6.5 Test on the EOF's representativity: Historic Data Eigenvectors.

In this section we shall present the fields of the three basic oceanographic variables as given by the EOFs extrapolation methods, but using the eigenvectors obtained from historic data.

We first present the first six leading vector modes that result from the summer and winter casts (Figure 6-76 and Figure 6-77). These figures consist of six frames: the left and right columns correspond to the non-standardized and standardized analyses, while the rows correspond, from top to bottom, to potential temperature, salinity and density.

Afterwards we will show the results for the different campaigns with three figures for each one of the variables:

a) Successive error profiles with the first 24 modes, for the non-standardized and standardized analyses. These figures give a clear idea of the convergence towards the data as additional modes are included.

b) Contour results at 10 m depth after interpolation with six, twelve, eighteen and twenty four modes for the non-standardized and standardized analyses (left and right columns) at 10 and

c) as in b), but at 100 m depth.

In order to easily compare the resulting contour fields and to examine their convergence towards actual data, contour figures obtained from observed profiles of potential temperature, salinity and density distributions at 10 and 100 m are included in Appendix 1. The figures at 10 m have already been shown in Chapter 5, but the ones at 100 m have not been shown previously. As in previous sections, potential temperature, salinity and density distributions are presented for FANS III and FANS II, while for FANS I and MEGO 94 we only show density distributions.

It is worth recalling that, as in the previous section, error profiles were estimated with all the available casts at every depth, and therefore errors can locally increase when an additional mode is added. While most of the x-axis in these error figures range from zero to 100, there are some exceptions (some huge error values are obtained with the first modes, particularly with the standardized analysis of density data).

Another important comment is that error profiles are estimated from non interpolated data, while the contours plots result after the SC interpolation procedure, and this may lead to some differences between both errors. When contour plots are discussed, any mention to error values will obviously refer to the values obtained from the interpolated fields. Squared brackets, ([% on-grid error values]) will be used to distinguish them.

Finally, we recall that one of the advantages of the EOF methodology is that data can be represented in a compact way, that is, with a few leading vector modes and their corresponding amplitudes. If these amplitudes are projected onto a grid with a suitable method, like successive corrections or optimal interpolation, then a complete three dimensional grid can be obtained that contains, in principle, a large amount of the variability of the field.

Following such procedure with the EOFs derived from historic eigenvectors, we realized that differences between estimated fields and observed fields were sometimes very different when computed at station points (without any spatial interpolation) and at grid points. For the standardized analysis of density, for instance, we obtained that on-grid errors at 100 m were extremely high compared to those obtained for the profiles, and this obviously reflected also on the contour plots (convergence towards actual data did not take place even using 24 modes). A closer analysis revealed that:

a) these discrepancies tended to increase with depth

b) they usually occurred when the amplitudes of the leading modes had magnitudes and sign strongly dependent on the bottom depth.

This situation is common to results obtained with the standardized analysis, but it was also observed in some cases with the non-standardized analysis.

We then proceeded to interpolate the estimated profiles themselves and obtained on-grid error results that were coherent with the error profiles. This improvement also became clear in the contour plots. By interpolating only the estimated profiles that actually contribute to every depth level, the number of data used as input for the SC algorithm at 100 m is lower than at 50 m, while the results at 10 m do not differ significantly, since the same number of data points are interpolated.

For the above reason, all the contour plots presented in this section were obtained from the interpolation of the estimated profiles instead of the amplitudes. We conclude that, even though the number of operations might increase significantly, the results highly justify this procedure. Equivalent results are obtained interpolating only the amplitudes from casts that reach (or surpass) the desired depth level.

