Climate Predictability Tool (CPT)



Ousmane Ndiaye and Simon J. Mason

cpt@iri.columbia.edu



International Research Institute for Climate and Society The Earth Institute of Columbia University



OVERVIEW

 The Climate Predictability Tool (CPT) provides a Windows package for :

- Ø seasonal climate forecasting
- Ø model validation
- Ø actual forecasts given updated data
- Uses ASCII input files
- Options :
 - Ø principal components regression (PCR)
 - Ø canonical correlation analysis (CCA)
- Help Pages on a range of topics in HTML format
- Options to save outputs in ASCII format and graphics as JPEG
- Program source code is now available for those using other systems (e.g., UNIX)



SELECTING THE ANALYSIS

Climate Predictability Tool, v. 6.03

_ D ×

File View Help

Canonical Correlation Analysis (CCA) Principal Components Regression (PCR)

CLIMATE PREDICTABILITY TOOL

Evaluating seasonal climate predictability Designed for MOS applications

Copyright 2003 International Research Institute for Climate and Society

IRI INTERNATIONAL RESEARCH INSTITUTE FOR CLIMATE AND SOCIETY

Choose the analysis to perform: PCR or CCA



INPUT DATASETS

📲 Climate Predictability Tool, v. 6.03 - Input Window	
File Edit Actions Options View Help	
Principal Compo	nents Regression
PROJECT: Explanatory (X) variables: Training data file: X input file: Number of gridpoints: Number of data in file: First year of data in file: First year of X training period: 1950 \$	Response (Y) variables: Training data file: X input file: browse Number of series: 1 First year of data in file: 1950 First year of Y training period: 1950
EOF modes:Minimum number of modes:1Maximum number of modes:1	Training data:Length of training period:30 \$Length of cross-validation window:5 \$

Both analysis methods require two datasets: "X variables" or "X Predictors" dataset; "Y variables" or "Y Predictands" dataset.



M	licrosoft Exc	el - Exampi	le_station							_ 🗆 🗙
	<u>File E</u> dit <u>V</u> i	ew <u>I</u> nsert	F <u>o</u> rmat <u>T</u> ools	<u>D</u> ata <u>W</u> in	dow <u>H</u> elp					_ 8 ×
D	🖻 🖬 🔒	BB	à 🛍 🖌	$\rightarrow \Sigma f_i$, <u>Z</u> [2			v 10	- B	
	H4	+	= -0.22			4				
	A	В	C	D	E	F	G	Н	1	J
1.	Stn	STN_A	STN_B	STN_C	STN_D	STN_E	STN_F	STN_G		
2	LAT	-33	-34	-31	-28	-24	-24	-28		
3	LONG	19	23	29	32	31	28	25		
4	1979	0.83	0	0.42	0.96	1.08	1.5	-0.22		
5	1980	-0.56	0	-0.37	-0.98	-0.39	-0.83	-0.58		
6	1981	1.19	0	-0.22	1.89	1.27	2.01	3.77		
7	1982	0.12	0	0.59	0	0.94	0.71	-0.17		
8	1983	2.16	0	2.88	2.05	2.34	0.81	0.63		
9	1984	-0.58	0	.1.15	0.96	.1.01	0.05	-0.25		
10	1985	-0.77	0	-0.99	-0.35	-0.76	-9999	-0.29		
11	1986	1.06	0	0.64	-0.32	0.42	1.18	0.79		
12	1987	0.21	0	-0.59	-9999	0.43	0.4	0.05		
13	1988	-0.44	0	-0.09	0.29	-0.03	-0.06	-0.43		
14	1989	0.17	0	-0.19	0.04	-0.44	0.11	-0.17		
15	1990	-0.29	0	1.77	-0.96	-0.91	-1.04	-0:78		
16	1991	1.33	0	-0.5	-0.58	-0.05	0.76	0.51		
17	1992	0.01	0	-1.11	-0.72	-1.1	-0.18	-0.02		
18	1993	-1.04	0	-0.57	-0.27	0.3	-0.44	-0.67		
19	1994	-0.54	0	-0.98	-1.4	-1.13	-1.22	-0.56		
20	1995	0.41	0	-0.25	0.31	-9999	0.55	-0.07		
21	1996	-0.94	0	-1.05	-1.04	-1.13	-0.63	-0.74		
22	1997	0.18	0	0.34	0.81	1.96	1.09	-0.51		
23	1998	-2.11	0	-1.24	-1.43	-1.41	-1.2	-0.81		
24	1999	1.22	0	1.38	-0.17	-0.69	-0.51	0:26		
25	2000	-0.24	0	1	0.89	-0.02	0.15	0.39		
26	2001	0.56	0	-0.48	-0.74	-0.79	0	1.26		
27	2002	-0.53	0	-1.04	-1	-0.71	-1.05	-0.66		
-78 I∢ ∢	► ► Exa	mple_statio	on /			1				
Dra	w • 🔓 🌀	AutoShap	es + 🔨 🔌		4 🙍	ð • 🔏 •	<u>A</u> - ≡	≡ ≓ •	7 .	
Rea	dy							NUM		

