Climate Predictability Tool (CPT)



Ousmane Ndiaye and Simon J. Mason

ousmane@iri.columbia.edu



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

SURVOL DU LOGICIEL

• Le Climate Predictability Tool (CPT) est un logiciel sous windows qui fait :

- Prévision saisonnière climatique
- Validation de modèle
- Vérification de modèle probabilistique
- Mise à jour de la prévision
- En quatre différentes langues
- Options d'Analyse statistique :
 - > Analyse en Corrélation Canonique (CCA)
 - > Régression en Composante Principale (RCP)
 - Regression Lineaire Multiple (RLM)
 - GCM validation (GCM)
 - Probabilistic Forecast Verification (PFV)
- Pages d'aide sur plusieurs rubriques en format HTML

Option de sauvegarde des sorties en fichier ASCII ou sous format graphique

COUNT RESERACION

CHARGER LES FICHIERS DE DONNEES

Variable Reponse (Y)	
summer reported (1)	Variables Prevues
navigateur	navigateur
Nom de fichier:	Nom de fichier:
Donnee Initiale: N/A	
Derniere donnee: N/A	Derniere donnee: N/A
Commencer à: 1	Commencer à: 1
Nombre de champs 0	Nombre de champs 0
Nombre de décalages 0	Nombre de décalages 0
Nombre de variables 0	Nombre de variables 0
Nombre utilisé 0	Nombre utilisé 0
	navigateur Nom de fichier: Donnée initiale: N/A Dernière donnée: N/A Commencer à: 1 1 + Nombre de champs 0 Nombre de décalages 0 Nombre de variables 0 Nombre utilisé 0

Utiliser le navigateur pour charger les deux fichiers d'entrée :

1. "Variables Explicatives (X)"

et

2. "Variables Réponse (Y)"



CHOIX D`UNE ANALYSE



On commence d'abord à choisir l'analyse à faire



FORMAT DES FICHIERS D'ENTRÉE SUR CPT

M	licrosoft Exe	cel - Exampl	e_station	** **						- 🗆 X
	<u>File E</u> dit <u>V</u> i	ew <u>I</u> nsert	F <u>o</u> rmat <u>T</u> ool	s <u>D</u> ata <u>W</u> in	idow <u>H</u> elp					_ B X
l D	🔗 🔲 🔒	ികിരി	a 🝙 🖌	$\nabla \star \Sigma f$, 7 1 (?)	» Aria		▼ 10	• B	• • »
	H4	- <u>-</u>	= -0.22	<u>, 1 - 1</u>		11				
	A	B	C	D	F	F	G	н	1	
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		
_28 € €	Exa	mple_statio	on /			•				
Dra	w • 🗟 🌀	AutoShap	es + 🔨 🔌		4 🙍	🕭 + 🏒 -	<u>A</u> - ≡	≡ ≓ •	7 .	
Rea	idy			Γ				NUM		

Fichier de type station :

Ce type de fichier contient sur les trois premières lignes :

- Nom des Station (sans espace; ≤16 caractères)
- 2) Latitude (en degré)
- 3) Longitude (en degré) Ensuite viennent les données Année (première colonne) et les Donnée (les données manquantes doivent être remplacée par la même valeur, -9999 par exemple)

Mots-clés : STN, LAT, LONG



FORMAT DES FICHIERS D'ENTRÉE SUR CPT

D	65 🖬 🔒			$\gamma \star \Sigma f$	× 🕌 🖸	¥ A	rial	• 10	• 8	<u> </u>
	A	B	= 0.94 C	D	F	F	G	н	1	
Ē	NAME	A	В	C	D	E	F	G	1) X:	
)	1979	0.83	0	0.42	0.96	- 1.0	18 1.5	-0.22	\$ 	-
L	1980	-0.56	Ő	-0.37	-0.98	-0.3	9 -0.83	-0.58		
	1981	1.19	Ő	-0.22	1.89	1.2	27 2.01	3.77		
	1982	0.12	0	0.59	0	0.9	0.71	-0.17		
ie.	1983	2.16	0	2.88	2.05	2.3	0.81	0.63		
2	1984	-0.58	0	1.15	0.96	1.0	0.05	-0.25		
É.	1985	-0.77	0	-0.99	-0.35	-0.7	6 -9999	-0.29		
E.	1986	1.06	0	0.64	-0.32	0.4	2 1.18	0.79		- î
)	1987	0.21	0	-0.59	-9999	0.4	3 0.4	0.05	0	
C.	1988	-0.44	0	-0.09	0.29	-0.0	.0.06	-0.43		
2	1989	0.17	0	-0.19	0.04	-0.4	4 0.11	-0.17		
3	1990	-0.29	0	1.77	-0.96	-0.9	-1.04	-0.78		
1	1991	1.33	0	-0.5	-0.58	-0.0	0.76	0.51		
5	1992	0.01	0	-1.11	-0.72	-1.	.1 -0.18	-0.02		
3	1993	-1.04	0	-0.57	-0.27	0.	3 -0.44	-0.67		
t	1994	-0.54	0	-0.98	-1.4	-1.1	3 -1.22	-0.56		- î
3	1995	0.41	0	-0.25	0.31	-999	9 0.55	-0.07	0	
3	1996	-0.94	Ö	-1.05	-1.04	-1.1	3 -0.63	-0.74		
)	1997	0.18	0	0.34	0.81	1.9	6 1.09	-0.51		
1	1998	-2.11	0	-1.24	-1.43	-1.4	1 -1.2	-0.81		
2	1999	1.22	0	1.38	-0.17	-0.6	9 -0.51	0.26		
3	2000	-0.24	0	1	0.89	-0.0	0.15	0.39		
4	2001	0.56	0	-0.48	-0.74	-0.7	9 0	1.26		
5	2002	-0.53	0	-1.04	-1	-0.7	1 -1.05	-0.66		
3										8
7										
2	IN INCO From					1.	1			