6.5.1 Historic Eigenvectors

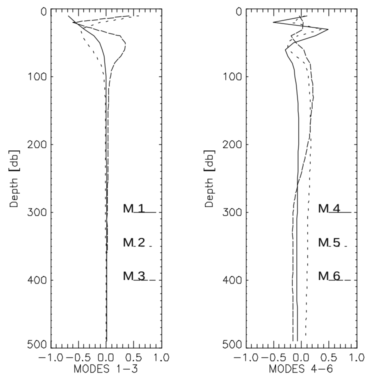
One of the first characteristics that are noticed is the smoothness of the vectors, which results from the larger data set from which they were obtained. This differs significantly from the eigenvectors that result from the campaigns casts, which show in general more spiky profiles (with the only exception of dynamic height, which result from a vertical integration process).

Comparing the left and right columns on both Figure 6-76 and Figure 6-77, the vertical variability is clearly larger in the standardized case (right column), with a rather remarkable similarity in the overall behaviour of the first three modes, in all the variables and for both seasons. The first mode is the closest one to a barotropic signal, which would have a uniform vertical value.

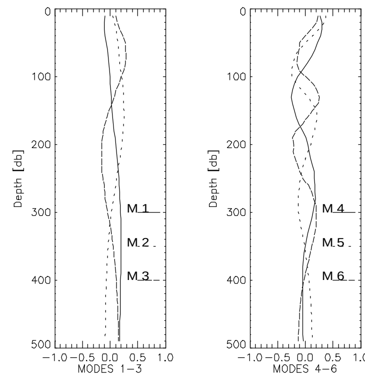
As for the non-standardized analysis (left column), there is a clear similarity between the temperature and density summer and winter vectors, but they differ in the depth at which the profiles tend to change their behaviour. This is very apparent for the first mode, and reflects the presence of the thermocline and pycnocline, which are found at larger depths during winter.

The above situation is not encountered in salinity, where the first mode for both seasons behaves like their corresponding standardized profiles. The vertical variability of all vector modes extends again to deeper levels in the winter eigenmodes.

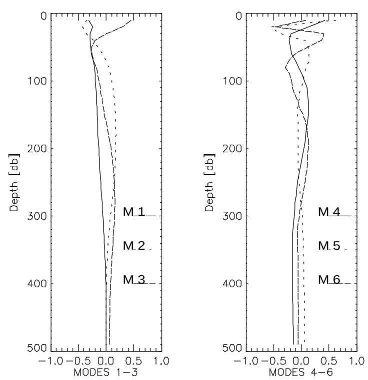
The above vector profiles might indicate a similarity between the temperature and density distributions within the historic data. These similarities are also observed in the average and standard deviation profiles (Figures 5-13 and 5-14).



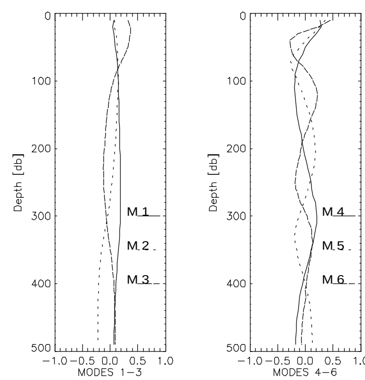
(a)



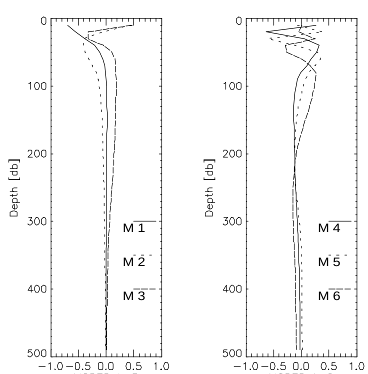
(b)



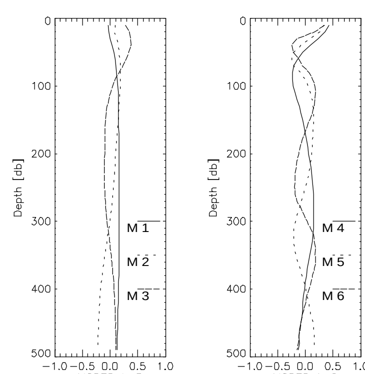
(c)



(d)



(e)



(f)

Figure 6-76 Historic summer eigenvectors for the first six modes. The left (right) column corresponds to the non-standardized (standardized) analysis, while the rows, from top to bottom, correspond to potential temperature (a-b), salinity (c-d) and density (e-f).

