**R\_TIDE V1.0 Matlab Toolbox Tutorial**

Huayang Cai1

1Institute of Estuarine and Coastal Research, School of Ocean Engineering and Technology, Sun Yat-sen University, Guangzhou, 510275, China

Correspondence: Huayang Cai ([caihy7@mail.sysu.edu.cn](mailto:caihy7@mail.sysu.edu.cn))

2022/09/21

A data-driven model to quantify the impact of river discharge on tide-river dynamics in river deltas.

Version 1.0 - September 2022

Provided by Huayang Cai

Institute of Estuarine and Coastal Research, School of Ocean Engineering and Technology, Sun Yat-sen University

Email contacts: caihy7@mail.sysu.edu.cn

How to cite:

- Huayang Cai, Bo Li, Erwan Garel, Tongtiegang Zhao, Feng Liu, and Suying Ou (2022), A data-driven model to quantify the impact of river discharge on tide-river dynamics in the Yangtze River estuary, Journal of Hydrology, submitted

**In this newly proposed version, we:**

**1) proposed a new definition of critical river discharge *Q*c defined as the value that leads to an apparent shift of tidal phases by approximately 100~180° with the increase of river discharge.**

**2) added an error estimation model.**

How to use R\_TIDE

1. **Download and install R\_TIDE toolbox**

**Users can download the latest R\_TIDE toolbox from Github:**

[**https://github.com/Huayangcai/R\_TIDE-V1.0-Matlab-Toolbox.git**](https://github.com/Huayangcai/R_TIDE-Matlab-Toolbox.git)

1. **R\_TIDE Demo**
   1. **Harmonic analysis driven by river discharge**

First of all, you need to load the data provided by R\_TIDE Toolbox (such as Data\_Yangtze\_river.mat). The demo can be executed **using the main program labelled by ‘R\_demo\_Yangtze.m’.**

**The data file ‘Data\_Yangtze\_river.mat’ contains 2 variables, including ‘stname’ and ‘ZQ’.**

‘stname’ denotes the name of tidal gauging stations, including 6 columns (e.g., TSG, JY, ZJ, NJ, MAS, WH, respectively).

For instance:

'TSG'

'JY'

'ZJ'

'NJ'

'MAS'

'WH'

‘*ZQ*’ denotes hourly data used for harmonic analysis. The data in the 1st column denote the time series of the input data in term of ‘datenum’. The data between the 2nd and the 7th column denote the water level series observed in the tidal stations mentioned above. For instance, there are 6 columns of water levels in this variable, the data in the 2nd column represent the water levels in TSG and the data in the 7th column represent the water levels in WH. The data in the 8th column denote hourly river discharge data used for harmonic analysis.

For instance:

731217.166666667 0.7818 -0.0565 1.4980 2.1796 1.9357 2.2544 12731.5430

731217.208333333 1.2703 0.6749 0.3553 1.3052 1.8316 2.1548 12770.1389

731217.250000000 1.4570 1.2471 0.1852 0.8740 1.7609 2.0815 12806.2174

731217.291666667 1.3719 1.5003 0.6139 0.7844 1.7172 2.0306 12839.8438

731217.333333333 1.1476 1.4408 1.2681 0.9349 1.6940 1.9980 12871.0829

……

**The syntax of the main subroutine is illustrated below:**

[nameu,fu,yout,st,ft,Eta,Phi,percent,si,cof,tidecon]=R\_tide(xin1,Q1,T1,lat1,ray,synth,Qc1(i,:),twin,sname,ipso,ipre,tsnr,tau(i));

**Descriptions of the inputs:**

xin1: hourly water level data used for harmonic analysis

Q1: hourly river discharge data used for harmonic analysis

T1: the corresponding time series of the input data in term of ‘datenum’

lat1: the latitude of the selected station

ray: Rayleigh criteria, the default value is 1, which indicates that the Rayleigh criteria is used to select tidal constituents. Otherwise, Rayleigh criteria is not used.

synth: signal noise ratio, the default value is 10. You can set it depending on your own time series.

Qc1: the critical discharge beyond which the tide is vanishing, the default value is the maximum of the *Q* variable. Generally, it can be set to be the value corresponding with a negligible tidal range last for more than 2 days.

twin: window spectrum, the default value is 366.

sname: the name of the selected station

ipso: the method you used for harmonic analysis. If ipso=1, it will invoke standard PSO (Particle Swarm Optimization) algorithm to optimize and save the optimized results. If ipso=2, it will directly invoke the optimized file derived from PSO last time. If ipso=3, it will invoke default Matlab FMINCON Function to optimize and save the optimized results. If ipso=4, it will directly invoke the optimized file derived from FMINCON function last time. **So if it is the first time to use it, this parameter should be set as 1 or 3.**

ipre: the method you used for prediction. If ipre=1, it will save the coefficients for prediction; If ipre=2, it will directly predict. Divide the time series into 2 parts, one of which is used to derive coefficient for prediction, the other is used for prediction. Users should set the length of time series in the file, namely, Rtide\_pre.m. **So if it is the first time to use it, this parameter should be set as 1.**

tsnr: decide whether you’d like to add an error estimation. If tsnr=0, it will output the tidal properties. If tsnr=1, it will add an error estimation invoking the function ‘Rtide\_harmonic\_witherr.m’;

tau: the default value of travelling time of river discharge propagating to the studied tidal gauging station.