1. STATION files:

This file-type contains :

Station_name (without spaces; ≤16 characters)

Latitude (in signed degrees)

Longitude (signed degrees)

Year (in the first column)

Data (missing values should be filled with the same value, -9999 for example)

<u>Keywords:</u> STN, LAT, LONG



12	licrosoft Exe	el - Examp	le_unrefere	nced						
픤	<u>File Edit Vi</u>	ew <u>I</u> nsert	Format Tool	s <u>D</u> ata <u>W</u> in	idow <u>H</u> elp					
D			Q. 🔁 💌	$\nabla \cdot \Sigma f$		🎽 Aria	d	• 10	• B	- 🛄
	F5	•	= 0.94							
	A	В	C	D	E	F	G	Н		J
1	NAME	A	В	C	D	E	F	G		
2	1979	0.83	0	0.42	0.96	1.08	1.5	-0.22		
3	1980	-0.56	0	-0.37	-0.98	-0.39	-0.83	-0.58		
4	1981	1.19	0	-0.22	1.89	1.27	2.01	3.77		-
5	1982	0.12	0	0.59	0	0.94	0.71	-0.17		_
3	1983	2.16	0	2.88	2.05	2.34	0.81	0.63		
7	1984	-0.58	0	1.15	0.96	1.01	0.05	-0.25		
3	1985	-0.77	0	-0.99	-0.35	-0.76	-9999	-0.29		
3	1986	1.06	0	0.64	-0.32	0.42	1.18	0.79		
0	1987	0.21	0	-0.59	-9999	0.43	0.4	0.05		
1	1988	-0.44	0	-0.09	0.29	-0.03	-0.06	-0.43		
2	1989	0.17	0	-0.19	0.04	-0.44	0.11	-0.17		
3	1990	-0.29	0	1.77	-0.96	-0.91	-1.04	-0.78		
4	1991	1.33	0	-0.5	-0.58	-0.05	0.76	0.51		
5	1992	0.01	0	-1.11	-0.72	-1.1	-0.18	-0.02		
6	1993	-1.04	0	-0.57	-0.27	0.3	-0.44	-0.67		
7	1994	-0.54	0	-0.98	-1.4	-1.13	-1.22	-0.56		
8	1995	0.41	0	-0.25	0.31	-9999	0.55	-0.07		
9	1996	-0.94	0	-1.05	-1.04	-1.13	-0.63	-0.74		
0	1997	0.18	0	0.34	0.81	1.96	1.09	-0.51		
1	1998	-2.11	0	-1.24	-1.43	-1.41	-1.2	-0.81		
2	1999	1.22	0	1.38	-0.17	-0.69	-0.51	0.26		
3	2000	-0.24	0	1	0.89	-0.02	0.15	0.39		
24	2001	0.56	0	-0.48	-0.74	-0.79	0	1.26		
25	2002	-0.53	0	-1.04	-1	-0.71	-1.05	-0.66		
26										
27										
8	Exa	mple unrel	ferenced /							
Dra	aw + Dr Co	AutoShap	es • \ \		4	3 - 1 -	A - =	= Ξ		
-	adv.	1	alaa yaa ka					NUM		

2. <u>UNREFERENCED or</u> <u>Indices files</u>:

The data are not referenced (no latitudes and longitudes):

Index_name (without spaces; ≤16 characters)

Year (in the first column)

Data (with missing data)