Fichier indice ou sans référence :

Les données sont sans coordonnées (pas de latitude et longitude):

Nom de l'indice (sans espace; ≤16 caractères)

Année (première colonne)

Données (avec les valeurs manquantes)

Mots-clés : NAME ou YEAR



FORMAT DES FICHIERS D'ENTRÉE SUR CPT

) 🖬 🄊 - 1) +	Name of Concession, or other	-	and Personnella	СРТ	tutorial_french1 [Mode de compatibili	té] - Microsoft P	owerPoint
6	1 - 1-	(²¹ -) =			De	ebitOMVS_ASO501	2 - Microsoft Excel		
Pre	Accueil Accueil Coller Ssse-papiers	Insertion Calibri G I S • F	Mise en page • 11 • A · · · · · · · · A · · · · · · · · · · · · · · · · · · ·		Enregistrer sous	Classeur Excel Classeur Excel (pre Classeur Excel bina Classeur Excel 97 - Données XML Page Web à fichier Page Web Modèle Excel Modèle Excel (prer Modèle Excel 97 -	nant en charge les macros) ire 2003 r unique nant en charge les macros) 2003		
1	A	В	С		👉 Favorites	Texte (séparateur :	tabulation)		
1	Stn	Oualia	BafingMakar	GO	Desktop	Texte Unicode	ML 2002		
2	LAT	13.6	12.55	13.	Downloads	Classeur Microsoft	t Excel 5.0/95		
3	LONG	-10.38333	-10.28333	-11	Recent Placer	CSV (séparateur: p	oint-virgule)		
4	1950	-999	-999	_	Necent Places	Texte (séparateur:	espace)		
5	1951	-999	-999		😑 uttantar	Texte (DOS)			
6	1952	-999	-999	-		CSV (Macintosh)			
7	1953	-999	-999	2.050	Documents	CSV (DOS)			
8	1954	866.4666667	-999	771	Music	SYLK			
9	1955	669.4666667	-999	673	Nom de fichier (Macro compléme	ntaire Excel		
10	1956	759.7333333	-999	645	Nom de fichier .	Macro compléme	ntaire Excel 97 - 2003		
11	1957	658.8333333	-999	317	Type:	Texte (séparateur :	tabulation)		•
12	1958	809.6666667	-999	609	Authors: c	ousmane	Tags: Add a tag		
13	1959	564.0333333	-999	475					
14	1960	336.6333333	-999	288					
15	1961	649.8	969.8666667	592	Cacher les dossiers	5	Outils 🔻 🔤 Ei	nregistrer	Annuler
16	1962	478.8333333	1077.166667	539	(20/41/0811)				
17	1963	387.3	749.3		366 -999	691.8666667	-999 235.5296982 556.4666	545	
18	1964	798.6666667	1062.5	614.9	-999	909.9333333	-999 458.0847118 953.2333	374	
19	1965	561.8	875.2	672.5	333333 -999	720.5	-999 297.63282	-999	

Les fichiers d'entrée peuvent être obtenus avec le logiciel Excel

en le sauvegardant au format :

"texte (separateur : tabulation".



Nom de fichier:	navigateur
Nom de fichier:	navigateur
Nom de fichier:	Nom do fishion
	Nom de nomer:
Donnée initiale: N/A	Donnée initiale: N/A
Dernière donnée: N/A	Dernière donnée: N/A
Commencer à: 1 ≑	Commencer à: 1 🗧
Nombre de champs 0	Nombre de champs 0
Nombre de décalages 0	Nombre de décalages 0
Nombre de variables 0	Nombre de variables 0
Nombre utilisé 0	Nombre utilisé 0
	Donnée initiale: N/A Dernière donnée: N/A Commencer à: 1 Nombre de champs 0 Nombre de décalages 0 Nombre de variables 0 Nombre utilisé 0

Pour sélectionner un fichier d'entrée il faut cliquer sur navigateur.



