**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))

2022/01/28

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

Version 1.0 - January 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, Water Resources Research, submitted

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]=R\_tide(xin,Q,T,lat1,ray,synth,Qc,twin,sname,ipso,ipre);

**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 and then invoke the program, namely, Rtide\_predict.m; 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.**

**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 coeffiicents adopted for each tidal bands derived from R\_TIDE

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

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

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

ray=1; % Rayleigh criteria

synth=10; % signal-to-noise ratio

twin=366; % days for spectrum window

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=1;

% If ipre=1, save the coefficients for prediction, invoke Rtide\_predict.m;

% 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.

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

dH=0.001;% The critical tidal range, if dH=0.001, the corresponding river discharge is the critical discharge Qc. If the river discharge exceeds this value, there is no tidal signal in this station.

Qc1=R\_Qtidal(xin1,Q1,T1,dH,Qc1); % calculate the critical river discharge Qc

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

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'); % save the results

N=1;

n=length(fu);% major tidal constituents

M1=length(Q1);

pl=cof(1,1);TauQ=fix(cof(1,2));

iq1=1:M1-TauQ;

iz1=iq1+TauQ;

z1=xin1(iz1,1);

time1=T1(iz1,1);

QQ1=Q1(iq1);

for j=1:n % coefficient matrix

x11(:,j)=cos(time1\*2\*pi\*fu(j)\*24);

x21(:,j)=sin(time1\*2\*pi\*fu(j)\*24);

x31(:,j)=QQ1.^pl.\*x11(:,j);

x41(:,j)=QQ1.^pl.\*x21(:,j);

end

xx=[ones(length(x11),1) QQ1.^pl x11 x21 x31 x41];

y=z1;

b1 = regress(y,xx);

y1=xx\*b1; % regression model to determine coefficients d and a

si=nanstd(y-real(y1)); % RMSE

percent=nanvar(real(y1))/nanvar(y)\*100; % R^2 denotes the Correlation coefficient

st(iz1,1)=[ones(length(x11),1) QQ1.^pl ]\*b1(1:2); % reconstructed s(t)

ft(iz1,1)=[x11 x21 x31 x41]\*b1(3:end); % reconstructed f(t)

y2(iz1,1)=y1; % regression to obtain water level

%%

figure1=figure;

plot(time1,y,time1,y1,'r',time1,st(iz1,1),'g')

title(['R^2= ' num2str(percent/100,3)],'fontname','Times New Roman')

xlabel('Date')

ylabel('\itZ \rm(m)')

dateFormat =2;

%title(str(k))

datetick('x',dateFormat)

xlim([min(time1) max(time1)])

set(gca,'fontname','Times New Roman')

legend('\itZ\_{\rmObs.}','\itZ\_{\rmSim.}','s(t)')

%%

% compute the time-dependent amplitudes and phases of n tidal constituents

for kk=1:n %the number of tidal constituents

Ak=sqrt(b1(kk+2).^2+b1(n+kk+2).^2); Bk1=sqrt(b1(2\*n+kk+2).^2+b1(3\*n+kk+2).^2);

Bk=QQ1.^pl.\*sqrt(b1(2\*n+kk+2).^2+b1(3\*n+kk+2).^2);

Alphak=atan2(b1(n+kk+2),b1(kk+2));

Betak=atan2(b1(3\*n+kk+2),b1(2\*n+kk+2));

Zk1=0.5\*(Ak\*cos(Alphak)+Bk.\*cos(Betak));

Zk2=0.5\*(Ak\*sin(Alphak)+Bk.\*sin(Betak));

Zfk=Zk1+Zk2\*j;

Zk=Zk1-j\*Zk2;

Eta2(iz1,kk)=abs(Zk)+abs(Zfk);

Phi2(iz1,kk)=90-atan2d(Zk2,Zk1);

m=find(Phi2(iz1,kk)<0);Phi2(m,kk)=Phi2(m,kk)+360;

ab(1:7,kk)=[Ak Bk1 Alphak Betak b1(1) b1(2) b1(3)];

end

jj=0;

for k=1:24:size(xin1,1)-24-floor(cof(2))

k1=k; jj=jj+1; k2=k1+23; kk1=k1:k2;

zm(jj,1)=mean(xin1(kk1,1));

ym(jj,1)=mean(yout(kk1,1));

stm(jj,1)=mean(st(kk1,1));

ftm(jj,1)=mean(ft(kk1,1));

qu(jj,1)=mean(QQ1(kk1));

for kk=1:n

Eta(jj,kk)=mean(Eta2(kk1,kk));

