INSTIITUTE OF ACCELERATING
SYSTEMS AND APPLICATIONS,
ATHENS, GREECE
WP10 Coordinator

MEDITERRANEAN OCEAN
FORECASTING SYSTEM:
TOWARD ENVIRONMENTAL
PREDICTION

MFSTEP

Project Deliverable Report
D6

WP10: Atmospheric Forcing
and
Air-Sea Interaction Studies

Project title	**MEDITERRANEAN OCEAN FORECASTING SYSTEM: TOWARD ENVIRONMENTAL PREDICTION**					
Project acronym	**MFSTEP**					
Contract number	**EVK3-CT-2002-00075**					
Deliverable number	**D6**					
Deliverable title	**Definition of the Intercomparison Protocol**					
Work package	**WP10 Atmospheric Forcing and Air-Sea Interaction Studies**					
Name of WP10 Coordinator	**Prof. George Kallos** **Institute of Accelerating Systems and Applications, Athens, Greece**					
Date of delivery	**September 1, 2003**					
Report status	**Version 1.0**	**Draft**		**Final**	**X**	
Project home page	**http://www.bo.ingv.it/mfstep/**					

1) Introduction

This report summarizes the work performed under the scopes of deliverable 6 of WP10. This work package contains the necessary activities to create and deliver the atmospheric surface fields to the WP8 and WP9 ocean modeling community, and to define and perform the Scientific Verification Period (SVP) intercomparison of atmospheric models. The aim of this sub-task (deliverable 6) is to define and document the protocols for atmospheric model intercomparison during the SVP. The AM&WFG/IASA group will perform the hindcasts of SVP using the nonhydrostatic SKIRON/Eta modelling system at high-resolution. IASA will archive the model output for the verification studies and it will be able to disseminate the raw data to the interested project partners through a project dedicated ftp server. The same will be true for the SVP forecasts performed at CHMI with the ALADIN-MFSTEP model's assimilation cycling and forecast. The results of the verification studies will appear in the MFSTEP-WP10 web page of IASA (http://forecast.uoa.gr).

2) SVP Hindcasts

In the framework of the SVP of WP10 simulations will be performed by IASA using the nonhydrostatic SKIRON/Eta modeling system in hindcast mode. The full volume atmospheric fields will be stored and inter-compared with the other atmospheric limited area model (LAM) of MFSTEP. Indeed CHMI will also produce similar forecasts.

During the SVP the atmospheric limited area models will produce 72-hour hindcasts initialized from the daily 0000 UTC ARPEGE analyses of January 2003. 31 simulations will be available for model inter-comparison and validation. For SKIRON the initialization of the soil moisture content and the soil temperature will be performed using the 24-hour forecast of the run of the previous day. The sea-surface temperature field will be forced by the daily 0.5° latitude x 0.5° longitude NCEP SSTs and it will remain fixed to its initial value throughout the simulation. The LAM lateral boundary conditions will be based on ARPEGE forecasts and will be updated every 3 hours. The ARPEGE analyses and forecasts that will be used for LAM initial and lateral boundary conditions will be provided by Meteo-France at a horizontal resolution of 0.25x0.25 degrees. For ALADIN the so-called 'blending' procedure already used for operational goals at CHMI will be used. This allows a scale-dependent mixture of the LAM guess forecast from the previous analysis (6 hour earlier hence) and of the ARPEGE analysis, both for the atmospheric and for the land-surface parts. The problem of soil initialization is thus avoided. The ARPEGE data will here be produced on a terrain-following grid mimicking an ALADIN configuration but with a 37 km mesh. The conversion to the nominal 9.5 km ALADIN-MFSTEP Grid will be performed prior to entering the blending or forecast procedure.

The computational model domain used by each center should include the corresponding dissemination domain. The computational model domain of IASA will cover the whole Mediterranean Region and part of Central Europe (Figure 1) while its dissemination domain will cover the Mediterranean Sea east of 18°E and the Black Sea (29°N-48°N, 18°E-42°E). The west boundary of the computational domain extends further west than the dissemination domain since it is well known that during all seasons synoptic systems or air-masses originating over western Mediterranean (Gulf of Genova, Atlas mountains etc.) strongly affect the weather of eastern Mediterranean (e.g. Kallos et al., 1997).

