) solute-laden flow passes through the conduit, and part of the solute is sequestered into the medium; (c) the conduit is quickly cleaned with fresh water, while solute in the medium remains inside; (d) the remaining solute is released from the medium, diluted and transported to the exit.

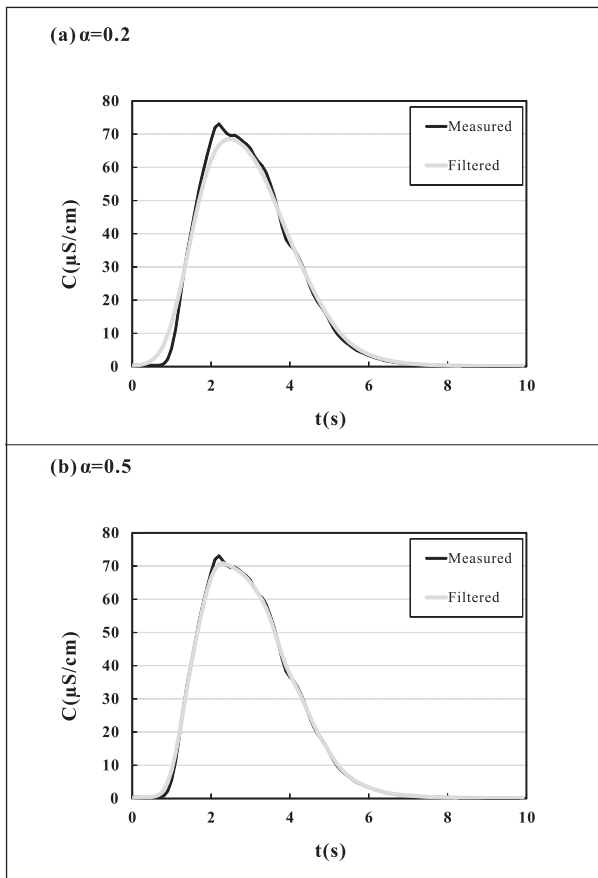


Fig. 4: Gaussian filtering on breakthrough curve (the volume of salty water into the matrix  $V_{in}$  at  $32.1 \text{ cm}^3$ ).

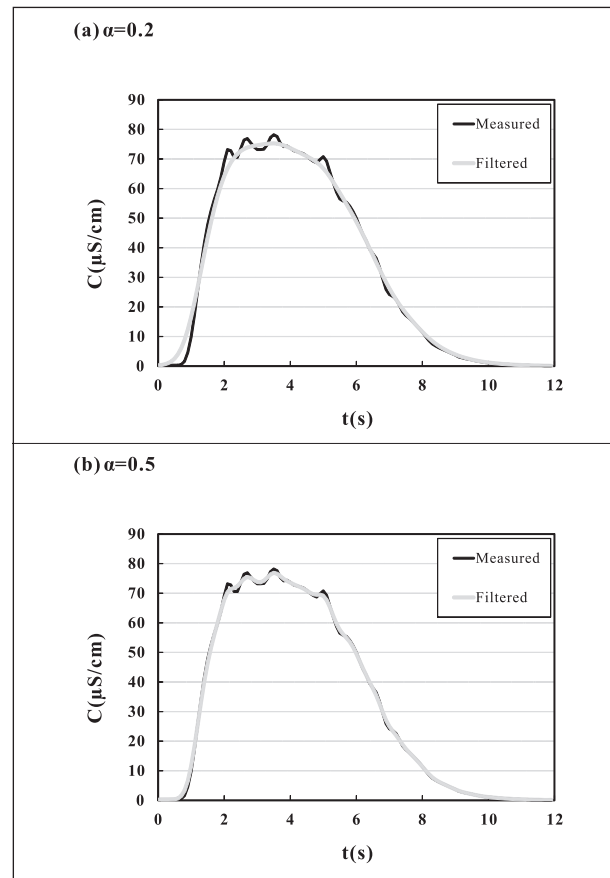


Fig. 5: Gaussian filtering on breakthrough curve (the volume of salty water into the matrix  $V_{in}$  at  $63.2 \text{ cm}^3$ ).

