Milestone M5.2: Prototype of an automatic system for forecast quality assessment of seasonal-to-decadal hindcasts

Introduction

Even more than in weather forecasts, the skill of seasonal forecasts depends very strongly on the geographical location and season of the year. Also, different forecast systems have strong and weak points in various regions and seasons. For a user of seasonal forecasts it is therefore very important to know what skill the different systems have achieved in past forecasts and hindcasts in order to be able to judge the quality of current forecasts for his/her region.

Verification of seasonal forecasts is available from the various centres (e.g., ECMWF, NCEP, IRI, UKMO) and research experiments (e.g., Demeter). However, these are static pictures, not complete (for all months and lead times) and not directly comparable due to the use of different verification measures and colour scales.

We introduce here the first web site that allows for the dynamic generation of skill score maps and diagrams from a variety of seasonal forecast models using different skill scores.

Description

Withing the ENSEMBLES project we have constructed a web site that lets the user generate verification plots in real time. The site is part of the KNMI Climate Explorer (<u>climexp.knmi.nl</u>). It consists of three parts

- 1. A full set of monthly data of the hindcasts of the ECMWF S2, NCEP CFS and IRI ECHAM4.5 forecasts systems, plus the Demeter research experiment.
- 2. A large set of deterministic and probabilistic verification routines: correlation, RMS, MAE, ROC, RPS, RPSS (with respect to climatology), Brier Score and decomposition, Brier Skill Score (with respect to climatology).
- 3. A set of web pages that allows anybody to generate verification curves and maps from 1. and 2.

In the following examples are shown of the current set-up. Planned extensions are

- 1. More data from other operational centres and the ENSEMBLES data servers
- 2. A connection to the public ECMWF seasonal to decadal ENSEMBLES server to allow the user to select data there and seamlessly verify at KNMI. This will allow for the first time verification studies of statistics of daily data, for instance the numbers of windstorms, high precipitation events, extreme snowfall.
- 3. More verification measurees, in particular skill scores such as the RPSS and BSS with respect to more sophisticated models than climatology and persistence: damped persistance, optimal normal correlactions and regression to Nino3.4.
- 4. A more user-friendly user-interface consisting of a single web page

rather than the current multi-step procedure. We envisage these extensions to be complete by summer 2006.

Example

On the web site climexp.knmi.nl, the data sets available for verification are under `seasonal forecast ensembles':

About KNMI Colloquia/Workshops Staff Publications Contact Links	Research ⊨ Data Center ⊨	About	KNMI	Colloquia	Worksho	ops St	aff Pu	ublicatio	ons C	ontaci	: Links	Dat
Home → dimexp → start		Home +	dimexp	→ selectie	ld seaens							Dat
Climate Explorer Starting point Welcome, Geert Jan van Oldenborgh from KNMI This paag care be added to vour bookmarketevortes to avoid logang in	Introduction, results News Examples Publications Effects of ENSO on the weather	Climate Explorer Select a monthly field Seasonal forecasts full ensembles									Introduction, results News Examples Publications Effects of ENSO on the wealth	
	Probability of tropical cyclones	Cal		ermble								Probability of tropical cyclon
Start by selecting a class of climate data from the right-hand menu. After you have selected the time series or fields of interest, you will be able to investigate it, correlate it to other data.	Select a time series								button			Select a time series
and generate derived data from it.	Daily station data	DEME		éo CER-		INGV	ECMW	= MPI	UKMO	all	3	Daily station data
If you are new it may be helpful to study the examples.	Daily climate indices Pentad climate indices			ICE FACS								Daily climate indices Pentad climate indices
Share and enjoy!	Monthly station data	var	start 1958 2001		1974- 2001	1987- 2001	1958- 2001	1969- 2001	1959- 2001		1959- 2001	Monthly station data
	Monthly climate indices Annual climate indices	t2m :	eb 🧹		~	~	0	~	~	0	0	Monthly climate indices Annual climate indices
	User-defined time series			. 0	0				- C	. 0	0	User-defined time series
corr Jan-Feb averaged NAO-Gibraltar index with Apr CRU 0.5o temperature 1901-2000	Upload your own time series	t2m i	nay 🤇	0 0	0	0	0	0	0	0	0	Upload your own time serie
60X	Select a field	t2m :	wg 🤇	0 0	0	0	0	0	0	0	0	Select a field
641	6-hourly fields	t2m i	10V C	0	0	0	0	0	0	0	0	6-hourly fields
638	Daily fields Monthly observations	DICD	ah d									Daily fields Monthly observations
eon	reanalistic fields			. 0	0				- U	. 0	0	reanalysis fields
57N - 🕎	seasonal forecast means seasonal forecast ensembles	prop i	nay (0 0	0	0	0	0	0	0	0	seasonal forecast mean
SHN C 2 C	scenario runs	prop a	wg 🤇	0 0	0	0	0	0	0	0	0	seasonal forecast enserr acenario runs
5IN SIN	Does fellord	prcp i	nov e	0	0	0	0	0	0	0	0	User-defined
488.	Upload your own field	slp	eb 🤇									Upload your own field
	Feedback		en C	0	С	0	0	0	0	0	0	Feedback
	Geert Jan's home page	slp	nay 🦿	0 0	0	0	0	0	0	0	0	Geert Jan's home page
TÔM Ô TÚC 20C 3ÔC 4ÔC SỐC ĐỘC TỐC		slp	aug 🤇	0 0	0	0	0	0	0	0	0	
		slp	10V C	0	0	0	0	0	0	0	0	
-3.6 -0.5 -0.4 -0.3 -0.2 0.2 0.3 0.4 0.5 0.6												