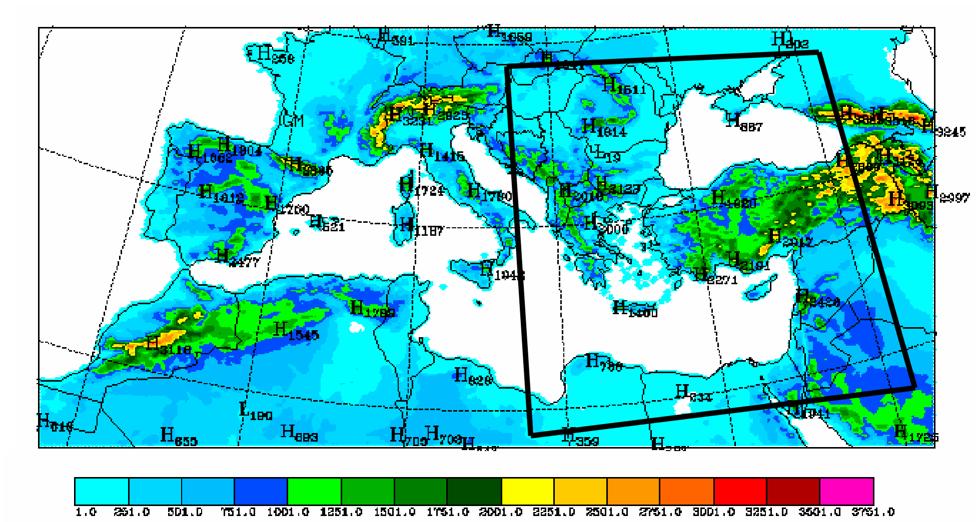

Figure 1. *The topography of the computational model domain of Skiron. The black frame indicates the dissemination domain.*

For ALADIN-MFSTEP the domain will use a conformal tangent Lambert projection of centre at 46.47°N/2.58°E and of SW corner at 26.73°N/18.62°W as well as NE corner at 46.08°N/49.25°E (Figure-2). The computational domain will be of 589 x 309 points with 37 unequally spaced levels in the vertical (from 17 meters above the ground to 5 hPa).

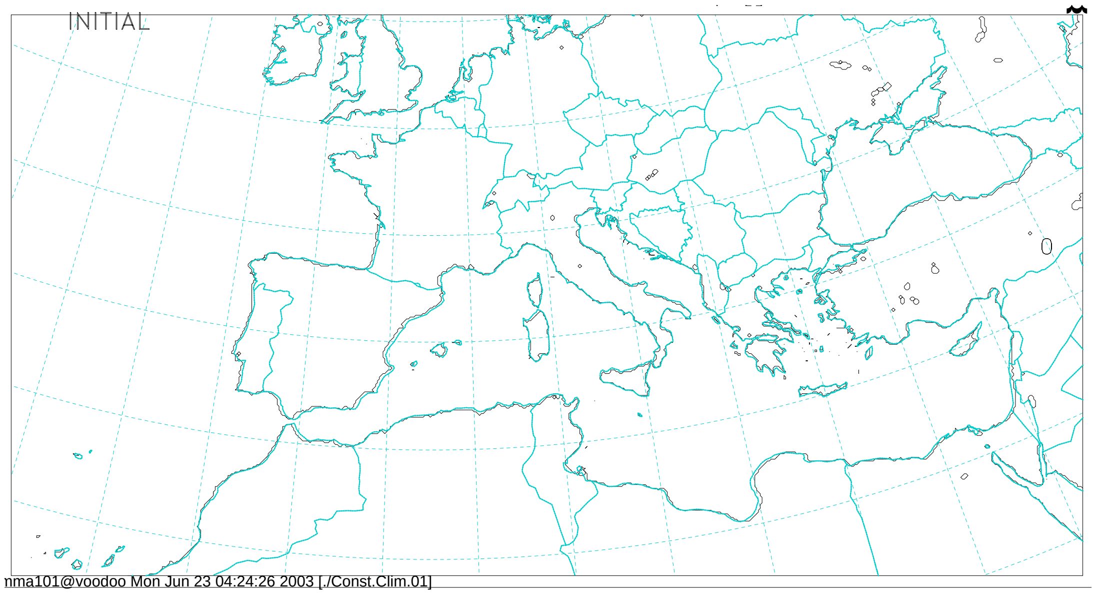

Figure 2. *The computational model domain of Aladin.*

In the horizontal, a grid increment of 0.1 degrees will be applied by all LAMs for their post-processing. The postprocessing domain of SKIRON will extend from 10°W to 42°E and from 29°N to 48°N while the similar domain of ALADIN will extend from 19°W to 37°E and from 30°N to 48°N in order to cover the whole OGCM domain. In the vertical, IASA uses 38 levels stretching from the ground to the model top at 25 mb (corresponding approximately to 25 km). The use of many vertical levels will provide a good vertical representation of the physical processes especially within the Atmospheric Boundary Layer.