**Descriptions of the outputs:**

nameu: the name of the selected tidal constituents

fu: the frequency of the selected tidal constituents (/h)

yout: the reconstructed water level derived from R\_TIDE, consisting of st and ft

st: the reconstructed water level derived by the residual water level model

ft: the reconstructed water level derived by the tidal-fluvial model

Eta: the time-dependent amplitude of each tidal constituent derived from R\_TIDE

Phi: the time-dependent phase of each tidal constituent derived from R\_TIDE

precent: the correlation coefficient between reconstructed and observed water levels

si: the Root Mean Square Error (RMSE) between reconstructed and observed water levels

cof: the regression coefficients adopted for each tidal band derived from R\_TIDE

tidecon: the tidal properties together with their errors;

**The code is demonstrated below, you can copy and run it. The instructions introduced above are also listed in the main program.**

% R\_DEMO - demonstration of capabilities used in Yangtze River Estuary,

% The parameters input:

% xin1 -- water level data used for harmonic analysis

% Q1 -- discharge data used for harmonic analysis

% T1 -- the corresponding time series of the input data in term of 'datenum'

% lat1 -- the latitude of the selected station

% ray -- Rayleigh criteria, the default value is 1, which indicates that the Rayleigh criteria is used to select tidal constituents. Otherwise, Rayleigh criteria is not used.

% synth -- signal noise ratio, the default value is 10. You can set it depending on your own time series

% Qc1 -- the critical discharge beyond which the tide is vanishing, the default value is the maximum of the Q variable. Generally, it can be set to be the value corresponding with a negligible tidal range last for more than 2 days

% twin -- window spectrum, the default value is 366.

% sname -- the name of the selected station

% ipso -- the method you used for harmonic analysis. If ipso=1, it will invoke PSO to optimize and save the optimized results. If ipso=2, it will directly invoke the optimized file derived last time. So if it is the first time to use it, this parameter should be set as 1.

% ipre -- the method you used for prediction. If ipre=1, it will save the coefficients for prediction and then invoke the program, namely, Rtide\_predict.m; If ipre=2, it will directly predict. Divide the data series into 2 parts, one of which is used to derive coefficient for prediction, another one is used to predict. Users should set the length of data series in the file, namely, Rtide\_pre.m. So if it is the first time to use it, this parameter should be set as 1.

% tsnr -- error estimation. if tsnr=0, derive tidal properties; if tsnr=1, add an error estimation.

% tau -- the default value of travelling time of river discharge propagating to the studied tidal gauging station.

% The parameters output:

% nameu -- the name of the selected tidal constituents

% fu -- the frequency of the selected tidal constituents (/h)

% yout -- the reconstructed water level derived from R\_TIDE, consisting of st and ft

% st -- the reconstructed water level derived by the residual water level model

% ft -- the reconstructed water level derived by the tidal-fluvial model

% Eta -- the time-dependent amplitude of each tidal constituent derived from R\_TIDE

% Phi -- the time-dependent phase of each tidal constituent derived from R\_TIDE

% precent -- the correlation coefficient between reconstructed and observed water levels

% si -- the Root Mean Square Error (RMSE) between reconstructed and observed water levels

% cof -- the regression coefficients adopted for each tidal constituent bands derived from R\_TIDE

% tidecon -- the tidal properties together with their errors;

% Reference:

% Contract: Huayang Cai (E-mail: caihy7@mail.sysu.edu.cn); Suying Ou (E-mail: ousuying@mail.sysu.edu.cn)

%%

clc,clear

close all

load Data\_Yangtze\_river.mat % Load the example.

mm=find(ZQ(:,1)==datenum(2009,1,1,0,0,0));

ii1=1:mm-1; % select the time series for calibration (2002-2008)

T1=ZQ(ii1,1); % time series from 2002 to 2008

Q1=ZQ(ii1,end); % river discharge from 2002 to 2008

Z1=ZQ(ii1,2:end-1); % water levels from 2002 to 2008

ii2=mm:length(ZQ);%select time series for prediction (2009-2012);

T2=ZQ(ii2,1); % time series from 2002 to 2008

Q2=ZQ(ii2,end); % river discharge from 2002 to 2008

Z2=ZQ(ii2,2:end-1); % water levels from 2002 to 2008

Qc1=max(Q1); % the default value of critical river discharge

ray=1; % Rayleigh criteria

synth=1; % signal-to-noise ratio

twin=366; % days for spectrum window

tsnr=0;

% if tsnr=0, derive tidal properties tidecon;

% if tsnr=1, add the error estimation together with the output tidecon;

ipso=2;

% if ipso=1, using PSO to optimize and save the optimized results.

% if ipso=2, directly using the optimized file derived from PSO last time.