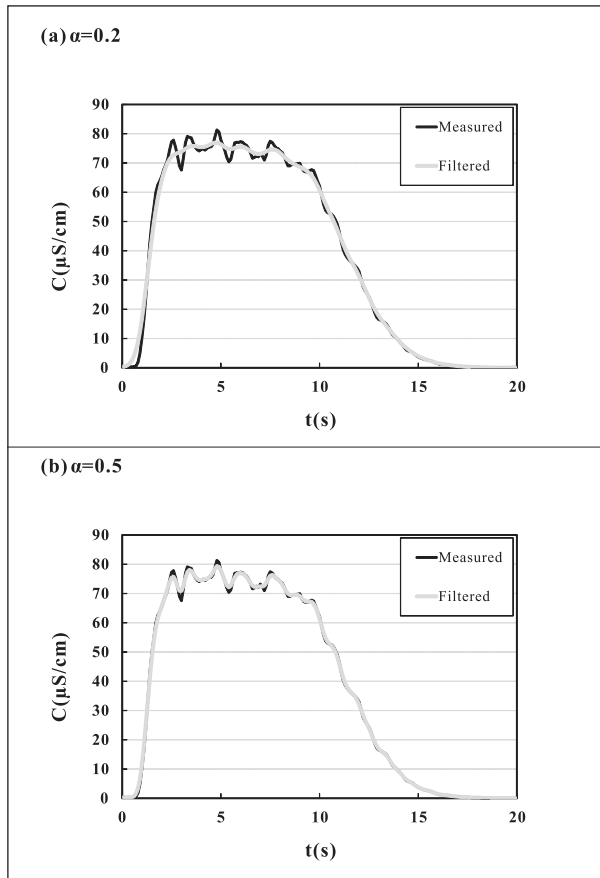


Fig. 6: Gaussian filtering on breakthrough curve (the volume of salty water into the matrix  $V_{in}$  at  $127 \text{ cm}^3$ ).

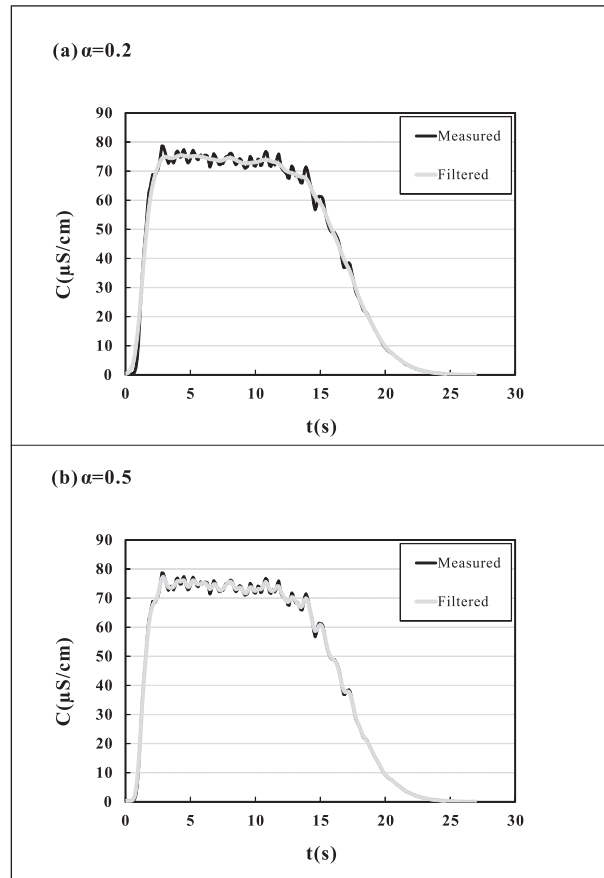


Fig. 7: Gaussian filtering on breakthrough curve (the volume of salty water into the matrix  $V_{in}$  at  $191.8 \text{ cm}^3$ ).

the surrounding matrix (Fig. 3b). Later the conduit is flushed by distilled water again to separate salty water in the conduit from that in the matrix (Fig. 3c). Finally, salt in the matrix is released back into the conduit, yielding breakthrough curve at the conduit exit that was measured by an electro-conductivity meter (Fig. 3d). These are the procedures of the experiments, and detail was documented carefully in Li (2004).