It is noteworthy that the setup of each limited area model (i.e. domain, horizontal resolution, number of vertical levels etc.) will be practically the same during the Scientific Validation Period and the Targeted Operational Period. This is necessary since the SVP aims to quantify the performance of the atmospheric models and to provide credibility to the meteorological fields that will be used operationally to force the ocean models.

The output of the LAM hindcasts will be archived on a 0.1°x0.1° regular latitude-longitude grid. More specifically, the surface fields will be stored for every hour because these are used by the ocean modelers and need to receive special attention. The fields that will be provided to the project partners during the Targeted Operational Period and need to be archived are the 10m wind, 2m temperature, 2m specific humidity, cloud coverage/fraction, mean sea-level pressure, total accumulated precipitation, surface radiative heat flux components (shortwave downward, shortwave upward, longwave downward, longwave upward, optionally also for a clear sky equivalent situation), surface latent heat flux, surface sensible heat flux, land-sea mask, and the sea-surface temperature. The atmospheric modelers need to validate, if possible, the ability of the LAMs to predict the above fields accurately. Moreover, the skill of the LAMs to predict upper-air fields such as the geopotential height, temperature and wind speed needs to be quantified. Hence, the upper-air fields will also be archived for every 6 hours.

IASA has setup an ftp server that will be dedicated to the dissemination of the SKIRON/Eta output. Its IP address is 195.134.91.8. A username and a password will be available to the partners after request. The SKIRON/Eta hindcasts will not be available online. However, they will be restored from the archive (if necessary) and will be downloaded from the above ftp server. A similar procedure for the raw products is being prepared at CHMI.

3) Methods for model Inter-comparison and Validation

The atmospheric model inter-comparison and validation that will be performed during SVP is an important task of WP10. The atmospheric modelers will be able to identify and improve model uncertainties (related to the model setup, the pre or post-processing modules, the utilized surface characteristics etc.). It will also provide credibility to the meteorological fields that will be used operationally to force the ocean models.

The validation will be performed with the use of well-known and widely accepted statistical methods. The statistical tests will quantify the relationship between the forecast fields and the actual state of the atmosphere. In this task the maximum available surface observations and the upper-air ARPEGE analyses will be used. The surface stations that provide operational surface observations (METAR and SYNOP) in the area of interest are depicted in Figure 3. This methodology allows the examination of the model performance in the whole period of the simulations, providing significant conclusions about the existence of systematic errors, the accuracy and the credibility of the model forecasts.

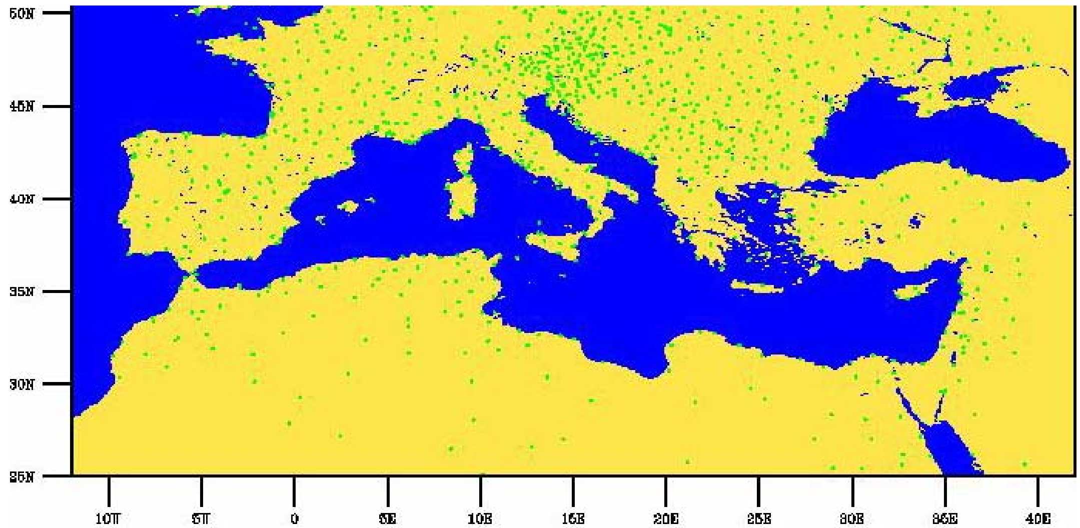

Figure 3*. The network of the surface weather stations that provide observations to the Global Telecommunication System in our area of interest (After Katsafados, 2003).*

The statistical analysis exhibits some differences depending on whether the meteorological variables are discrete or continuous. Discrete variables are allowed to take on only a finite number of values, whereas continuous variables may take on any of the infinitely many real values within their range. The rainfall, snowfall and the cloud cover are considered to be discrete variables, while the temperature, the wind speed and the mean sea-level pressure are continuous variables. On the surface of the earth, the forecasts and the observations of continuous meteorological variables result from a finite number of discrete values. The statistical study of grid-point meteorological fields is also important. The surface pressure, the geopotential heights and the upper-air temperatures appear in the form of grid-point fields, with the use of objective analysis methods in International Meteorological Centers (ECMWF, NCEP, Meteo-France and others).