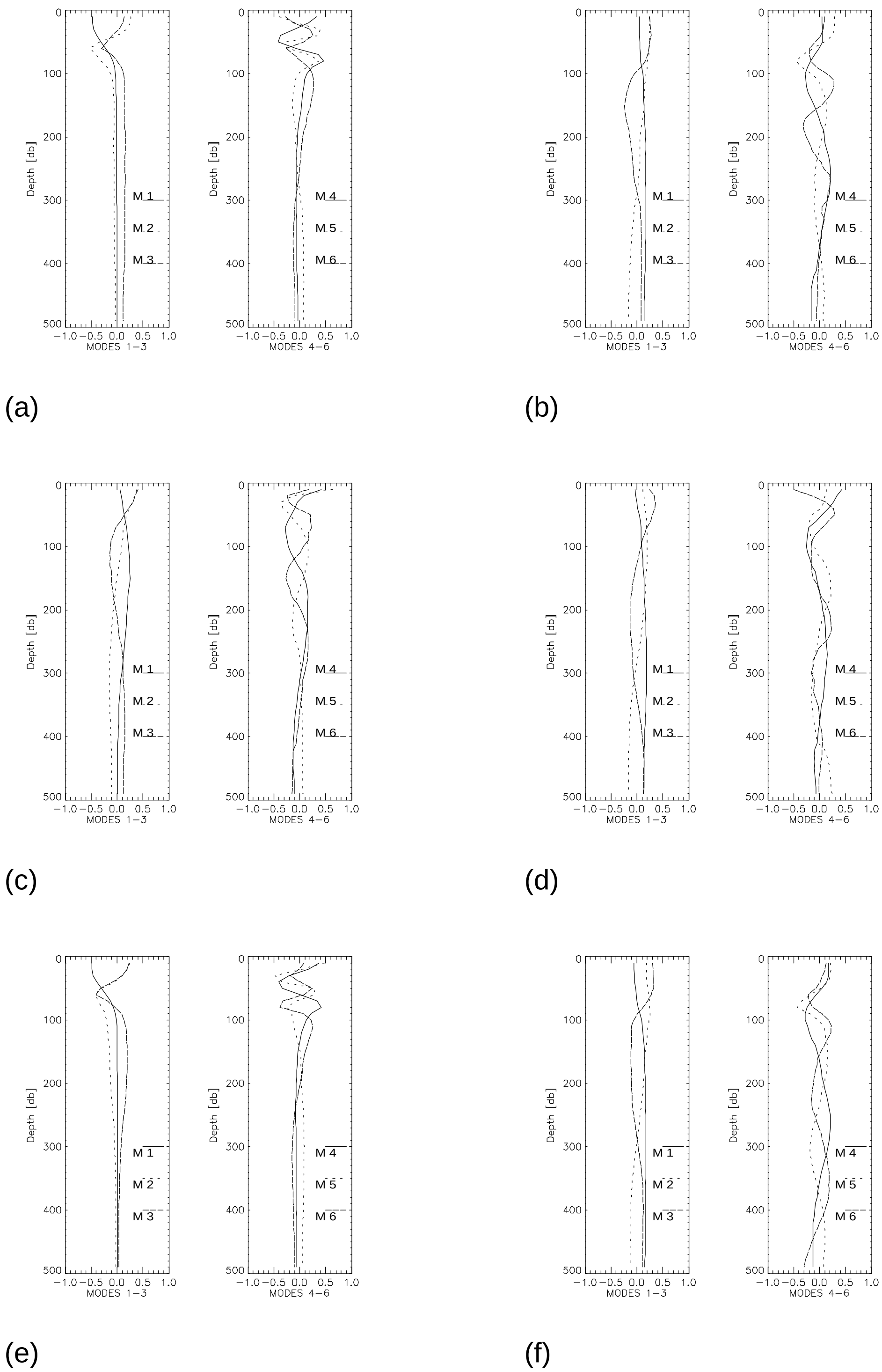


Figure 6-77 Historic winter eigenvectors for the first six modes. The left (right) column corresponds to the non-standardized (standardized) analysis, while the rows, from top to bottom, correspond to potential temperature (a-b), salinity (c-d) and density (e-f).

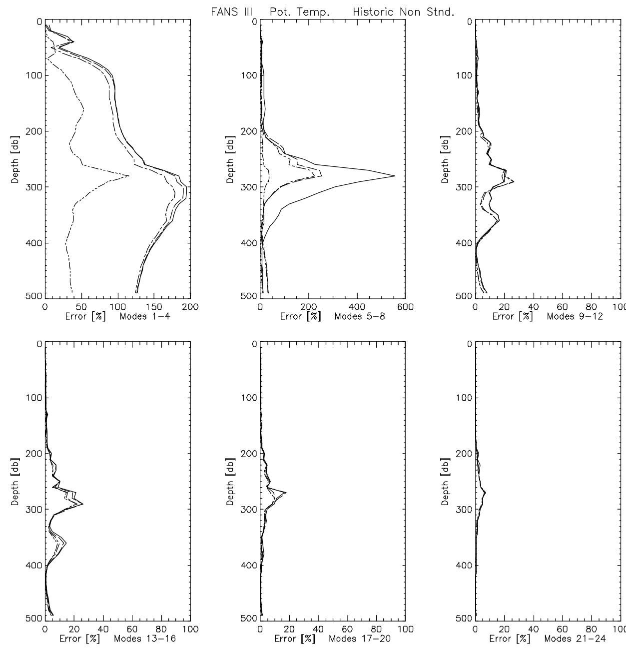
6.5.2 FANS III – Historic Summer Eigenvectors

Since the vectors used as basis were obtained from historic data casts, it is not surprising that error profiles differ significantly from the previous section. Here we are more interested on the convergence towards the data as additional modes are included, as this will indicate to what extent the analysed data and historic data obey a similar statistics.

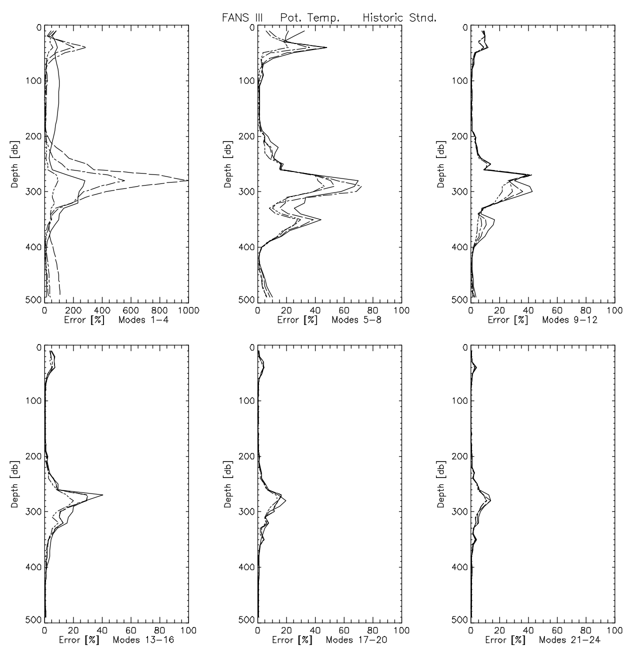
The potential temperature error profiles (Figure 6-78) show some cases in which the errors increase as additional modes are added. In the non-standardized case this happens with the fifth mode, while the addition of the second and third mode increase significantly the error of the standardized analysis. As in the previous section, this larger errors usually take place at some depth levels only, while at others they decrease. Since the non-standardized analysis considers mainly the upper layer variability, fewer modes are needed to get low error values, while the opposite is true for the bottom layers. In this particular case, the standardized analysis results in a mid-depth layer with higher errors after the first seven modes, and this situation holds up to the 24 modes considered.