% if ipso=3, using Fmincon function to optimize and save the optimized results.

% if ipso=4, directly using the optimized file derived from Fmincon function last time.

ipre=2;

% if ipre=1, save the coefficients for prediction.

% if ipre=2, directly predict. Divide the time series into 2 parts, one of which is used to derive coefficient, the one is used for prediction.

% Users should set the length of time series in Rtide\_pre.m.

ic=1;

% if ic=0, ignore the influence of large river discharge on tidal properties;

% if ic=1, need a given critical Qc, the default value is the maximum value,

% then calibrate them according to the results derived from the first test

Qc1(1:6,1)=66000;

Qc1(1:6,2)=66000;

if ic==1

Qc1(1:6,1)=[66000;66000;66000;66000;66000;66000];

else

Qc1(1:6,2)=max(Q1);

end

tau=[15;13;9;7;5;4]; % the default value of tau, representing the default value of travelling time of river discharge propagating to the studied tidal gauging station.

% for i=1:size(Z1,2)

% xin1=Z1(:,i); % read the water level of selected time series in each tidal gauging station

% dH=0.05;% The cofficient of critical tidal range, the corresponding river discharge is the critical discharge Qc2. If the river discharge exceeds this value, there is no tidal signal in this station.

% Qc0=R\_Qtidal(xin1,Q1,T1,dH); % calculate the critical river discharge Qc2,but for many staion,it have no

% Qc1(i,1)=Qc0;%mean no second critical ;

% end

lat=[32.04044;31.9260;32.2293;32.0518;31.7026;31.3291]; % latitude of the tidal gauging stations

% note: different from T\_tide and NS\_tide, the variables are not conveyed via invoking function, but assigned values directly in the main program

% This program does not provide more options such as the solutions to system of linear algebraic equations, and does not add analysis of inference tide

% end of the input variables

%% calibration

% harmonic analysis

for i=1:size(Z1,2)

xin1=Z1(:,i); % read the water level of selected time series in each tidal gauging station

sname=char(stname(i));% read the name of tidal gauging station

lat1=lat(i); % read the latitude of tidal gauging station

[nameu,fu,yout,st,ft,Eta,Phi,percent,si,cof,tidecon]=R\_tide(xin1,Q1,T1,lat1,ray,synth,Qc1(i,:),twin,sname,ipso,ipre,tsnr,tau(i));% run R\_tide program

per(i)=percent;rmse(i)=si;

if (ipso==1 || ipso==2)

fname1=['R\_TIDE\_calibration\_for\_PSO\_' sname '.mat'];

end

if (ipso==3 || ipso==4)

fname1=['R\_TIDE\_calibration\_for\_fmin\_' sname '.mat'];

end

save(fname1,'T1','xin1','Q1','st','ft','percent','Eta','Phi','cof','si','yout','Qc1','nameu','tidecon'); % save the results

%discrete case

%[y1,y2,yok,yms]= Rtide\_discrete(xin1,Q1,T1,sname)

%% validation

% In this part, we use the coefficients derived from calibration period

% as well as the river discharge records in the rest time series to

% reconstruct the water level series, which reveal the influence of the

% alteration in river discharge, ignoring the geometric change

% \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

%prediction

if ipre==2

[z1,y2,percent1,si1,st,ft]=Rtide\_pre(Z2(:,i),Q2,T2,sname);

if (ipso==1 || ipso==2)

fname2=['R\_TIDE\_outputforper\_for\_PSO\_' sname '.mat'];

end

if (ipso==3 || ipso==4)

fname2=['R\_TIDE\_outputforper\_for\_fmin\_' sname '.mat'];

end

save (fname2,'z1','y2','cof','percent1','st','ft','Q2','si1')

end

end

It is worth noting that we assume that the geometric boundary during the validation period is more or less the same as during the calibration period, which is not completely true owing to the river discharge modulation (such as dam’s operation) as well as intensive human interventions (such as dredging or land reclamation). Thus, the aim of validation is to quantify the impacts of external forcing on water level dynamics, including the combined effects due to geometric change and mean sea level, as well as the impact of river discharge regulation caused by dam’s operation.

The model performances in terms of RMSE and the coefficient of determination (R2) are shown in Table 1, for both the calibration and validation periods. The RMSE values for both the calibration and validation periods were always less than 0.30 m, and the values of R2 were always larger than 0.91, which suggests that the model can successfully reproduce the water level dynamics along the estuary.

Table 1 Model performances for the 6 gauging stations along the Yangtze River estuary

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Period | Parameters | TSG | JY | ZJ | NJ | MAS | WH |
| Calibration  (2002-2008) | RMSE/m | 0.24 | 0.23 | 0.25 | 0.25 | 0.26 | 0.26 |
| R2 | 0.93 | 0.93 | 0.96 | 0.97 | 0.98 | 0.99 |
| Validation  (2009-2012) | RMSE/m | 0.26 | 0.25 | 0.26 | 0.28 | 0.30 | 0.29 |
| R2 | 0.91 | 0.92 | 0.95 | 0.97 | 0.97 | 0.98 |