Open X Input File				
Librar	ies 🕨 Documents 🕨	 ✓ 49 Search Documents 		
om de fi Organize 👻 New f	older		i≡ • 🔟 📀	
ernière (🚖 Favorites ommenc 📃 Desktop	Documents library Includes: 2 locations	Arran	ge by: Folder ▼ #1	
ombre d 🛛 🚴 Downloads	Name	Date modified	Туре	
ombre d 🤤 Recent Places	Bluetooth Exchange Folder	30/10/2012 18:16	File folder	
ombre u 🔤 Libraries	= figure01	28/11/2012 15:24	JPEG image	
	sstavril	29/11/2012 15:13	JPEG image	
A Music	sstavril_cor	29/11/2012 15:13	Text Document 😑	
lle de la	sstjul	29/11/2012 15:07	JPEG image	
Videos	📄 sstjul_cor	29/11/2012 15:07	Text Document	
grès:	📄 sstjun	29/11/2012 15:08	JPEG image	
tions : Computer	sstjun_cor	29/11/2012 15:09	Text Document	
📇 OS (C:)	属 sstmay	29/11/2012 15:11	JPEG image	
CD Drive (E:) e-e:	• • •	00 MA (004 0 4 0 4 0	- · · · · ·	
Fi	le name:	▼ X Input Files (*,*) 👻	

CPT ouvre un navigateur genre 'windows explorer' qui permet de localiser et de charger le fichier de données.



ichier Editer Actions Outils Op	ions Montrer Aide				
Variables Explicat navigateur Nom de fichier: Donnée initiale: N/A Dernière donnée: N/A Commencer à: 1 +	X Domain Selection Please specify domain lim (Limites dados aproximat	Fichiers d'entrée Variable Réponse (navigateur de fichier:	0 N	Variables I naviga om de fichier:	Prévues — teur
Nombre de décalages Nombre de variables Nombre utilisé Taille de la période d'apprentis Progrès: 0% Actions:	Ext (Southern latitudes and we Northernmost latitude (8 Southernmost latitude (- Westernmost longitude (Easternmost longitude (tension X: estern longitudes negatives) t8): 90 e80: -90 (0): 0 358): 360		X Domain	2
	<u></u> k	(<u>D</u> ata Limits	Redraw Map	e isse issow issow asow asow	

Pour des données en grille ou station, CPT te propose de choisir ton domaine spatial sur lequel s'effectuera l'analyse PCR/CCA ou le domaine d'affichage des résultats (Y). On peut utiliser la souris ou taper directement les coordonnées limites

du domaine.

RI

Fichier Editer Actions Outils Option	s Montrer Aide	Fichiers d'entrée	
navigateur Nom de fichier: ERSST_Jul5012; Donnée initiale: 16 Jul 1950 Dernière donnée: 16 Jul 2012 Commencer à: 1950 ÷ Nombre de champs Nombre de décalages Nombre de points de	glo_oldfrmt.tsv 1 1 5580	navigateur Nor Nom de fichier: Na Donnée initiale: N/A Dernière donnée: N/A Commencer à: 1 1 ÷ Nombre de champs 0 Nombre de décalages 0 X Mode Options	om de fichier: onnée initiale: ernière donnée ommencer à: ombre de cham ombre de décal
Taille de la période d'apprentissag Progrès: 0% Actions:	0 e: 0 +	Nombre de modes (X) Nombre maximum de modes 5 Nombre minimum de modes 1 QK Avancée Annuler	de prévis

Ensuite il faut choisir le nombre de modes à utiliser dans l'analyse. CPT va chercher la combinaison pour trouver le nombre optimal de modes.





Pour des fichiers de type station, CPT utilise par défaut les coordonnées dans le fichier pour délimiter la zone néanmoins il y`a une option de redéfinir le domaine.



CHOISIR LA PERIODE D'APPRENTISSAGE

navigateur	
in an	navigateur
Ltsv Nom de fichier: DebitOMVS_ASO5012_txt	Nom de fichier: ERSST_Jul5012glo_oldfrmt.tsv
Donnée initiale: 1962	Donnée initiale: 16 Jul 1950
Dernière donnée: 2012	Dernière donnée: 16 Jul 2012
Commencer a 1962 ≑	Commencer à: 2012 ≑
Nombre de champs 1	Nombre de champs 1
Nombre de décalages 1	Nombre de décalages 1
Nombre de stations 8	Nombre de points de 5580
Nombre utilisé 0	Nombre utilisé 0
	Nom de fichier: DebitOMVS_ASO5012_txt Donnée initiale: 1962 Dernière donnée: 2012 Commencer à: 1962 ‡ Nombre de champs 1 Nombre de décalages 1 Nombre de stations 8 Nombre utilisé 0

Par défaut CPT commence l'analyse à partir de la première année dans les fichiers X et Y; même si ces années sont différentes.

On devrait normalement choisir la même année en général l'année la plus récente entre les deux fichiers.

Faire attention lorsque la saison est à cheval sur deux années comme par exemple les saisons DJF ou JFM. Dans ce cas l'année de début du fichier X doit être décalée d'une année par rapport à celle du fichier Y.