As for the contour distributions at 10 m (Figure 6-79), the similarity with the data considering the first six modes is excellent for the non-standardized analysis, and additional modes do not change the resulting contours nearly at all. On the other hand, 18 modes are required to get an 8% error (calculated with the interpolated values) with the standardized analysis.

At 100 m (Figure 6-80) the differences are less notorious, and both analyses result in very similar distributions with the first 12 modes (errors of 5% and 7.5% respectively).



(a)



(b)

Figure 6-78 FANS III potential temperature error profiles for the non-standardized (a) and standardized (b) analyses, considering the addition of successive modes, from 1 to 24.

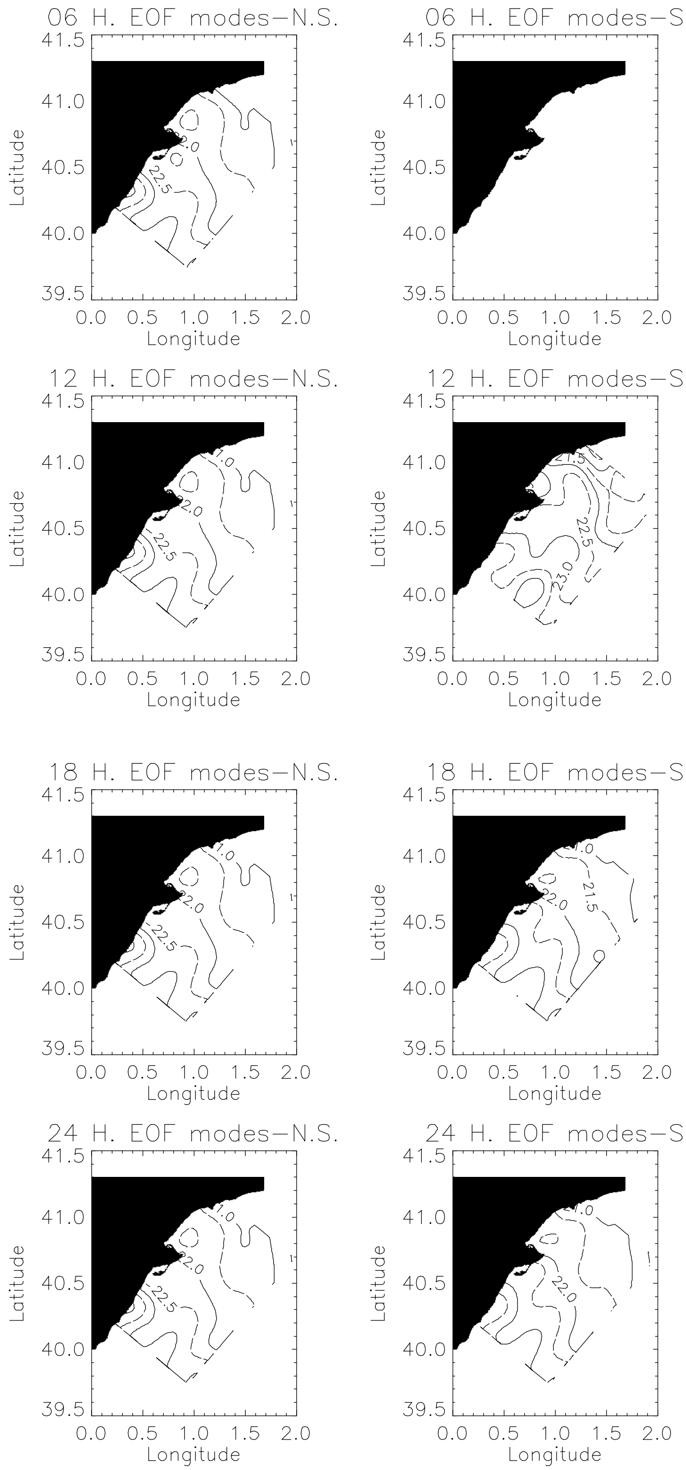


Figure 6-79 FANS III potential temperature contours at 10 m considering 6, 12, 18 and 24 modes, for the non-standardized (left) and standardized (right) analyses.