Keywords: NAME or YEAR



	licrosoft Exc	el - Example_sta	tion							
	<u>File Edit Vie</u>	ew <u>I</u> nsert F <u>o</u> rmat	Tool	s <u>D</u> ata <u>W</u> in	ndow <u>H</u> elp					<u>- 8 ×</u>
	🗳 Open	Ctrl+O	6 ×	$\gamma \cdot \Sigma f$	× 🕌 🕐	🎇 🛔 Aria	1	→ 10	- B	⊞ • *
	Save <u>A</u> s	¢.	2							
1				D	E	F	G	H		J
1	<u>1</u> Example	e_unreferenced	B	STN_C	STN_D	STN_E	STN_F	STN_G	2	
2	<u>2</u> Example	e_station	-34	-31	-28	-24	-24	-28		
3	<u>3</u> sahind6	898	23	29	32	31	28	25		
4	4 \work\c	otisation	0	0.42	0.96	1.08	1.5	-0.22	2	
5		×	0	-0.37	-0.98	-0.39	-0.83	-0.58		
6 -	1001	1.10	0 1	-0.22	1.89	1.27	2.01	3.77		
7	1982	0.12	0	0.59	0	0.94	0.71	-0.17		
8	1983	2.16	0	2.88	2.05	2.34	0.81	0.63		
9	1984	-0.58	0	1.15	0.96	1.01	0.05	-0.25		
10	1985	-0.77	0	-0.99	-0.35	-0.76	-9999	-0.29		
11	1986	1.06	0	0.64	-0.32	0.42	1.18	0.79		
12	1987	0.21	0	-0.59	-9999	0.43	0.4	0.05		
13	1988	-0.44	0	-0.09	0.29	-0.03	-0.06	-0.43		
14	1989	0.17	0	-0.19	0.04	-0.44	0.11	-0.17		
15	1990	-0.29	0	1.77	-0.96	-0.91	-1.04	-0.78		
16	1991	1.33	0	-0.5	-0.58	-0.05	0.76	0.51		
17	1992	0.01	0	-1.11	-0.72	-1.1	-0.18	-0.02		
18	1993	-1.04	0	-0.57	-0.27	0.3	-0.44	-0.67		
19	1994	-0.54	0	-0.98	-1.4	-1.13	-1.22	-0.56		
20	1995	0.41	0	-0.25	0.31	-9999	0.55	-0.07		
21	1996	-0.94	0	-1.05	-1.04	-1.13	-0.63	-0.74		
22	1997	0.18	0	0.34	0.81	1.96	1.09	-0.51		
23	1998	-2.11	0	-1.24	-1.43	-1.41	-1.2	-0.81		
24	1999	1.22	0	1.38	-0.17	-0.69	-0.51	0.26		
25	2000	-0.24	0	1	0.89	-0.02	0.15	0.39		
26	2001	0.56	0	-0.48	-0.74	-0.79	0	1.26		
27	2002	-0.53	0	-1.04	-1	-0.71	-1.05	-0.66	\	
28	K	onle station				14				
Dra		AutoShanes - >			3 📶 👩	Br.H.	A . =	= = =		ل ت السب
Dea	dv	1. Teopulation (

The input files could be easily made using a spreadsheet such as Excel



Mic Mic	rosoft Exc	el - Example_	station							_ 🗆 ×
Eil	e <u>E</u> dit <u>V</u> ie	ew Insert For	mat <u>T</u> ools	Data <u>W</u> indo	w <u>H</u> elp					_ 8 ×
Save /	١s								? ×	>
9	Save in:	DATA			★ €		光 🎫 🔻 To	ools +		<u></u>
Jave git Image: Price and Pric										
		File <u>n</u> ame:	Example_st	ation			•	Save		
F	laces	Save as <u>t</u> ype:	Text (Tab d	elimited)			•	Cano	el	
20 21 22	1995 1996 1997	0.41 -0.94 0.18	Microsoft Excel Workbook 41 Web Page 34 Template 18 Text (Tab delimited) -0.74 -0.74 -0.51							
23	1998	-2.11	Microsoft Ex	Unicode Text Microsoft Excel 5.0/95 Workbook -0.81						
24	1999	1.22	0	1.38	-0.17	-0.69	-0.51	0.26		
25	2000	-0.24	0	1	0.89	-0.02	0.15	0.39		
26	2001	0.56	0	-0.48	-0.74	-0.79	1.05	1.26		
28	2002	-0.03		-1.04		-0.71	-1,00	-0.00		-
	• • Exar	mple_station /	/ 				· <u> </u>	·	7	
Draw Ready	• 13 (5) /	AutoShapes	• \ `\[·· <u>·</u> · ·				

In Excel the file should be saved as: "Text, tab delimited"



Climate Predictability Tool, v. 6.03 - Input Window	
File Edit Actions Options View Help	
Principal Compo	nents Regression
PROJECT: Explanatory (X) variables: Training data file: X input file: Number of gridpoints: First year of data in file: First year of X training period: 10 10 1950 \$	Response (Y) variables: Training data file: X input file: browse Number of series: 1 First year of data in file: 1950 First year of Y training period: 1950
EOF modes:Minimum number of modes:1Maximum number of modes:1	Training data:Length of training period:30 \$Length of cross-validation window:5 \$

To select input files just click on browse.



Climate Predictability Tool, v. 6.03 - Inp File Edit Actions Options View Help	ut Window	×
Open X Input File	?:	sion
Look in: Data Angola Botswana ECHAM ERSST2 SAWS CPT_sen ECMVF_FMA Example_gridded Example_station Example_station Example_unreference index301 index310 JFM_99-03 File name: Files of type: X Input	KatumaniONDfreq KatumaniONDRain KatumaniONDRain KatumaniONDRain KenyaONDPredictands KenyaONDPredictands MakinduONDFreq MakinduONDRain MakinduONDRain MakinduONDRain MakinduONDRain SADC_FMA SADC_FMA_unref SADC_FMA_unref SADC_OND SADC_OND SADC_OND Sahcmap6803 SST_reformat RAIN RAIN_missing RainfallOND5180_15x25 Read_ud Files (*.*) Open	SION (f) variables: g data file: fg data: fg da

CPT opens a browser, which by default looks for data in: C:\Documents and Settings\user\Application Data\CPT\DATA\ You can search for data from any other directory.