CHOISIR LA PERIODE D'APPRENTISSAGE

	Variable Réponse (Y)	Variables Prévues
navigateur Nom de fichier: ERSST_Jul5012glo_oldfrmt.tsv Donnée initiale: 16 Jul 1950 Dernière donnée: 16 Jul 2012 Commencer à: 1952	navigateur Nom de fichier: DebitOMVS_ASO5012txt Donnée initiale: 1962 Dernière donnée: 2012 Commencer à: 1962 -	navigateur Nom de fichier: ERSST_Jul5012glo_oldfrmt.ts Donnée initiale: 16 Jul 1950 Dernière donnée: 16 Jul 2012 Commencer à: 2012 ▲
Nombre de champs 1 Nombre de décalages 1 Nombre de points de 5580 Nombre utilisé 0	Nombre de champs 1 Nombre de décalages 1 Nombre de stations 8 Nombre utilisé 0	Nombre de champs 1 Nombre de décalages 1 Nombre de points de 5580 Nombre utilisé 0

La taille de la période d'apprentissage (longueur) est spécifiée en mettant le nombre d'année (49) ainsi que la largeur de la fenêtre de la validation croisée (5) où se fera la validation du modèle.



DONNEE MANQUANTES - MISSING VALUES

nier Editer Actions Outlis	ptions Montrer Alde	
Variables Expl navigat Nom de fichier: ERSST_, Donnée initiale: 16 Jul 19 Dernière donnée: 16 Jul 20 Commencer à: 1962 ‡ Nombre de champs Nombre de décalages Nombre de points de Nombre utilisé	Options des Modes X Options des Modes Y Options de l'ACC Options de GCM Goodness Index Période Climatologique Personnalisé Configuration de Reéchantillonnage Verification Settings Configuration de Prévision	Fichiers d'entrée lariable Réponse (Y) navigateur DebitOMVS_ASO5012_txt 1962 2012 1962 2012 1962 3 nps 1 lages 1 ons 8 0
aille de la période d'appre	Données	Transformer les Données Y Borné à Zéro
	Graphique	Sort Ensemble Members
rogrès: 0%		Valeurs Manquantes

Si on a des données manquantes dans le fichier, CPT offre un menu de remplissage avec plusieurs options : Options => Données => Valeurs Manquantes



DONNEE MANQUANTES - MISSING VALUES

Nombre de champs		1	Nombre de champs	
rs Manquantes			Manton in Analogue	
	atives (X): —	_	Variable Réponse (Y):
Missing value flag:	-999	*	Missing value flag:	-999
Maximum % of missing values:	10	-	Maximum % of missing values:	10
Maximum % of missing points de gr	ille:	10	Maximum % of missing stations:	10
Number of near-neighbours:	1	-	Number of near-neighbours:	1
Missing Value Replacement Select method: • Long-term means • Long-term medians • Random numbers • Best near-neighbours			Missing Value Replacement Select method: • Long-term means • Long-term medians • Random numbers • Best near-neighbours	
		<u>0</u> K	Annuler	

Il faut d'abord indiquer le code/chiffre qui représente une donnée manquante (Missing value flag), ainsi on peut :

limiter le pourcentage maximum de donnée manquante (Maximum % of missing value) au delà duquel la station/grille est enlevée de l'analyse
 Aussi pour chaque année on peut spécifier le pourcentage maximum (Maximum % of missing points de grille) de station manquante au delà duquel toute l'année est enlevée de l'analyse.

Pour remplacer une donnée manquante on peut limiter le nombre de station proche à utiliser (Number of near-neighbours)



DONNEE MANQUANTES - MISSING VALUES

			de la	Nombre de champs	a	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
anquantes				Manhoo do dischaper		X
	tives (X): —	_		Variable Réponse (Y):	
sing value flag:	-999	-		Missing value flag:	-999	÷
ximum % of missing values:	10	-		Maximum % of missing values:	10	÷
ximum % of missing points de gril	lle:	10	÷	Maximum % of missing stations:	10	÷
mber of near-neighbours:	1	-		Number of near-neighbours:	1	-
Aissing Value Replacement: elect method Long-term means Long-term medians Random numbers Best near-neighbours		QK	Annu	Missing Value Replacement: Select method: • Long-term means • Long-term medians • Random numbers • Best near-neighbours		
	Anquantes Variables Explicat sing value flag: ximum % of missing values: ximum % of missing points de grit mber of near-neighbours: Aissing Value Replacement: elect method Long-term means Long-term medians Random numbers Best near-neighbours	Variables Explicatives (X): ssing value flag: -999 ximum % of missing values: 10 ximum % of missing points de grille: mber of near-neighbours: 1 Aissing Value Replacement: elect method Long-term means Long-term medians Random numbers Best near-neighbours	Variables Explicatives (X): ssing value flag: symmetry of missing values: 10 ximum % of missing points de grille: 10 mber of near-neighbours: 1 Aissing Value Replacement: elect method Long-term means Long-term medians Best near-neighbours	Image: Second state Variables Explicatives (X): Second state ssing value flag: Second state 999 ÷ ximum % of missing values: 10 ÷ 10 ÷ ximum % of missing points de grille: 10 ÷ 10 ÷ mber of near-neighbours: 1 ÷ 10 ÷ Aissing Value Replacement: 1 elect method Long-term means Long-term medians Best near-neighbours Best near-neighbours OK	Variables Explicatives (X): Variable Réponse (Y ssing value flag: 999 ximum % of missing values: 10 ximum % of missing points de grille: 10 mber of near-neighbours: 1 Missing Value Replacement: 1 elect method Long-term medians Long-term medians Random numbers Best near-neighbours Best near-neighbours Image: Maximum % of missing values: Missing Value Replacement: Best near-neighbours Best near-neighbours	Variables Explicatives (X): Variable Réponse (Y): ssing value flag: 999 ximum % of missing values: 10 10 + ximum % of missing points de grille: 10 mber of near-neighbours: 1 1 + Missing Value Replacement: 0 Long-term means 1 Long-term medians 0 Random numbers - Best near-neighbours - QK Annuler