The partners have chosen a number of statistical methods that will provide a robust validation of the models. The use of wrong statistical methods may lead to misleading results. The statistical methods that will be used in order to validate the model performance are described below:

Continuous meteorological fields (Surface and upper air):

The surface fields that will be examined in the SVP validation studies are the mean sea-level pressure, 2m temperature and 10m wind speed. The proposed statistical methods, applied to discrete variables, are the bias (BIAS) and the root mean square error (RMSE), while an optional test could be the Frequency of Forecast Errors.

The upper-air fields that will be examined in the SVP validation studies are the geopotential height, the temperature, and the wind speed at the isobaric levels of 850 and 500 hPa. The choice of these levels was based on our interest for study of the model performance in the atmospheric boundary layer and in heights that correspond to the free atmosphere without major influence from the surface fields. The proposed statistical methods are the BIAS and the RMSE, and optionally the Anomaly Correlation (AC) and the frequency of forecast errors.

Bias (BIAS):

The bias estimates the correspondence between the mean value of the forecast (F) and the observation (O). This measure calculates the sum of the differences in a total of N values:

$$bias = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i) = \overline{F} - \overline{O}$$

If *bias*<0 (>0), the model underestimates (overestimates) the specific variables.

Root Mean Square Error (RMSE):

This measure is considered to be one of the most popular in the estimation of the forecast accuracy (Wilks, 1995; Katsafados, 2003 and others). It is mostly used in grid-point fields and it is expressed by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2}$$

RMSE takes values greater than or equal to zero. This value is not dimensionless but it exhibits the same units as the validated field. It is an important measure as it provides a quantitative measure of the model performance.

Frequency of forecast errors:

This measure may be applied in the validation of the minimum and maximum temperature values (see Katsafados, 2003). Initially, the differences between the observed and the predicted values are classified in groups of 2°K. The errors of the maximum/minimum predicted temperatures are calculated in smaller time periods for every simulation and they are grouped to the corresponding categories (Brooks et al., 1996). The ideal histogram of error frequencies follows the Gaussian distribution with maximum values around 0. Negative asymmetry of the histogram indicates underestimate of the temperature, while the corresponding category with maximum frequency indicates the magnitude of the underestimation (Mao et al., 1999).

Anomaly Correlation (AC):

The anomaly correlation is considered as a common statistical method for the estimation of the relation between gridded fields. The (AC) calculation is based on the transformation of the forecasted and observed fields in climatological anomalies. In more details, the (AC) estimation is resulted from the abstraction of the corresponding mean climatological values from the prognostic and analysis values respectively. In a gridded field, with IM, JM dimensions, the anomaly correlation coefficient is calculating according to (Wilks, 1995):

$$AC = \frac{\sum_{j=1}^{JM}\sum_{i=1}^{IM}\left[(y_{(i,j)} - C_{(i,j)})(o_{(i,j)} - C_{(i,j)})\right]}{\left[\sum_{j=1}^{JM}\sum_{i=1}^{IM}(y_{(i,j)} - C_{(i,j)})^2 \sum_{j=1}^{JM}\sum_{i=1}^{IM}(o_{(i,j)} - C_{(i,j)})^2\right]^{1/2}}$$

where, $y_{(i,j)}$ denotes the prognostic value of the *(i,j)* grid point, $o_{(i,j)}$ denotes the relevant analysis value and $C_{(i,j)}$ corresponds in the climatological value of each particular meteorological variable. The threshold of 0.6 suggests a characteristic lower limit to successful forecasts.

<u>Discrete meteorological fields (Precipitation)</u>:

The validation of discrete meteorological variables, such as the accumulated precipitation, is based on an I×J contingency table of absolute frequencies, or counts, of the I×J possible combinations of forecast and event pairs. Every element of this table represents the number of cases where the predicted and the observed value of precipitation (for example) exceed the predefined precipitation amounts in the same time periods. In the case of precipitation IASA proposes the use of the predefined values of 0.5, 2, 4, 6, 10, 16, 24, 36 mm per 12 hours. Figure 4 illustrates the essential equivalence of the contingency table and the joint distribution of forecasts and observations for the simple, I=J=2, case.