We use the results (i.e., four breakthrough curves) from a set of experiments in which conduit flow had a Reynolds number of 1278 (a higher number indicating a more turbulent and mixing flow), when a volume of salty water ( $V_{in}$ ) was injected from the conduit into the surrounding matrix. Later, a lateral water flux released salt from the matrix back into the conduit which mixed with distilled water from the conduit entrance. Finally the diluted salty water was recorded by the electro-conductivity meter at the conduit exit. The four breakthrough curves are plotted in Figs. 4–7. Please note that the volume of salty water ( $V_{in}$ ) from the conduit into the matrix controlled the thickness of salty water into the matrix and thus the lasting period of the breakthrough curve. In

the experiments, salty water from the conduit into matrix had a concentration of  $0.260 \text{ kg/m}^3$ , corresponding to an electro-conductivity of  $0.520 \text{ mS/cm}$ . Accordingly, the breakthrough curves in term of electro-conductivity can be converted to those of concentration linearly because for such a very dilute solution, electro-conductivity is almost proportional to concentration.

From Equation (1), Gaussian parameter  $\alpha$  is the only parameter of Gaussian filter that controls filtering. For each breakthrough curve,  $\alpha$  is set to be 0.2 and 0.5 to see the filtering effect on the curve. The filtered breakthrough curves are plotted in Figs. 4–7.

The high frequency fluctuations at the peak amplitude are not desirable, as they are interferences produced by small holes on the copper pipe in the experiments (Li *et al.* 2008). With  $\alpha = 0.5$ , those fluctuations are effectively suppressed and the filter has no effect on the long tailings. However, the initial arrival is artificially shifted forward due to the earlier arrival of the tracer. The filter with  $\alpha = 0.5$  has no such artificial effect on the curves, but the suppressing effect on those undesired high fluctuations is not as good as the filter with  $\alpha = 0.2$ .

Concentration fluctuations in laboratory arise from apparatus design, measurement technique or instrumental setup, having nothing to do with the flow processes on the field. Moreover, field tests have time scale much longer than laboratory experiments, due to the large spatial scale. The former is more realistic.

As our second example, a breakthrough curve of Uranine measured in a karst aquifer in the United States is processed. Fig. 8 depicts how Gaussian parameter  $\alpha$

controls the filter's effect in acquiring the low frequency trend. Compared with the above four breakthrough curves measured at laboratory, the high frequency fluctuations in this case are much denser and more severe. With  $\alpha$  exceeding 1, the filter has only small filtering effect. With  $\alpha < 0.1$ , the high frequency fluctuations have been effectively removed away and the primary trend becomes evident.

## DISCUSSION

Gaussian filter is widely used in processing seismograms in which the recorded seismic waves are prone to noises and/or interferences induced by the complex structure of crust and the instrumental response of seismometers. Receiver functions are a frontier technology, in which

raw seismograms are transformed to help identification of seismic phases (Langston 1977, 1979). It is well admitted that Gaussian filter with  $\alpha$  of 2 is capable of effectively filtering away undesired high-frequency components in receiver functions.