Climate Predictability Tool, v File Edit Actions Options View	. 6.03 - Input Window Help	_ 🗆 ×
PROJECT: Explan: Tri X input file: Number of gridpoints First year of data in f First year of X trainin	X domain X domain: (southern latitudes and western longitudes negative) Northernmost latitude (88): 5 Southernmost latitude (88): -15 Westernmost longitude (0): -50 Easternmost longitude (358):	× les: browse 1 1950 1950 ↓
Minimum number of Maximum number of	<u>O</u> K <u>C</u> ancel	30 +

For gridded and station datasets, CPT lets you choose the spatial domain over which you want to perform your EOF or CCA analysis. In general the domain is known in advance through experience.



Principal C	compo	nents Regression	
ROJECT:			
————— Explanatory (X) variables: —		Response (Y) variables:	
Training data file:		Training data file:	
X input file: ECMWF_FMA.tsv	browse	X input file: br	owse
Number of gridpoints:	81	Number of series:	
First year of data in file:	1958	First year of data in file: 🛛 🖊 19	50
First year of X training period:	1958 🗘	First year of Y training period:	950 😫
EOF modes:		Training data:	

You proceed in the same way to select your file containing the Y variables (predictands).



SETTING THE TRAINING PERIOD

Climate Predictability Tool, v. 6.03 - Input Window	
e Edit Actions Options View Help	
Principal Comp	onents Regression
PROJECT: Explanatory (X) variables: Training data file: X input file: ECMWF_FMA.tsv browse	Response (Y) variables: Training data file:
First year of X training period:	First year of Y training period:
EOF modes:	Training data: Length of training period: 30 + Length of cross-validation window: 5 +

By default CPT usually starts the analysis from the first years in the X and Y files; note that these years could be different. You would normally set them equal to the latest year in the two files.

You should make sure the lag is correct if you cross the calendar year while using the DJF or JFM season, for example. In this case the starting year for file X may need to be one year earlier than for file Y.



SETTING THE TRAINING PERIOD

Climate Predictability Tool, v. 6.03 - Input Window	
	nents Regression
PROJECT:	lents Regression
Explanatory (X) variables:	Response (Y) variables:
X input file: ECMWF_FMA.tsv Number of gridpoints: 81	Training data file: X input file: NE_Brazil.txt Number of stations: 71
First year of data in file:1958First year of X training period:1971	First year of data in file:1971First year of Y training period:1971
EOF modes: Minimum number of modes: Maximum number of modes: 1 +	Training data: Length of training period: 27 Length of cross-validation window: 5

You have to specify the length of the training period as well as the length of the cross-validation window.



SETTING ANALYSIS OPTIONS

Climate Predictability Tool, v. 6.03 - Input Window	
File Edit Actions Options View Help	
Principal Compo	nents Regression
PROJECT: Explanatory (X) variables: Training data file: X input file: ECMWF_FMA.tsv browse Number of gridpoints: 81 First year of data in file: 1958 First year of X training period: 1971 \$	Response (Y) variables: Training data file: X input file: NE_Brazil.txt Number of stations: 71 First year of data in file: 1971 First year of Y training period: 1971
EOF modes: Minimum number of modes: Maximum number of modes:	Training data:Length of training period:27Length of cross-validation window:5

You have to choose the number of EOFs for the predictor fields used to fit the model. If you set the minimum to be less than the maximum, CPT will find the optimum number of modes between the two numbers. However, if you set the minimum equal to the maximum, then CPT will use that number of modes.



MISSING VALUES

🖁 Climate Predictability Tool, v. 6.03 - Input Window	
File Edit Actions Options View Help	
X EOF Options Transform Y Data Missing Values Default Settings PROJECT	nents Regression
Explanatory (X) variables:	Response (Y) variables:
Training data file:	Training data file:
X input file: ECMWF_FMA.tsv browse	X input file: NE_Brazil.txt browse
Number of gridpoints: 81	Number of stations: 71
First year of data in file:	First year of data in file: 1971
First year of X training period:	First year of Y training period:
EOF modes:	Training data: Length of training period:
Maximum number of modes:	Length of cross-validation window: 5

If you have missing values in your dataset, you need to specify what you want CPT to do with them.



MISSING VALUES

Principal Compo	onents Regression
ing ¥alues	
Explanatory (X) variables: Missing value flag: Maximum % of missing values: 10 Missing Value Replacement: Select method: Output Long-term means Cong-term medians Random numbers Best nearest neighbour	Reponse (Y) variables: Missing value flag: -999 Maximum % of missing values: 10 Missing Value Replacement: 10 Select method: Long-term means Long-term medians Random numbers Best nearest neighbour Best nearest neighbour

Next to the Missing value flag box, you need to specify the number in your dataset that represents a missing value.

You can choose the maximum % of missing values. If a station has more than that percentage of missing values, CPT will not use that station in its model. You can also choose which method you want CPT to use to replace the values.