Pour remplacer une donnée manquante on a le choix d'utiliser :

- Ia moyenne des valeurs existantes (Long-term mean)
- Ia médiane des valeurs existantes (Long-term medians)
- des données aléatoires qui ne changent pas le caractère de la série (Random numbers)

par la valeur de la station la plus proche de la même année (Best near-neighbours)



SAUVEGARDE DES CHOIX DE L'ANALYSE

Vouvea Duvrir	(Ctrl+N) (Ctrl+O)	ves (X)	Fichiers d'entrée Variable Réponse (Y)	Variables Prévues
auver	(Ctrl+S)		navigateur	navigateur
auver Comme		5012glo_oldfrmt.tsv	Nom de fichier: DebitOMVS_ASO5012.txt Donnée initiale: 1950	Nom de fichier: ERSST_May5012glo_oldfrmt.ts Donnée initiale: 16 Mai 1950
ermer X			Dernière donnée: 2012	Dernière donnée: 16 Mai 2012
ermer Y			Commencer à: 1962 ≑	Commencer à: 2011 ≑
ermer Forecast		1	Nombre de champs 1	Nombre de champs 1
) ácultate		1	Nombre de décalages 1	Nombre de décalages 1
resultats		6480	Nombre de stations 8	Nombre de points de grille 6480
ortie	(Ctrl+Q)	4774	Nombre utilisé 5	Nombre utilisé 0
		_	1844	

Une fois que les fichiers d`entrée sélectionnés de même que les choix sur l'analyse il est recommandé de sauvegarder ces choix dans un projet pour les appeler plus tard avec le menu : Fichier => Sauver



LANCER L'ANALYSE

Nom de fichier: ERSST_Apr5012gro_ordination		navigateur
Donnée initiale: 16 Avr 1950 Dernière donnée: 16 Avr 2012 Commencer à: 1962 Iombre de champs 1 Iombre de décalages 1	Nombre de champs 1 Nombre de décalages 1	Nom de fichier: ERSST_Apr5012glo_oldfrmt.tsv Donnée initiale: 16 Avr 1950 Dernière donnée: 16 Avr 2012 Commencer à: 2011 Nombre de champs 1 Nombre de décalages 1
Nombre de points de grille 6480 Nombre utilisé 0	Nombre de stations 8 Nombre utilisé 0	Nombre de points de grille 6480 Nombre utilisé 0

Maintenant on peut commencer l'analyse en allant au menu : Actions => Calculer => Validation Croisée



L'ANALYSE

Progrès: 1 Actions:	00%		
Beginning analysis . Optimizing cross-val: Période d'apprentissa	 idated performan age: 1962 à 2010	ce	
CURREI	чт	OPTIMUM	
Number of Modes	5 Goodness Index	Number of Modes	Goodness Index
1 2 3	-0.200 -0.108 0.161	1 2 3	-0.200 -0.108 0.161
4 5	0.218 0.190	4 4	0.218 0.218
Identifying categoric Fait!	sing full traini ?5	ng period (1962 a 20	טרע)

CPT commence l'analyse et on peut voir les étapes et les procédures d'optimisation du modèle.



L'ANALYSE

	(URRENT		OPTIMUM	
Number	of	Modes	Goodness Index	Number of Modes	Goodness Index
	1		-0.200	1	-0.200
	2		-0.108	2	-0.108
	3		0.161	3	0.161
	- 4		0.218	4	0.218
	5		0.190	4	0.218
Constructing Identifying (Fait!	mod ate	lel using egories .	full training 	g period (1962 à 20	010)

Optimisation du modèle :

1. CPT utilise d'abord 1 seule composante principale (mode 1) comme prédicteur pour faire une validation croisée et ensuite il calcule un indice de bonne prévision "goodness index" indiquant la qualité de la prévision en général (1 étant la meilleure). Ensuite CPT utilise les modes 1 et 2 pour faire la prévision et calcule à nouveau le goodness index, et ainsi de suite jusqu'à utiliser le maximum de mode (5 modes).