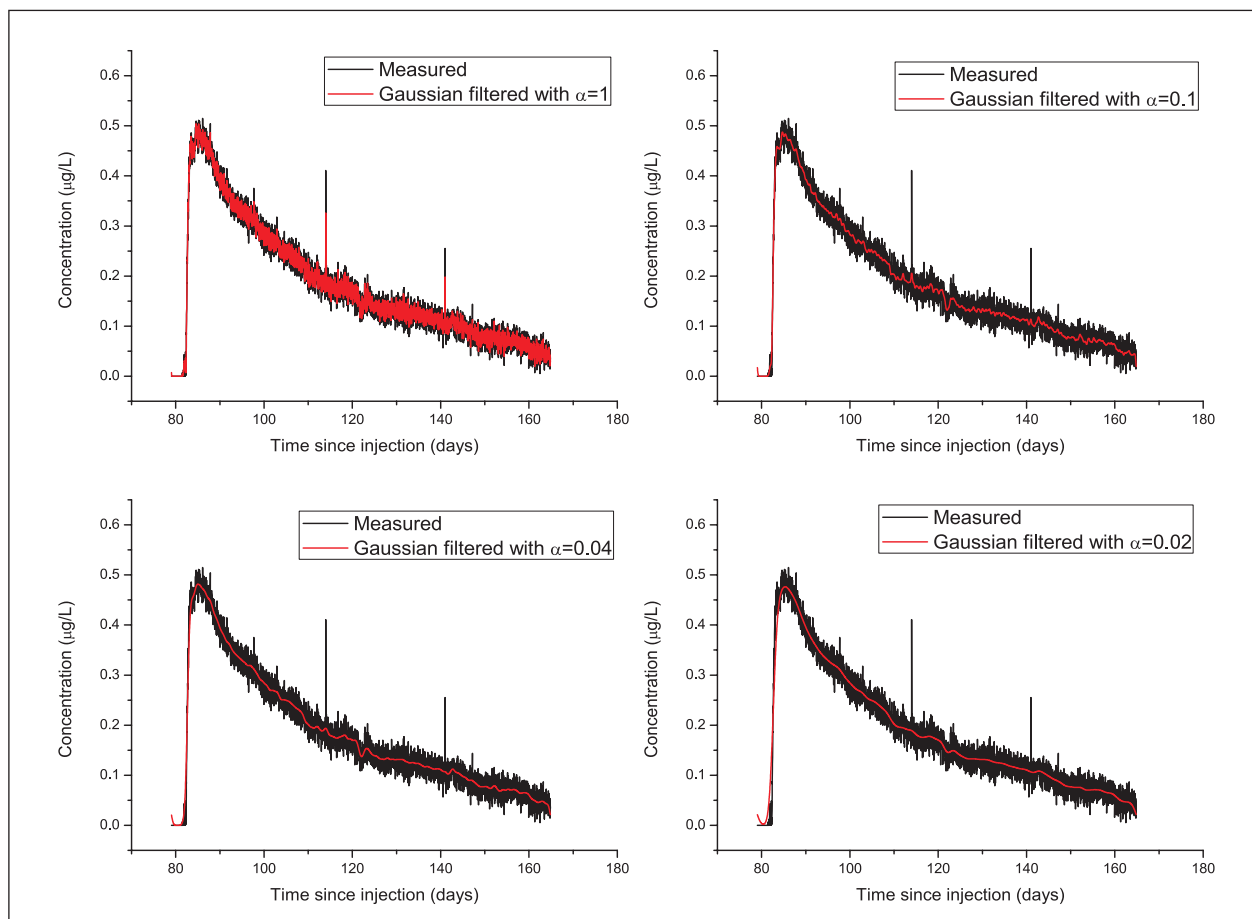


Fig. 8: Gaussian filtering on an Uranine breakthrough curve measured by field test. Note that the small (red) falling-down at 80 days is caused by singularity (the discontinuity circlely between the end of the measured signal at 164 days and the beginning of the signal at 80 days). This is because Fourier Transform always yields signal (that is continuous circlely between its end and its beginning), but the measured data is discontinuous circlely between its end and its beginning.

Seismograms have both positive and negative values due to the oscillations of the pendulum inside seismometer. In contrast, contaminant breakthrough curves can only have zero or positive values because tracer concentration cannot be negative. Interestingly, our results show that the filtered curves have no any negative values appearing. This physical reasonability is an advantage of Gaussian filter.

In our laboratory experiments, the small fluctuations on the breakthrough curves arose from the heterogeneity of the pipe (with small holes) to simulate karst conduit (Li *et al.* 2008), because the scale of glass beads (diameter at 0.256 mm on average) simulating rock matrix is too small to generate those fluctuations.

Whether in laboratory experiments or field tests, breakthrough curves often present a fast rising limb followed by a much slower falling limb (Fig. 8). The major cause is a variety of retention and retarding mechanisms on solute transport such as vortex at conduit enlargements (Hauns *et al.* 2001), small-scale convection in the conduit subsurface and the associated two-ways exchange of water between conduit and matrix (Thibodeaux & Boyle 1987; Elliot & Brooks 1997), reversible seepage between conduit and matrix and temporarily solute staying in pores or fissures of the ambient rock (Martin & Dean 2001; Screaton *et al.* 2004; Birk *et al.* 2006; Li *et al.* 2008; Li & Field 2016), or passive solute exchange between mobile- and immobile-water regions like 2RNE (Toride *et al.* 1993; Field & Pinsky 2000).