SAVING PROGRAM SETTINGS

Climate Predictability Tool, v. 6.03 - Input Window	
File Edit Actions Options View Help	
New (Ctrl+N) Open (Ctrl+O) Save (Ctrl+S) Save As	nents Regression
Exit (Ctrl+Q)	Response (Y) variables:
X input file: ECMWF_FMA.tsv browse	X input file: NE_Brazil.txt browse
First year of data in file:1958First year of X training period:1971	First year of data in file:1971First year of Y training period:1971
EOF modes: Minimum number of modes: Maximum number of modes: 4	Training data: Length of training period: 27 Length of cross-validation window: 5

Once you have selected the input files and your settings it is a good idea to save these settings in a project file to recall them later:

File => Save

By default, CPT saves all the project files in the subdirectory C:\Documents and Settings\user\Application Data\CPT\Projects\

RUNNING CPT

📲 Climate Predictability Tool, v. 6.03 - Input Window	
File Edit Actions Options View Help	
Calculate Cross-validated	
Retroactive al Compo	onents Regression
	onents Regression
PROJECT:	
Explanatory (X) variables:	Response (Y) variables:
Training data file:	Training data file:
X input file: ECMWF FMA.tsv browse	X input file: NE Brazil.txt browse
Number of gridpoints: 81	Image: Number of stations: 71
First year of data in file: 1958	First year of data in file: 1971
First year of X training period:	First year of Y training period:
EOE modes:	
Minimum number of modes:	Iraining data:
	Length of training period:
Maximum number of modes: 4	Length of cross-validation window: 5

Then you can run the analysis: Actions => Calculate => Cross-validated



DATA ANALYSIS

1	📲 Climate Predictability Tool, v. 6.03 - Results Window					
	File Tools Customise Help					
	Progress: 100 Actions:	%				
	Reading C:\Documents an Checking for missing va Reading C:\Documents an Checking for missing va Data read successfully	d Settings\ma lues d Settings\ma lues	ster09\Application	Data\CPT\Data\ECMW Data\CPT\Data\NE_B	F_FMA.tsv razil.txt	
	Optimizing cross-valida Training period: 1971 t	ted performan o 1997	ice			
	CURRENT		OPTIMUM			
	Number of Modes	Goodness Index	Number of Modes	Goodness Index		
	1	0.504	1	0.504		
	2	0.488	1	0.504		
	3	0.469	1	0.504		
	4	0.459	1	0.504		
	Cross-validating model Constructing model usin Identifying categories Calculating climatologi Done !	 g full traini es	ng period (1971 to	1997)		

CPT begins the specified analysis in a new "Results Window". Here you can see the steps of the analysis and of the optimization procedure.



DATA ANALYSIS

CURRENT	CURRENT		OPTIMUM		
Number of Modes	Goodness Index	Number of Modes	Goodness Index		
1	0.504	1	0.504		
2	0.488	1	0.504		
3	0.469	1	0.504		
4	0.459	1	0.504		

Optimizing the number of EOF modes:

CPT uses EOF #1 to make cross-validated forecasts then calculates a "goodness index" summarizing how good all the forecasts are (the closer to 1 the better). Then CPT uses EOF #1 and #2 to remake cross-validated forecasts and calculates a new goodness index for these, and so on until using all four EOFs.
 At each step CPT compares the goodness indices and retains under the column "OPTIMUM" the highest goodness index and the corresponding number of EOFs (in the example above, 1).
 CPT uses this number of EOFs (i.e. 1) to build the model.



RESULTS : graphics



The menu Tools => Graphics => Scree plots displays the percentage of variance associated with each EOF plotted.



RESULTS : graphics



1. The menu Tools => Graphics => X EOF loadings and scores

displays the loading pattern of each EOF and the temporal series.

2. CPT allows you to customize and save each graphic by:

right-clicking on the mouse selecting the graphic to customize / save



CHANGING THE TITLE



To change the title of the graph

- 1. right-click the mouse
- 2. go to EOF Loadings
- 3. click on Title



SAVING GRAPHICS



You can choose the name of the graphic output file by clicking on browse. You can adjust the quality of the JPEG graphic as well. All the output files are saved by default under: C:\CPT\Output\



SHOWING HIGH RESOLUTION MAPS



If you want to get a better quality map, you can change the setting to high resolution.