One selects a field from the collection of Demeter, ECMWF, NCEP and IRI data, in this case the 1Feb T2m forecasts from the ECMWF System-2 operational seasonal forecast system was selected (only visible when one scrolls down the page). The next page gives a list of options to investigate this field, among them is `verify against observation'. This brings up the main verification form, with choices for the dataset to verify against (only relevant ones are shown), the verification measure, the threshold, season and plot options:

Sitemap Searc	ph Q	KNMI Sitemap Search	
About KNMI Colleguia/Workshops Staff Publications Contact Link	ks Research → Data Center →	About KNMI Collequia/Workshops Staff Publications Contact Links	Res Data C
			Date of
lome → climexp → select		Home → climexp → regionverificationform	
Dimate Explorer	Introduction, results	Climate Explorer	Introduction, results
leid	News	Field verification	News
insemble ECMWF-2 1feb T2m	Examples	ensemble ECMWF-2 1 feb T2m	Examples
	Publications		Publications
ound ensemble members 0 to 39 cmwl2_t2m_feb_00.dat	Effects of ENSO on the weather Probability of tropical cyclones	Converting ensemble ECMWF-2 1feb T2m from K to Celsius	Effects of ENSO on the weather Probability of tropical cyclones
axis: whole world in 240 1.50° steps. first point at 0.00° E	Select a time series	Verifying Temperature field	Select a time series
f axis: regular grid with 121 1.50° steps. first point at 90.00° S	Daily station data	Terroretaria	Daily station data
f at 1.00	Daily climate indices	Temperature C 1856-now anomalies: HadCRUT2 (Jones & Parker T2m/SST analysis)	Daily climate indices
Monthly data available from Feb1987 to Jul2005 (222 months)	Pentad climate indices	1857-now variance adjusted. C HadCBUT1, C HadCBUT1v	Pentad climate indices
Variable t2m (** surface 2 metre temperature) in K. defined at Z level 0	Monthly station data		Monthly station data
	Monthly climate indices	Land C 1851-now anomalies: CRUTEM2 (Jones T2m analysis) C number of stations,	Monthly climate indices
Extract timeseries	Annual climate indices	C variance adjusted, C CRUTEM1, C nr, C CRUTEM1v.	Annual climate indices
atitude: -90 °N 90 °N (leave second field blank for one point)	User-defined time series	V variance adjusted, ≦ CRUTEM1, S nr, S CRUTEM1v.	User-defined time series
	Upload your own time series	1901-2000: CRU2 New 0.5° analysis(land only) C Old World, C New World	Upload your own time series
ongitude: 0 °E-360 °E		1901-2000: <u>CRU2 New 0.5° analysis</u> (land only) [*] Old World, [*] New World (LARGE)	
ocundaries halfway grid points	Select a field	A	Select a field
	6-hourly fields	Air C 1800-1997: COADS Tair	6-hourly fields
make C average C set of grid points	Daily fields	Temperature 1800-1997: COADS Tair	Daily fields
units Convert to Celsius Cleave in K	Monthly observations	T2n 0 0	Monthly observations
	reanalysis fields	T2m C 1.5° C 2.5° ERA-40	reanalysis fields
Make time series	seasonal forecast means	<u>^</u>	seasonal forecast means
	seasonal forecast ensembles scenario runs	C NCEP/NCAR	seasonal forecast ensembles
Create a field with derived data	User-defined		scenario runs