2. A chaque étape CPT compare les goodness indices et retient dans la colonne "OPTIMUM" le meilleur modèle cad celui qui a l'indice le plus élevé. (voir l'exemple ci-dessus avec #4).



RESULTATS

Váriábi Vérification Rétroactive Bootstrap Graphique de Dispersion Iom de fichier: E Nom de fichier: Debit0 Debit0 Cartes de Skill-scores Ide fichier: ERSST_Apr5012glo_oldfrmt.tsv Nom de fichier: 1 Nom de fichier: Debit0 Dernière donnée: 16 Avr 1950 Dernière donnée: 1 Nombre de champs 1 Nombre de champs 1 Nombre de champs 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de points de grille 6480 Nombre de stations 8 Nombre de grille 6480 Nombre utilisé 4774 Données d'apprentissage 5 Données d'apprentissage	Variation Rétroactive Bootstrap Image: Commencer à: Table de contingence Nom de fichier: Deprive de Dispersion Nom de fichier: Deprive donnée: 1 Donnée initiale: 1950 Dernière donnée: 1 Commencer à: 1962 ÷ Nombre de champs 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de stations 8 Nombre utilisé 4774 Données d'apprentissage 3 ÷ aille de la période d'apprentissage: 49 ÷	Mastahl	Validation •	Validation Croisée	Mesures de Performance	Marticles Defense
Image: Control of the contingence Image: Control of the contingence Image: Control of the	Image: State de contingence Image: State de content de conte	Variabi	Vérification	Rétroactive	Bootstrap	navigateur
Dernière donnée: 1 Prévisions Commencer à: Prévisions Nombre de champs 1 Nombre de champs 1 Nombre de décalages 1 Nombre de points de grille 6480 Nombre utilisé 4774 Données d'apprentissage 0	Dernière donnée: 1 Prévisions Prévisions Commencer à: 1 Nombre de champs 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de points de grille 6480 Nombre de stations 8 Nombre de stations 8 Nombre de la période d'apprentissage: 49 Dernière donnée: 2012 Commencer à: 1962 ÷ Dernière donnée: 16 Avr 2012 Commencer à: 2011 ÷ Dernière donnée: 16 Avr 2012 Commencer à: 2011 ÷ Nombre de champs 1 Nombre de décalages 1 Nombre de stations Nombre de la période d'apprentissage: Données d'apprentissage 1 Nombre de la fenêtre de validation croisée: 3 Nombre de prévisions: 2	lom de fichier: E)onnée initiale: 1	Modes	Nom de fichier: DebitC Donnée initiale: 1950	Graphique de Dispersion Cartes de Skill-scores	de fichier: ERSST_Apr5012glo_oldfrmt.tsv
Nombre de champs 1 Nombre de champs 1 Nombre de champs 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de points de grille 6480 Nombre de stations 8 Nombre de points de grille 6480 Nombre utilisé 4774 Nombre utilisé 5 Nombre utilisé 0	Nombre de champs 1 Nombre de champs 1 Nombre de champs 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de points de grille 6480 Nombre de stations 8 Nombre de points de grille 6480 Nombre utilisé 4774 Nombre utilisé 5 Nombre de points de grille 6480 Nombre utilisé 5 Données d'apprentissage 0 0 aille de la période d'apprentissage: 49 + Largeur de la fenêtre de validation croisée: 3 + Nombre de prévisions: 2 +)ernière donnée: 1 Commencer à:	Prévisions	Dernière donnée: 2012 Commencer à: 1962		Dernière donnée: 16 Avr 2012 Commencer à: 2011 🖨
Nombre de décalages 1 Nombre de décalages 1 Nombre de décalages 1 Nombre de points de grille 6480 Nombre de stations 8 Nombre de points de grille 6480 Nombre utilisé 4774 Nombre utilisé 5 Nombre utilisé 0	Nombre de décalages 1 Nombre de décalages 1 Nombre de points de grille 6480 Nombre de stations 8 Nombre utilisé 4774 Nombre utilisé 5 Données d'apprentissage 1 Nombre de prévisions: 2 aille de la période d'apprentissage: 49 4 Largeur de la fenêtre de validation croisée: 3 Nombre de prévisions: 2	lombre de champs	1	Nombre de champs	1	Nombre de champs 1
Nombre utilisé 4774 Données d'apprentissage	Nombre utilisé 4774 Nombre utilisé 5 Nombre utilisé 6480 Nombre utilisé 5 Nombre utilisé 0 Données d'apprentissage 1 argeur de la fenêtre de validation croisée: 3 Nombre de prévisions: 2	lombre de décalages	1	Nombre de décalages	1	Nombre de décalages 1
Données d'apprentissage	Données d'apprentissage aille de la période d'apprentissage: 49 💠 Largeur de la fenêtre de validation croisée: 3 🐥 Nombre de prévisions: 2 🐥	lombre utilisé	4774	Nombre utilisé	° 5	Nombre utilisé 0
aille de la période d'apprentissage: 10 🔺 👘 Largeur de la fenêtre de validation croisée: 2 🔺 Nombre de prévisions: 2 🔺		aille de la nériode d'a	upprontiesago: 10	Données d'app	rentissage	Nombre de prévisions:

Pour voir les résultats il faut aller au menu Outils => Validation => Validation Croisée :

 Mesures de Performance : montre les séries prévue et observée sur chaque station et une mesure de performance (corrélation ...).
 Cartes de Skill-scores : montre la performance des modèles sur

l'ensemble des stations.