Nonetheless, even a solute plume with spatial symmetry will also result in a (temporal) breakthrough curves with fast rising and slower falling at down streams; see the Green's function for the standard advection-dispersion problem (Abramowitz & Stegun 1970; Kreft & Zuber 1978). The physical reason is that when the plume

front reaches the sampling point, it has a relative sharp (rising) gradient. In contrast, when the plume back arrives at the same point, a certain time has passed and the plume the moment will have a smaller spatial gradient due to the spreading action of hydrodynamic dispersion (Taylor 1954; Bear 1972).

Recognition of major patterns and discern of dominant processes are always an interesting topic in science and even art. This paper is an endeavour of applying Gaussian filter to karst hydrogeology. Nevertheless, it helps pattern recognition, not meaning that the filtered curves can take place of raw data.

At last, in our program we use dimensionless frequency, namely,  $\frac{f}{f_h} = \frac{i}{N}$ , wherein  $f$  is dimensional frequency,  $f_h = \frac{1}{dT}$  ( $dT$  is the sampling interval in the time domain),  $i$  and  $N$  are nodal numbers and the total number of nodes, respectively. According to Fig. 2, if  $G(w) < 0.2$  is the filtering criteria, then the cutting-off frequency ( $f_c$ ) will be as follows.

$$\frac{f_c}{f_h} > 0.4\alpha. \quad (3)$$

For example, in Fig. 6a,  $\alpha = 0.2$  and  $f_h = 10$  Hz, so that the filter removes away frequencies higher than 0.8 Hz, according to Equation (3). Because the high frequency fluctuations have frequency approximately at 1 Hz, they have been removed away in that figure. In Fig. 6b,  $\alpha = 0.5$  and  $f_h = 10$  Hz, so that the filter removes away frequencies higher than 2 Hz. Because the high frequency fluctuations have frequency approximately at 1 Hz, they are still present in Fig. 6b.

## CONCLUSIONS

Signal processing is important because signals monitored by instruments are often a combination and mixture of innate information, noises and interferences. It is often necessary to remove undesired noises and regular interferences away and retain the innate information to clearly reveal the dominant physical/chemical processes that produce phenomena.

Gaussian filter is an advanced tool toward the target, and its filtering effect is controlled by Gaussian parameter alone. According to the upper/temporal equation in Fig. 1, the weighting kernel is Gaussian

function in Equation (2). In contrast, the weighting kernel of the moving-average filters is a series of simple constants which are not as flexible as Gaussian filter. More important, as demonstrated at the end of Discussion Section, Gaussian parameter well controls the cutting-off frequency. Along this line, Equation (3) provides an estimate of the optimal Gaussian parameter. From the sampling interval in the time domain and the frequency of high frequency fluctuations to be removed away, Gaussian parameter can be well ascertained via the equation. In this regard, Gaussian fil-

ter is well controllable and very efficient in removing away the high frequency fluctuations, being advantageous over the average filters in which the relationship between their weighting factors and frequency is unclear.

At last, splines method may yield negative/unrealistic values of tracer concentration. In contrast, the tracer concentration curves after Gaussian filter keep positive and no any negative values appear.

## ACKNOWLEDGMENTS

This research was sponsored by National Science Foundation of China under grants 41571010 and 41371040. The authors are deeply grateful to Editor-in-Chief, Franci

Gabrovšek, whose sharp comments and insightful suggestions tremendously helped us improve the paper.