		Observed		
		Yes	No	
Forecast	Yes	a	b	a+b
	No	c	d	c+d
		a+c	b+d	n=a+b+c+d

Figure 4. *The contingency table for discrete meteorological variables. The letters a to d correspond to all the possible pairs Forecast/Observation. (After Wilks, 1995).*

The term “a” defines the number of cases where the predicted and observed values exceed a predefined precipitation amount, and it corresponds to a successful forecast. The term “b”, corresponding to false alarm cases, represents the number of cases where only the predicted value (and not the observation) exceeds a predefined precipitation amount. The term “c” represents the number of cases that the observed value (and not the predicted) exceeds a predefined amount, and it corresponds to missed cases. The total number of combinations is n=a+b+c+d.

The contingency table is the base for the use of statistical tests for discrete variables and it will be used in some of the proposed tests of SVP.

The proposed statistical methods, applied to discrete variables, are the bias (B) and the root mean square error (rmse), while an optional test could be the Equitable Threat Score (ETS).

Bias (B):

The bias, or comparison of the average forecast with the average observation, of categorical forecasts, is usually represented as a ratio. The bias is simply the number of ‘YES’ forecasts to the number of ‘YES’ observations. In terms of figure 4, the bias ratio is:

$$B = \frac{a+b}{a+c}$$

Unbiased forecasts exhibit B=1, indicating that the event was forecast the same number of times that it was observed. Bias greater (less) than one indicates that the event was forecast more (less) often than observed.

Equitable Threat Score (ETS):

This measure can be used as an index of the model ability to predict a predefined precipitation amount. In terms of figure 4, the Equitable Threat Score is:

$$ETS = \frac{a - a_r}{a + b + c - a_r}$$

where $a_r = \dfrac{(a+b)\cdot(a+c)}{a+b+c+d}$ defines the random frequency of successful predictions that exceed a predefined precipitation amount (Schaefer, 1990).

The Equitable Threat Score takes values between –1 and +1 (included). In the case of perfect forecasts ETS=1, while ETS is near 0 in constant or random forecasts.

Root Mean Square Error (rmse):

This statistical measure quantifies the range of the differences between the predictions and the observations that exceed a predefined precipitation amount (Colle et al. 1999, 2000) when it is applied to discrete variables. The rmse is expressed by the equation

$$rmse = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - X_i)^2}$$

where P_i and X_i are the predicted and observed precipitation amount in a total number of N observations that exceed the predefined value. The importance of this measure lies on the fact that it provides a quantitative measure of the model performance although it cannot indicate overestimates or underestimates of the forecasts.

4) Future Work

The works planned for the second half of 2003 include the finalization of the tests and the system automation. The SVP hindcasts will be produced and the model will be verified. Finally, the high resolution forecasting products will be distributed to the partners at the required format.

References

Brooks, H.E., and C.A. Doswell III, 1996: A comparison of measures-oriented and distributions-oriented approaches to forecast verification. *Wea. Forecasting*, **11**, pp.288-303.

Colle, B. A., K. J. Westrick, and C. F. Mass, 1999: Evaluation of MM5 and Eta-10 Precipitation Forecasts over the Pacific Northwest during the Cool Season**.** *Wea. Forecasting,* **14**, 137–154.

Colle B.A., C.F. Mass and K.J. Westrick, 2000: MM5 precipitation verification over the Pacific Northwest during the 1997-1998 cool seasons. *Wea. Forecasting,* **15**, pp. 730-744.

Kallos, G., V. Kotroni, K. Lagouvardos, A. Papadopoulos, M. Varinou, O. Kakaliagou, M. Luria, M. Peleg, A. Wanger, and M. Uliasz, 1997: Temporal and spatial scales for transport and transformation processes in the Mediterranean. Proc. of the 22nd NATO/CCMS Int. Techn. Meeting on Air Pollution Modeling and Its Application, 2-6 June 1997, Clermont Ferrand, France. Edited by Sven-Erik Gryning and Nandine Chaumerliac, Plenum Press, New York, Vol 20, pp.8.

Katsafados, P., 2003: Factors and parameterizations that determine the performance of limited area models in long-range weather forecasts. PhD Thesis, Dept. of Physics, Univ. of Athens, Athens, Greece. pp. 257 (in Greek).

Mao, Q., R.T. McNider, S. Mueller, and H.M.H. Juang, 1999: An optimal model output calibration algorithm suitable for objective temperature forecasting. *Wea. Forecasting*, **14**, pp.190-202.

Schaefer, J.T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, pp.570-575.

Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences, Academic Press NY, pp. 467.