INDICATEURS D'INCERTITUDE

	Validation	• V	alidation Croisée	Mesures	de Performa	nce		
– Variabl	Vérification	▶ R	étroactive	 Bootstra 	р			Variables Prévu
	Table de contingence	•	1	Graphiq	ue de Dispers	ion	11 1001010	navigateur
nier: E	Modes	> Nor	n de fichier: Del	oitC Cartes d	e Skill-scores		de fichier:	ERSST_Apr50
iale: 1	Cartes Climatologique	Doi	inee initiale: 195	0			hee initiale:	16 Avr 1950
onnee: 1	Prévisions	> Der	niere donnee: 201	2		Deri	niere donne	e: 16 Avr 2012
			innencera: 19	62 👻	702	Con	intencer a:	2011 -
champs	1	Nor	nbre de champs		1	Non	nbre de char	nps
🛄 Cross-vali	dated scores							
Station: Station	: Dualia 13.60N, 10.3	38W	— Sample: ——		Confidence	limits:		- P-value: —
Station: Statior Contin Cor Sco & C Rap Bia	t Dualia 13.60N, 10.5 Score:	38W	Sample: 8.4455 8.4744 66.17 19.84 8.2553 -4.97 -4.97	Confidence 0 5 0	Confidence 1euel: 9 .2289 to .2066 to 6.87% to 5.24% to 5.24% to 1294 to 60.09 to	limits:	6.69	• P-value:
Station: Statior Contin Cor Sco Rap Bia RCC Err	t Dualia 13.60N, 10.3 Score: vous measures: rélation de Pearson rélation de Spearman re 2AFC e variance port de variance is moyen EM eur absolue moyenne	38₩	Sample: 8.4455 8.4744 66.17 19.84 8.2553 -4.97 195.95 154.93	Confidence 9 0 5 - 1 1	Confidence 1evel: 9 .2289 to .2066 to 6.87% to 5.24% to .1294 to 60.09 to 49.63 to 21.89 to	limits: 5.000% 9.0000 0.0000 0.00% 0.00% 0.00%	0.00 0.00 0.00 0.00	• P-value: • . 002 0 • . 0000 • . 002 0 • . 004 0 1 . 0000 N/A N/A
Station: Statior Contin Cor Sco & C Rap Bia RCQ Err Catego	t Dualia 13.60N, 10.3 Score:	38W		Confidence Ø Ø 5 - 1 1	Confidence 1evel: 9 .2289 to .2066 to 6.87% to 5.24% to .1294 to 60.69 to 49.63 to 21.89 to	limits:	0.00 0.90 0.90 0.90	- P-value: 9.0020 9.0090 9.0040 1.0000 N/A N/A 9.0260
Station: Statior Contin Cor Cor Sci Rap Bia RCQ Err Catego Sci Ski	t Dualia 13.60N, 10.3 Score: vous measures: rélation de Pearson rélation de Spearman re 2AFC e variance port de variance is moyen EM eur absolue moyenne rical measures: re de réussite 11 score de réussite	38W		Confidence Ø Ø 5 - 1 1 3 3	Confidence 1evel: 9 .2289 to .2066 to 6.87% to 5.24% to .1294 to 60.09 to 49.63 to 21.89 to 1.82% to 4.55% to	limits:	0.00 0.90 0.00 0.00	- P-value: 9.0920 9.0900 9.0920 9.0940 1.0990 1.0990 N/A N/A 9.0260 9.0260
Station: Statior Contin Cor Cor Sco % co Rap Bia RCC Err Catego Ski LEF	t Dualia 13.60N, 10.3 Score:	38W	Sample: 8.4455 8.4744 66.17 19.84 8.2553 -4.97 195.95 154.93 47.73 21.59 38.15	Confidence 9 0 5 - 1 1 1 3 - 1 1	Confidence 1evel: 9 .2289 to .2066 to 6.87% to 5.24% to .1294 to 60.09 to 49.63 to 21.89 to 1.82% to 4.55% to 1.45% to	limits:	0.00 0.00 0.00 0.00	 P-value: 0.0020 0.0000 0.0020 0.0040 1.0000 1.0000 N/A N/A 0.0260 0.0000
Station: Statior Contin Cor Cor Scc % c Rap Bia RCC Err Catego Scc Ski LEF Scc	t Oualia 13.60N, 10.3 Score: uous measures: rélation de Pearson rélation de Spearman re 2AFC e variance port de variance is moyen EM eur absolue moyenne rical measures: re de réussite 11 score de réussite S re de Gerrity	38W	Sample: 8.4455 8.4744 66.17 19.84 8.2553 -4.97 195.95 154.93 47.73 21.59 38.15 28.46	Confidence 0 9 - 1 1 3 - 1	Confidence 1evel: 9 .2289 to .2866 to 5.24% to 5.24% to .1294 to 49.63 to 21.89 to 1.82% to 4.55% to 1.45% to 9.88% to	limits:	0.99 0.99 0.99 0.99	 P-value: 0.0020 0.0020 0.0020 0.0040 1.0000 N/A N/A 0.0260 0.0260 0.0000 0.0020
Station: Station Contin Cor Scc % C Rap Bid RCC Err Catego Scc Ski LEF Scc 200	: Dualia 13.60N, 10.3 Score: vous measures: rélation de Pearson rélation de Spearman re 2AFC e variance port de variance is moyen EM eur absolue moyenne rical measures: re de réussite ll score de réussite S re de Gerrity C (forecast categories (corrigente)	38W		Confidence 0 9 - 1 1 1 3 - 1 1 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	Confidence 1evel: 9 .2289 to .2866 to 6.87% to 5.24% to .1294 to 49.63 to 21.89 to 1.82% to 4.55% to 1.45% to 9.88% to 8.38% to	limits:	0.99 0.99 0.99 0.99	- P-value: 0.6026 0.6090 0.6020 0.6040 1.6090 1.6090 N/A N/A 0.0266 0.6026 0.6060 0.6090 0.6090 0.6090 0.6090 0.6090
Station: Station Contir Cor Scc & C Rap Bia RCC Err Catego Scc Ski LEF Scc Scc 206 Pir	: Dualia 13.60N, 10.3 Score: vous measures: rélation de Pearson rélation de Spearman re 2AFC e variance port de variance is moyen EM eur absolue moyenne rical measures: re de réussite 11 score de réussite S re de Gerrity C (forecast categories C (continuous forecast e de ROC (au-descous c	38W	Sample: 0.44455 0.4744 66.17 19.84 0.2553 -4.97 195.95 154.93 47.73 21.59 30.15 28.46 69.53 67.13 8.659	Confidence 0 9 - 1 1 1 1 3 - 1 5 6 8	Confidence 1evel: 9 .2289 to .2866 to 6.87% to 5.24% to .1294 to 60.09 to 49.63 to 21.89 to 1.82% to 4.55% to 1.45% to 9.88% to 8.38% to 2.93% to .4429 to	limits:	0.99 9.99 9.99 9.99	- P-value: 0.0020 0.0020 0.0020 0.0040 1.0000 1.0000 N/A N/A N/A 0.0260 0.0260 0.0020 0.0020 0.0020 0.0100 0.0100 0.0100