Customize => Graphics => High Resolution Map



RESULTS

🐮 🖪 Clin	nate Predictability Tool, v.	5.06 - Results Wi	ndow		
File 1	Tools Customise Help				
Pi Ai	Validation C Contingency Tables R Graphics Forecast	ross-validated + etroactive +	Performance Measures Bootstrap Skill Map Scatter Plots		
Read Check Read Check Data	ing C:\CPT\DATA\ECM king for missing va ing C:\CPT\DATA\NE_ king for missing va read successfully	WF_FMA.tsv lues Brazil.txt lues			
Trai	mizing cross-valida ning period: 1971 t	ced performa o 1997	ince		
CURRENT		OPTIMUM			
	Number of Modes	Goodness Index	Number of Mode	s Goodness Index	
	1	0.504	1	0.504	
	2	0.488	1	0.504	
	3	0.469	1	0.504	
	4	0.459	1	0.504	
Cros Cons Iden Done	s-validating model tructing model usin tifying categories !	 g full train 	ing period (1971 t	o 1997)	

To see the results go to the menu "Tools":

Validation : shows skill, hindcasts and observed series Contingency Tables : shows contingency tables Graphics : shows the EOFs time series, loading patterns and scree plot



RESULTS



To see the series forecasted and observed at each station/grid go to: **Tools** => Validation => Cross-Validated => Performance Measures This menu displays some statistics of the forecast, such as correlation coefficient, RMSE, ROC etc (for more details refer to the help page).



REVERSING THE COLORS



Customise => Graphics => Reverse Colors

If you are forecasting temperature instead of precipitation, then it would be more intuitive to have red (hot/above) and blue (cold/below), so you might want to invert the default colors. You might also want black and white images if they are to be included in a report or publication.



INDICATIONS OF UNCERTAINTY

oss-validated scores				
ation 1				
2.885, 40.12W				
Score:	— Sample: —	Confidence	limits:	P-value: -
Continuous measures:		Confidence level: 9	5.000%	
Pearson's correlation	0.6879	0.4751 to	0.8199	0.0000
Spearman's correlation	0.6288	0.3130 to	0.8107	0.0000
Mean squared error	67755.04	45905.87 to	94133.70	0.0000
Root mean squared error	260.30	214.26 to	306.81	0.0000
Mean absolute error	224.57	177.75 to	277.83	0.0000
Bias	-4.73	-100.96 to	87.31	N/A
ategorical measures:				
Hit score	62.96	44.44% to	81.48%	0.0000
Hit skill score	44.44	16.67% to	72.22%	0.0000
LEPS score	51.85	21.42% to	79.16%	0.0020
Gerritu score	50.00	19.15% to	76.13%	0.0020
ROC area (below-normal)	0.7654	0.5489 to	0.9387	0.0140
	0 9119	0 6000 to	0.0715	0 000

For indications of uncertainty in the performance measures go to: Tools => Validation => Cross-Validated => Bootstrap



ADJUSTING THE BOOTSTRAP SETTINGS

oss-validated scores					
tation 1 🛟	Resampling Options		<u>×</u>	1	
2.885, 40.12		- Bootstrapping: —			
Continuous mea Pearson's c	Number of bootstra Confidence level (ap samples: %):	500	ce limits: 95.000% p. 0.8199	- P-value:
Spearman's Mean square Root mean s Mean absolu	Number of permut	- Permutations: — ations:	500 🗘	0.8107 0 94133.70 0 306.81 0 277.83	0.0000 0.0000 0.0000 0.0000
Bias Categorical me Hit score Hit skill s	[<u>O</u> K <u>C</u> ancel		D 87.31	N/A 0.0000 0.0000
LEPS score Gerrity score ROC area (bel	ow-normal)	51.85 50.00 0.7654	21.42% 19.15% 0.5489	to 79.16% to 76.13% to 0.9387	0.0020 0.0020 0.0140

Customize => Resampling Settings CPT allows you to adjust the bootstrap settings.



RESULTS : data files

📲 Climate	Predictability Tool, v	. 6.03 - Results Wind	low		
File Tools	Customise Help				
Save	(Ctrl+S)				
Save As		0%			
Open For	recast File				
Data Out	put				
			-+		•
Close	T FOR DICCIDO I	nd Settings\ma	steruy\Application v	ata\CPI\Data\ECMWF_FMH.	tsv
Reading	C:\Documents a	alues nd Settinas\ma	ster09\Annlication D	ata\CPT\Data\NF_Brazil.	txt
Checking	q for missing v	alues			
Data rea	ad successfully				
0-1-1-1-1					
Uptimiz: Training	LNG Cross-Valid Deriod: 1071	ated performan to 1007	ce		
11 grurui	g period. 1971	0 1777			
	CURRENT		OPTIMUM		
N	umber of Modes	Goodness	Number of Modes	Goodness	
		Index		Index	
	1	0.504	1	0.504	
	2	0.488	1	0.504	
	3	0.469	1	0.504	
	. 4	0.459	1	0.504	
Cross-va	alidating model			007)	
Identify	cting model usi vina estoaccioc	ng full traini	ng perioa (1971 to 1	997)	
Calculat	ting categories	ies			
Done!	cing crimatorog				

The menu File => Data Output allows you to save output data:

- 1. EOFs: time series, loading patterns, variance
- 2. The parameters (coefficients) of the model (example: $Y = \underline{A}x + \underline{b}$)
- 3. The input data (with the missing values filled)
- 4. Cross-validated forecasted time series



SAVING OUTPUT FILES

Climate Predictability Tool, v. 6.03 - Results Window File Tools Customise Help Output file definitions	
Pile Edit Help Output files Historical Predictions EDF Prefiltering PCR Results Input Data Predictions Drowse Drowse	CMWF_FMA.tsv E_Brazil.txt
C C Identifying categories Calculating climatologies Done!	