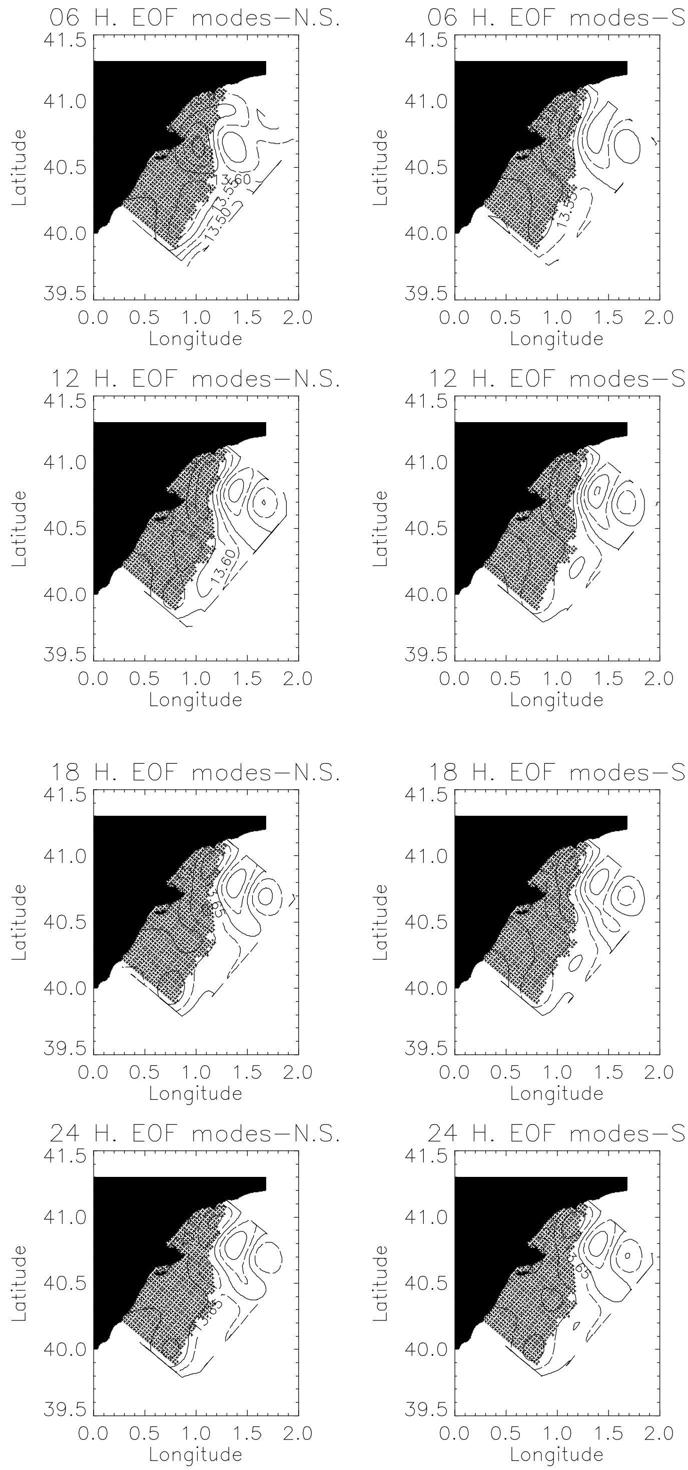
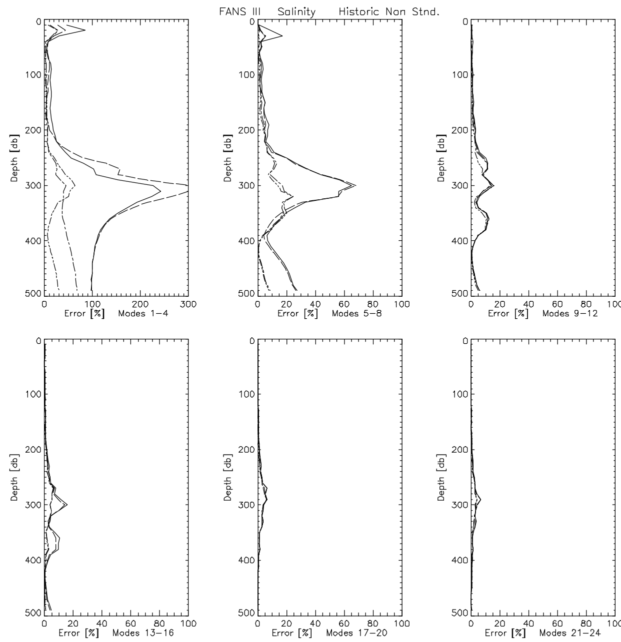