	User-defined Upload your own field		User-defined
lew time scale: yearly	opidad your own lield	Map verification measures	Upload your own field
Vew variable: number of t2m11eb T2m	Investigate this field	C Correlation of the ensemble mean	Investigate this field
Threshold: greater than 25 K	Plot difference with a field		Plot difference with a field
	Compute mean and higher moments	Root mean square error (RMSE) of the ensemble mean	Compute mean and higher mome
Make new field	Correlate with a time series	C Mean absolute error (MAE) of the ensemble mean	Correlate with a time series
	Pointwise correlations with a field	O Brier score (O alternative)	Pointwise correlations with a field
Apply yearly high/low-pass filter	only observations		only observations
	only reanalyses	C BSS wrt climatology	only reanalyses
Apply a yearly high-pass running-mean filter	only seasonal forecasts	C Besolution	only seasonal forecasts
with cut-off value of 2 years	only scenario runs only user-defined fields	C Reliability	only scenario runs
	Spatial correlations with a field		only user-defined fields
Filter	only observations	C Uncertainty	Spatial correlations with a field only observations
	only reanalyses	C Tercile RPS, C alternative, C quintile RPS	only observations only reanalyses
ownload ensemble ECMWF-2 1feb T2m	only seasonal forecasts		only reanalyses only seasonal forecasts
expload ensemble ECMWF-2 freb 12m he ECMWF member states do not permit us to give you access to the raw data. Please	only scenario runs	C Tercile RPSS wrt climatology	only scenario runs
onsult the ECMWF ERA or seasonal forecasting website for further information.	only user-defined fields	C Area under under the ROC curve, 💿 alternative1, C alternative2	only user-defined fields
ansas are commission of account intercacing website in interesting mediate in	Verify field against observations	Only compute the netodf files with observations and forecasts	Verify field against observations
	Make EOFs		Make EOFs
	Feedback Geert Jan's home page	Timeseries verification measures	Feedback
	crear own a nome balle	C Pict likelihood	Geert Jan's home page
D KNMI		C Deterministic scores for the ensemble mean (correlation, root mean squared error, and mean absolute error)	
		C Brier score	
		C Plot reliability diagram	
		C Compute Ranked Probability Score for terciles	
		C Plot ROC curve for number of ensemble members below some threshold (C)	

alternative1 © alternative2) © Plot ROC curve varying the model threshold © Only compute the observations/forecasts table

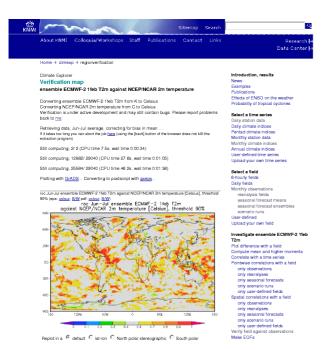
bins 🗖 show R logile

ire a threshold, use 90 %

Threshold

Use 10

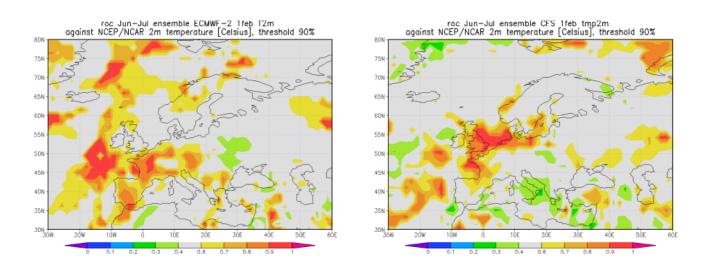
I chose the area under the ROC curve for very warm (90%) June-July, starting from Feb.1 analyses (this is not visible on the screendump, one has to scroll down to see all options). The verifying dataset is the NCEP/NCAR reanalysis, the units are automatically converted to agree with each other. The production of the verification map take less than two minutes, a typical time; most of this time is needed to read the forecast data from the data store:



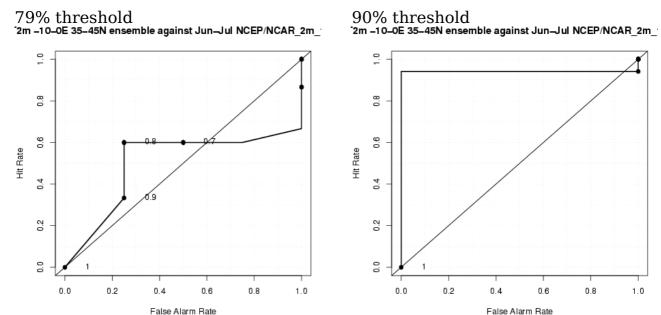
The resulting map is available in a variety of formats (PNG, PDF, EPS), and also the underlying data can be downloaded if the user prefers her own graphics software.

This score show quite good skill in predicting heat waves in southwestern Europe (the Iberian peninsula and France), indicated by the red colours for ROC areas larger than 0.5 (which is obtained by a system without any skill). There is very little skill in these areas in predicting the mean climate (e.g., van Oldenborgh et al, 2005). We suspected that this skill is due to soil moisture depletion: when the winter rains have been much weaker than normal, the soil dries up in summer, leading to higher temperatures on average (see also Ferranti and Viterbo, 2006), but first some cross-checks have to be made.

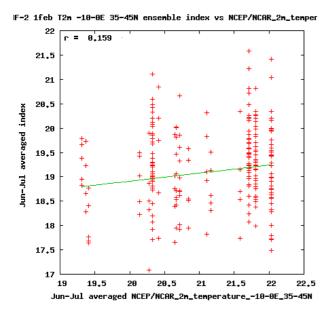
The same map can easily be made for a different forecasting system, the NCEP CFS. That system does not show skill in this area, casting doubt on the ECMWF skill score.



Further analysis of the area-averaged temperature over the Iberian Peninsula shows that the skill only is high when one chooses a threshold that is much higher than 80%, which excludes the years in the hindcast period with only 5 ensemble members. The curves shown are ROC vcurves that plot the Hit Rate against the False Alarm Rate for different sensitivities (number of ensemble members needed to issue a forecast). A system with no skill give a diagonal line with equal Hit Rates and False Alarm Rates. The area under the curve is a measure of the quality of the system.



There is essentially no skill in the 79% threshold ROC curve. The same can be seen in scatterplots of the data: the reason that the warm summers in Spain and Portugal were forecast well is that they all fell in the last few years, for which there are 40 ensemble mebers. The apparent skill in the maps is therefore due to a coincidence of recent warm years and a changing ensemble size.



Conclusions

The Climate Explorer web verification system is the first web-based verification system that allow anyone to quickly and conveniently

- compute skill scores for various forecast systems, as maps or time series,
- compare the skill of different systems,
- investigate apparent skill in certain areas.

We plan to add data of more forecast systems (ENSMEBLES data when it becomes available), make the system more user-friendly, and add confidence intervals to the skill scores.

Geert Jan van Oldenborgh & Caio Augusto Dos Santos Coelho, 7 March 2006

References

L. Ferranti and P. Viterbo, 2006: The European summer of 2003: sensitivity to soil water initial conmditions, submitted to J. Climate.

G.J. van Oldenborgh, M.A. Balmaseda, L. Ferranti, T.N. Stockdale en D.L.T. Anderson, *Evaluation of atmospheric fields from the ECMWF seasonal forecasts over a 15 year period* J. Climate, 2005, 18, 16, 2970-2989, corr. J. Climate, 2005, 18, 5188-5198.