In order to save the outputs in separate files, you have to specify a file name by clicking on browse. By default CPT saves the output files under: C:\CPT\Output\



📲 Climate Predictability Tool, v.	6.03 - Results Wind	ow		
File Tools Customise Help				
Save (Ctrl+S)				
Save As	%			
Open Forecast File				
Data Output				
Class	d Sattings) ma	ctor(00) Application (
	iu sectinys (ma: ilies	sceres/Hhbicacion i	Vacatori (Dacateonwr_rnm.cs)	' ,
Reading C:\Documents an	d Settinos\ma	ster09\Application [)ata\CPT\Data\NE_Brazil.txt	
Checking for missing va	lues			
Data read successfully				
Uptimizing cross-valida Training period: 1071 t	ited performan o 1007	ce		
fraining period. 1971 (.0 1997			
CURRENT		OPTIMUM		
Number of Modes	Goodness	Number of Modes	Goodness	
	Index		Index	
1	0.504	1	0.504	
2	0.488	1	0.504	
3	0.469	1	0.504	
4	0.459	1	0.504	
Cross-validating model	····			
Constructing model usin	ig full traini	ng period (1971 to ²	1997)	
Identifying categories	•••			
Done •	.es			
Done:				

Once your model is built, you can make a forecast using a forecast file with new records of the X variables: File => Open Forecast File



Forecast Input F	ile
nple	
— Forecast data file: —	
CMWF_FMA.tsv	browse
file:	1958
ch to forecast:	1971 🗘
forecast:	1 🛟
	nple — Forecast data file: — CMWF_FMA.tsv file: ch to forecast: forecast:

A new window is opened. By default CPT selects the same input predictor file. You can change it by clicking browse.



cast series		
	Forecast Input File	
ROJECT: E	xample	
	——— Forecast data file: ———	
Forecast file:	ECMWF_FMA.tsv	browse
First year of dat	ta in file:	1958
First year from	which to forecast:	2000
Number of year	s to forecast:	1

You then select:

- (a) the starting year of the forecasts
- (b) the number of years to forecast



🐮 Climate Predictability Tool, v	. 6.03 - Results Wind	ow				
File Tools Customise Help						
Validation 🕨						
👝 Contingency Tables 🕨 📊	1%					
Graphics 🕨 🂾						
A Forecast	Series					
	Exceedances					
Reading C:\Documents	Maps 🔹 🕨 🔪 🗛	ster09\Application D)ata\CPT\Data\ECMWF_FMA.tsv			
Checking for missing v	alues					
Reading C:\Documents a	nd Settings\mag	ster09\Application [<pre>Data\CPT\Data\NE_Brazil.txt</pre>			
Deta read successfully	alues					
Data read Successfully						
Optimizing cross-valid	ated performanc	:e				
Training period: 1971 to 1997						
ouppeut		00771000				
CURRENT		UPIIMUM				
Number of Modes	Goodness	Number of Modes	Goodness			
Number of Modes	Goodness Index	Number of Modes	Goodness Index			
Number of Modes	Goodness Index	Number of Modes	Goodness Index			
Number of Modes	Goodness Index 0.504	Number of Modes	Goodness Index 0.504			
Number of Modes 1 2	Goodness Index 0.504 0.488	Number of Modes 1 1	Goodness Index 0.504 0.504			
Number of Modes 1 2 3 4	Goodness Index 0.504 0.488 0.469 0.459	Number of Modes 1 1 1	Goodness Index 0.504 0.504 0.504 0.504			
Number of Modes 1 2 3 4 Cross-validating model	Goodness Index 0.504 0.488 0.469 0.459	Number of Modes 1 1 1 1	Goodness Index 0.504 0.504 0.504 0.504 0.504			
Number of Modes 1 2 3 4 Cross-validating model Constructing model usi	Goodness Index 0.504 0.488 0.469 0.459 ng full trainin	Number of Modes 1 1 1 1 1 ng period (1971 to 1	Goodness Index 0.504 0.504 0.504 0.504 0.504			
Number of Modes 1 2 3 4 Cross-validating model Constructing model usi Identifying categories	Goodness Index 0.504 0.488 0.469 0.459 ng full trainin	Number of Modes 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Goodness Index 0.504 0.504 0.504 0.504 0.504			
Number of Modes 1 2 3 4 Cross-validating model Constructing model usi Identifying categories Calculating climatolog	Goodness Index 0.504 0.488 0.469 0.459 ng full trainin ies	Number of Modes 1 1 1 1 ng period (1971 to 1	Goodness Index 0.504 0.504 0.504 0.504 0.504			

Once the file is selected and the years to forecast are chosen go to the menu Tools => Forecast => Series or Maps.