Pour avoir des indications d'incertitude sur la performance des mesures (corrélation) il faut aller à :

Outils => Validation => Validation Croisée => Bootstrap On a ainsi des limites de confiance (Confidence limites) et la probabilité due au hasard (P-value).

RESULTATS



Un clique droit sur la carte permet d'afficher l'intervalle de confiance sur la prévision : Personnaliser => Intervalles de Prévisions



Pour voir il faut aller à : Outils => Prévisions => Dépassements : permet de voir la probabilité de dépasser une certaine valeur (cote) pour la climatologie (en rouge) et pour la prévision.

RESULTATS : CARTES PREVUES



Outils => Prévisions => Cartes => Probabilités :

montre la probabilité de chaque catégorie à chaque point ainsi que la distribution spatiale de la prévision.

Sur cet exemple il est clair que la catégorie au dessus de la normale (Above) est la plus probable.





RESULTATS : dans un fichier

Sauver	(Ctrl+O) (Ctrl+S)	ves (X)		Hindcasts Données d'Entrée Climatologies Modes PCR Results Prévisions
Sauver Comme Fermer X Fermer Y		012glo_oldfrmt.tsv	Nom de fic Donnée ini Dernière de Commence	Cross-Validated Predictions:
Fermer Forecast Résultats	*	1 Sauver Goodness In	ndex	
Sortie	(Ctrl+Q)	Résultats		•

- CPT offre l'option de sauvegarder les résultats dans un fichier avec le menu Fichier => Résultats => Résultats :
- ✓ Les valeurs prévues par la validation croisée Hindcasts
- Les données d'entree après avoir rempli les valeurs manquantes Données d'Entree
- ✓ Les moyennes et les seuils Climatologies
- ✓ Les composantes principales PCR Results
- ✓ Les prévisions du modèle Prévisions



CONCLUSIONS

• Pour plus de détails, lire la page d'aide sur chaque menu et sur chaque option.

 Souscrire à la liste des utilisateurs pour être notifié aux mise à jour :

http://iri.columbia.edu/outreach/software/

 Nous aimerions vous entendre. Vos commentaires et vos questions nous aide a améliorer le CPT donc n'hésiter pas á nous écrire á :

cpt@iri.columbia.edu

