%**Step 1**, data preparation%

cd 'C:\PARAFAC\EEM'

filetype=1;ext = 'csv';RangeIn='A1..AL262';headers=[1 1];display\_opt=0;outdat=1;

[X,Emmat,Exmat,filelist\_eem,outdata]=readineems(filetype,ext,RangeIn,headers,display\_opt,outdat);

Ex=Exmat(1,:);

Em=Emmat(:,1);

cd 'C:\PARAFAC\WATER'

filetype=1;ext = 'csv';RangeIn='A1..AL262';headers=[1 1];display\_opt=0;outdat=1;

[X\_b,Emmat\_b,Exmat\_b,filelist\_b,outdata\_b]=readineems(filetype,ext,RangeIn,headers,display\_opt,outdat);

Exb=Exmat\_b(1,:);

Emb=Emmat\_b(:,1);

cd 'C:\PARAFAC\RAMAN'

RamEx=350;

filetype='Raman275';

ext = 'csv';RangeIn='A1..B166';display\_opt=0;outdat=1;

[S\_R,W\_R,wave\_R,filelist\_R]=readinscans(filetype,ext,RangeIn,display\_opt,outdat);

cd 'C:\PARAFAC\UV'

filetype='Abs'; ext='csv'; RangeIn='A1..B351';

display\_opt=0; outdat=1;

[S\_abs,W\_abs,wave\_abs,filelist\_abs]=readinscans(filetype,ext,RangeIn,display\_opt,outdat);

cd 'C:\PARAFAC\Correction'

Excor=csvread('CorEX.csv');

Emcor=csvread('CorEM.csv');

cd 'C:\PARAFAC'

[LogNUM,LogTXT]=xlsread('SampleID.xlsx','SF-6-P');

Log\_EEMfile=LogTXT(:,7);

Log\_ABSfile=LogTXT(:,8);

Log\_Blkfile=LogTXT(:,9);

Log\_RAMfile=LogTXT(:,10);

Pair\_EEM\_abs=[Log\_EEMfile Log\_ABSfile];

Pair\_EEM\_R=[Log\_EEMfile Log\_RAMfile];

Pair\_EEM\_bk=[Log\_EEMfile Log\_Blkfile];

Pair\_BLK\_R=[Log\_Blkfile Log\_RAMfile];

[Xabs,PL]=matchsamples(filelist\_eem,filelist\_abs,Pair\_EEM\_abs,X,S\_abs);

[Sr,PLr]=matchsamples(filelist\_eem,filelist\_R,Pair\_EEM\_R,X,S\_R);

[Xbk,PLbk]=matchsamples(filelist\_eem,filelist\_b,Pair\_EEM\_bk,X,X\_b);

[Sbk\_R,PLbk\_R]=matchsamples(filelist\_b,filelist\_R,Pair\_BLK\_R,X\_b,S\_R);

A=[wave\_abs;Xabs]

B=Xbk;

W=[wave\_R;Sr];

T=[wave\_R;Sbk\_R];

[XcRU Arp IFCmat BcRU XcQS QS\_RU]=FDOMcorrect(X,Ex,Em,Emcor,Excor,W,[],A,B,[],[],[]);

originaldata.Ex=Ex';

originaldata.Em=Em;

originaldata.X=XcRU;

originaldata.nEx=length(Ex);

originaldata.nEm=length(Em);

originaldata.nSample=size(X,1);

Xs=smootheem(originaldata,[18 15],[15 15],[17 18],[19 18],0);

%**Step 2**, check the plot%

PlotEEMby1(1:40,Xs,'R.U.')

eemview(Xs,'X',[5 4])

%**Step 3** PARAFAC modeling and validation%

[AnalysisData]=SplitData(Xs);

[AnalysisData]=SplitHalfAnalysis(AnalysisData,(3:7),'MyData.mat');

SplitHalfValidation(AnalysisData,'1-2',3);

SplitHalfValidation(AnalysisData,'3-4',3);

[AnalysisData]=RandInitAnal(AnalysisData,3,5);

SplitHalfValidation(AnalysisData,'1-2',4);

SplitHalfValidation(AnalysisData,'3-4',4);

[AnalysisData]=RandInitAnal(AnalysisData,4,5);

SplitHalfValidation(AnalysisData,'1-2',5);

SplitHalfValidation(AnalysisData,'3-4',5);

[AnalysisData]=RandInitAnal(AnalysisData,5,5)

SplitHalfValidation(AnalysisData,'1-2',6);

SplitHalfValidation(AnalysisData,'3-4',6);

[AnalysisData]=RandInitAnal(AnalysisData,6,5)

SplitHalfValidation(AnalysisData,'1-2',7);

SplitHalfValidation(AnalysisData,'3-4',7);

[AnalysisData]=RandInitAnal(AnalysisData,7,5)

%**Step 4**, compare the residual of different model based on the component number%

PARAFACcomparison (AnalysisData,3,4,5,6,7)

% **Step 5** modelout for 3 component %

ComponentEEM(AnalysisData,3);

[FMax,B,C]=ModelOut(AnalysisData,3,'C:\outputname.xlsx')

% modelout for 4 component %

ComponentEEM(AnalysisData,4);

[FMax,B,C]=ModelOut(AnalysisData,4,'C:\outputname.xlsx')