The option Series shows the predicted values (cross) for the current station as well as forecast possibilities, confidence limits for the forecast and, in the "Thresholds" box, the "category thresholds" as well as the climatological probabilities for the 3 categories.



CHANGING CATEGORY DEFINITIONS



CHANGING CATEGORY DEFINITIONS





EXPRESSING THE FORECAST AS ANOMALIES

Forecasts Station 2 4 6.57S, 40.12W	
Threshold Settings 1971-1997 clir Thresholds upper Standardization: lower Select method: Climatological No standardization above Anomalies Downal Standardization Probabilities: Probabilities:	s-Validated Hindcasts
Year Probabil B t 2000 33 20 Absolute thresholds: 0.2 Upper: 200	
For Iower: 100 Year Forecast QK Cancel 2000 373.991 Image: Cancel 1970 1975 1980	1985 1990 1995 2000 2005

The forecast can be expressed as anomalies, or as standardized anomalies, rather than as absolute values: Customize => Thresholds



EXPRESSING THE FORECAST AS ANOMALIES



The thresholds, as well as the forecast ranges, are now defined as anomalies. If absolute thresholds are set, CPT assumes that these are defined as anomalies.



PREDICTION INTERVALS



To draw error bars on the forecast, right click on the graph: Customize => Prediction Intervals An error bar is indicated.



CHANGING THE PREDICTION INTERVAL



You can also change the width of the prediction interval. Customize => Forecast Settings

The default setting of 68.2% gives standard error bars.



SAVING FORECASTS



To save the forecasts, right click, and specify the required output files.



CHANGING THE CLIMATOLOGICAL PERIOD

File Tools Customis	bility Tool, v. 6.03 - R e Help	esults Window		
Progress: Actions:	100%	•		
Reading C:\D	limatological Period	2	✓ ication	Data\CPT\Data\ECMWF_FMA.tsv
Checking for Reading C:\D Checking for Data read su	Please specify c l First year:	limatological period:	ication	Data\CPT\Data\NE_Brazil.txt
Optimizing c Training per	Last year:	1997 🗘		
	<u>0</u> K	<u>C</u> ancel	OPTIMUM	
Number	1.000			Goodness
		Index		Index
	1	0.504	1	0.504
	2	0.488	1	0.504
	3	0.469	1	0.504
0-0 0000000000000000000000000000000000	4	0.459	1	0.504
Cross-validati Constructing r Identifying ca Calculating cl Done!	ing model model using ful ntegories limatologies	l training period	(1971 to	1997)

By default, the forecast probabilities are calculated relative to a climatological period that is the same as the training period. To change the climatological period go to: Customise => Climatological Period



FORECAST MAPS

Forecast probabilities maps

Forecast: 2000

	— Forecast	s: ——		
1971-1971 cl	imatology			
Station	Below	Normal	Above	
2	28%	39%	34%	
4	32%	28%	40%	
11	25%	51%	24%	_
12	25%	36%	39%	
13	28%	36%	37%	
14	3 0%	41%	3 0%	
15	32%	26%	41%	
17	29%	27%	44%	
18	29%	39%	32%	
22	32%	38%	3 0%	
23	25%	53%	22%	
25	32%	40%	29%	
26	31%	23%	46%	
27	29%	26%	45%	
28	26%	43%	31%	
33	24%	38%	37%	•





Tools => Forecast => Maps

The option Maps lets you see maps of your forecasts – either maps of the probabilities or maps of the actual forecast values.

The forecast probabilities map lists the probabilities for each category at each location as well as the spatial distribution of the probabilities.

In this example it is evident that in 2000 the below-normal category has the lowest probability over most of northeast Brazil.



FORECAST MAPS

Forecast maps Forecast: 2000

Forecasts: 1971-1997 climatology Station Forecast Lower Upper 2 788.277 526.834 1049.721 4 373.998 248.826 499.170 11 816.287 562.244 1070.330 12 610.649 389.141 832.157 13 321.125 656.384 488.755 14 485.848 314.575 657.121 15 612.732 419.801 805.663 17 637.091 467.779 298.467 18 591.334 421.816 760.852 22 565.975 407.657 724.292 23 665.312 438.059 892.565 25 415.180 282.608 547.752 26 600.462 420.799 780.126 27 822.902 500.413 1145.391 28 353.259 224.892 481.627 33 466.224 869.382 667.803





The forecast values map lists the actual forecast values for each category at each location as well as the spatial distribution of the values.



EXCEEDANCE PROBABILITIES



To draw the probabilities of exceedance go to: Tools =>Forecast => Exceedances



CONCLUSIONS

- For further details, read the help page of each menu and option.
- Subscribe to the user-list to be advised of updates: <u>http://iri.columbia.edu/outreach/software/</u>

 We want to hear from you. Your comments and questions help us to improve the CPT so do not hesitate to write to us at: <u>cpt@iri.columbia.edu</u>