## REFERENCES

- Abramowitz, M. & I.A. Stegun, 1970: *Handbook of Mathematical Functions*.- 9-Revised edition. Dover Publications, pp. 1046, New York, Dover.
- Bear, J., 1972: *Dynamics of Fluids in Porous Medium*.- American Elsevier Publishing Company, pp. 764, New York, Dover,
- Birk, S., Liedl, R. & M. Sauter, 2006: Karst spring responses examined by process-based modeling.- *Ground Water*, 44, 6, 832–836, doi:10.1111/j.1745-6584.2006.00175.x.
- Blinchikoff, H.J. & A.I. Zverev, 2001: *Filtering in the Time and Frequency Domains*.- SciTech Publishing, pp. 520.
- Elliot, A. H. & N.H. Brooks, 1997: Transfer of nonsorbing solutes to a streambed with bed forms: Laboratory experiments.- *Water Resources Research*, 33, 137–151.
- Erlebacher, G., Hussaini, M.Y. & L.M. Jameson, 1996: *Wavelets: Theory and Applications*.- Oxford University Press., pp. 528.
- Field, M. & P. Pinsky, 2000: A two-region nonequilibrium model for solute transport in solution conduits in karstic aquifers.- *Journal of Contaminant Hydrology*, 44, 3, 329–351. DOI: 10.1016/S0169-7722(00)00099-1
- Hauns, M., Jeannin, P.-Y. & O. Atteia, 2001: Dispersion, retardation and scale effect in tracer breakthrough curves in karst conduits.- *Journal of Hydrology*, 241, 3-4, 177–193. DOI: 10.1016/S0022-1694(00)00366-8
- Hamming, R.W., 1983: *Digital Filters*.- 2nd edition, Prentice-Hall Press, pp. 257, Michigan.
- Kreft, A. & A. Zuber, 1978: Physical meaning of dispersion equation and its solution for different initial and boundary conditions.- *Chemical Engineering Science*, 33, 11, 1471–1480.
- Langston, C.A., 1977: Corvallis, Oregon, Crustal and upper mantle structure from teleseismic P and S waves.- *Bulletin of Seismological Society of America*, 67, 713–724.
- Langston, C.A., 1979: Structure under Mount Rainier, Washington, inferred from teleseismic body wave.- *Journal of Geophysical Research*, 84, 4749–4762.
- Li, G., 2004: *Laboratory Simulation of Solute Transport and Retention in a Karst Aquifer*.- PhD thesis. Florida State University, Tallahassee, pp. 194.
- Li, G., Loper, D.E. & R. Kung, 2008: Contaminant sequestration in karstic aquifers: Experiments and quantification.- *Water Resources Research*, 44, W02429, DOI: 10.1029/2006WR005797.
- Li, G. & M.S. Field, 2016: Solute migration from the aquifer matrix into a karst conduit and the reverse.- *Ground Water*, 54, 5, 699–708. DOI: 10.1111/gwat.12416
- Martin, J.B. & R.A. Dean, 2001: Exchange of water between conduits and matrix in the Floridan Aquifer.- *Chemical Geology*, 179, 145–166, DOI: 10.1016/S0009-2541(01)00320-5.
- Nixon, M.S. & A.S. Aguado, 2008: *Feature Extraction and Image Processing*.- Academic Press, pp. 424.
- Oppenheim, A.V. & R.W. Schaffer, 2009: *Discrete-Time Signal Processing*.- 3rd edition, Pearson, pp. 1144.
- Ross, S.M., 1996: *Stochastic Processes*.- 2nd edition, John Wiley and Sons, 528 pp.
- Savitzky, A. & M. Golay, 1964: Smoothing and differentiation of data by simplified least squares procedure.



- dures.- Analytical Chemistry, 36, 1627–1639, DOI: 10.1021/ac60214a047.
- Screaton, E., Martin, J.B., Ginn, B. & L. Smith, 2004: Conduit properties and karstification in the Santa Fe river sink-rise system of the Floridan aquifer.- Ground Water, 42, 338–346, DOI: 10.1111/j.1745-6584.2004.tb02682.x.
- Taylor, G.I., 1954: The dispersion of matter in turbulent flow through a pipe.- Proceedings of Royal Society of London, Ser. A, 223, 446–468. DOI: 10.1098/rspa.1954.0130
- Thibodeaux, L.J. & J.D Boyle, 1987: Bedform-generated convective transport in bottom sediment.- Nature, 325, 341–343.
- Toride, N., Leij, F.J. & M.T. van Genuchten, 1993: A comprehensive set of analytical solutions for nonequilibrium solute transport with first-order decay and zero-order production.- Water Resources Research, 29, 7, 2167–2182. DOI: 10.1029/93WR00496
- Yu, X., Ghasemizadeh, R., Padilla, I., Kaeli D. & A. Alshwabkeh, 2016: Patterns of temporal scaling of groundwater level fluctuation.- Journal of Hydrology, 536. DOI: 10.1016/j.jhydrol.2016.03.018.
- Ziegler, H., 1981: Properties of digital smoothing polynomial (DISPO) filters.- Applied Spectroscopy, 35, 1, 88–92.