Phi(jj,kk)=mean(Phi2(kk1,kk));

end

yy1(iz1,1)=y1;

end

% compute variances for s(t)\_zr and f(t)\_ztr through hourly data and daily averaged data

varhour(k,1)=var(y);

varhour(k,2)=var(st(:,1));

varhour(k,3)=var(ft(:,1));

varday(k,1)=var(zm(:,1));

varday(k,2)=var(stm(:,1));

varday(k,3)=var(ftm(:,1));

clear z1 u1 timet1 x11 x21 x31 x41 x51 x61

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,'xin1','y2','Eta','Phi','ab','cof','b1','percent','st','ft','QQ1')

%% 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

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

ii2=mm:size(ZQ,1); % time series from 2009 to 2012

z2=ZQ(ii2,2:end-1);

xin2=z2(:,i);% water level from 2009 to 2012

T2=ZQ(ii2,1);% time series from 2009 to 2012

Q2=ZQ(ii2,end); % river discharge record from 2009 to 2012

M2=length(Q2);

pl=cof(1,1);TauQ=fix(cof(1,2));

iq2=1:M2-TauQ;

iz2=iq2+TauQ;

z2=xin2(iz2,1);

time2=T2(iz2,1);

QQ2=Q2(iq2);

for j=1:n

x12(:,j)=cos(time2\*2\*pi\*fu(j)\*24);

x22(:,j)=sin(time2\*2\*pi\*fu(j)\*24);

x32(:,j)=QQ2.^pl.\*x12(:,j);

x42(:,j)=QQ2.^pl.\*x22(:,j);

end

clear j

xx=[ones(length(x12),1) QQ2.^pl x12 x22 x32 x42];

yy2=xx\*b1; %regression model to determine coefficients d and a

Bias=nanmean(z2-yy2);

si=nanstd(z2-real(yy2)); %RMSE

%st(iz,k)=[ones(length(x1),1) q1.^pl ]\*b1(1:2); %reconstructed s(t)

%ft(iz,k)=[x1 x2 x3 x4]\*b1(3:end); %reconstructed f(t)

YY2(iz2,1)=yy2; %regression to obtain water level

percent=nanvar(real(yy2))/nanvar(z2)\*100; %R2:Correlation coefficient

for kk=1:n

Ak=ab(1,kk);

Bk=QQ2.^pl\*ab(2,kk);

Alphak=ab(3,kk);

Betak=ab(4,kk);

Zk1=0.5\*(Ak\*cos(Alphak)+Bk.\*cos(Betak));

Zk2=0.5\*(Ak\*sin(Alphak)+Bk.\*sin(Betak));

Zfk=Zk1+Zk2\*j;

Zk=Zk1-j\*Zk2;

Etay(iz2,kk)=abs(Zk)+abs(Zfk);

Phiy(iz2,kk)=90-atan2d(Zk2,Zk1);

%m=find(Phi2(iz,kk)<0);Phi2(m,kk)=Phi2(m,kk)+360;

end

bb1(1:3)=ab(5:7,1,1);

sty(iz2,1)=[ones(length(z2),1) QQ2.^pl ]\*bb1(1:2)';

jj=0;

% calculate the dailymean value

m=find(Etay==0);Etay(m)=nan;Etam=reshape(nanmean(reshape(Etay(1:end-20,:),24,[])),[],length(fu));

m=find(Phiy==0);Phiy(m)=nan;Phim=reshape(nanmean(reshape(Phiy(1:end-20,:),24,[])),[],length(fu));

m=find(xin2==0);xin2(m)=nan;ym=reshape(nanmean(reshape(xin2(1:end-20,:),24,[])),[],1);

m=find(YY2==0);YY2(m)=nan;yym=reshape(nanmean(reshape(YY2(1:end-20,:),24,[])),[],1);

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

fname3=['R\_TIDE\_validation\_for\_PSO\_' sname '.mat'];

end

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

fname3=['R\_TIDE\_validation\_for\_fmin\_' sname '.mat'];

end

save(fname3,'Etay','Etam','Phiy','Phim','xin2','ym','yym','nameu')

clear z2 u2 timet1 x12 x22 x32 x42 x52 x62

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 0.27 m, and the values of R2 were always larger than 0.90, 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 | 0.24 | 0.25 | 0.25 | 0.26 | 0.24 |
| R2 | 0.93 | 0.93 | 0.96 | 0.98 | 0.98 | 0.99 |
| Validation  (2009-2012) | RMSE/m | 0.27 | 0.26 | 0.26 | 0.26 | 0.27 | 0.27 |
| R2 | 0.92 | 0.92 | 0.95 | 0.97 | 0.98 | 0.98 |