Figure 6-80 FANS III potential temperature contours at 100 m considering 6, 12, 18 and 24 modes, for the non-standardized (left) and standardized (right) analyses.

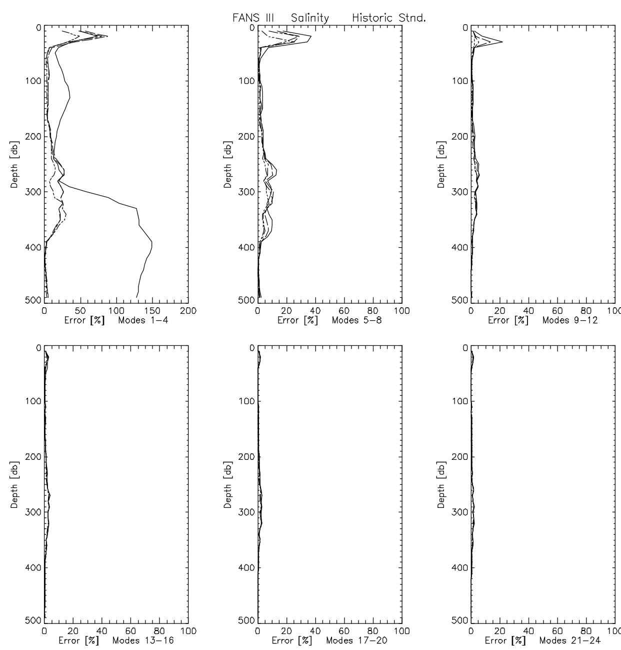
The salinity error profiles (Figure 6-81) have a behaviour that conforms to what has already been written. While the non-standardized analysis resolves faster the upper layer, the standardized one does so for the lower layer (though the variability is always higher closer to the surface).

While the contour plots at 10 m (Figure 6-82) show a very close approach to the data with six modes for the non-standardized analysis, the on-grid error values with both analyses are very close when considering 12 modes, and in fact, the error with 24 modes turns out to be slightly lower in the standardized analysis [3.5% vs 4.5%].

At 100 m (Figure 6-83), the standardized analysis with six modes gives a better approach to the data [8.4% vs 15%], but the convergence to the data is faster for the non-standardized analysis, with an error of 1.7% vs 5.3% with 12 modes. When 18 modes are considered, both error values are below 2%, and these are even lower than 1% with 24 modes.



(a)



(b)

Figure 6-81 FANS III salinity error profiles for the non-standardized (a) and standardized (b) analyses, considering the addition of successive modes, from 1 to 24.

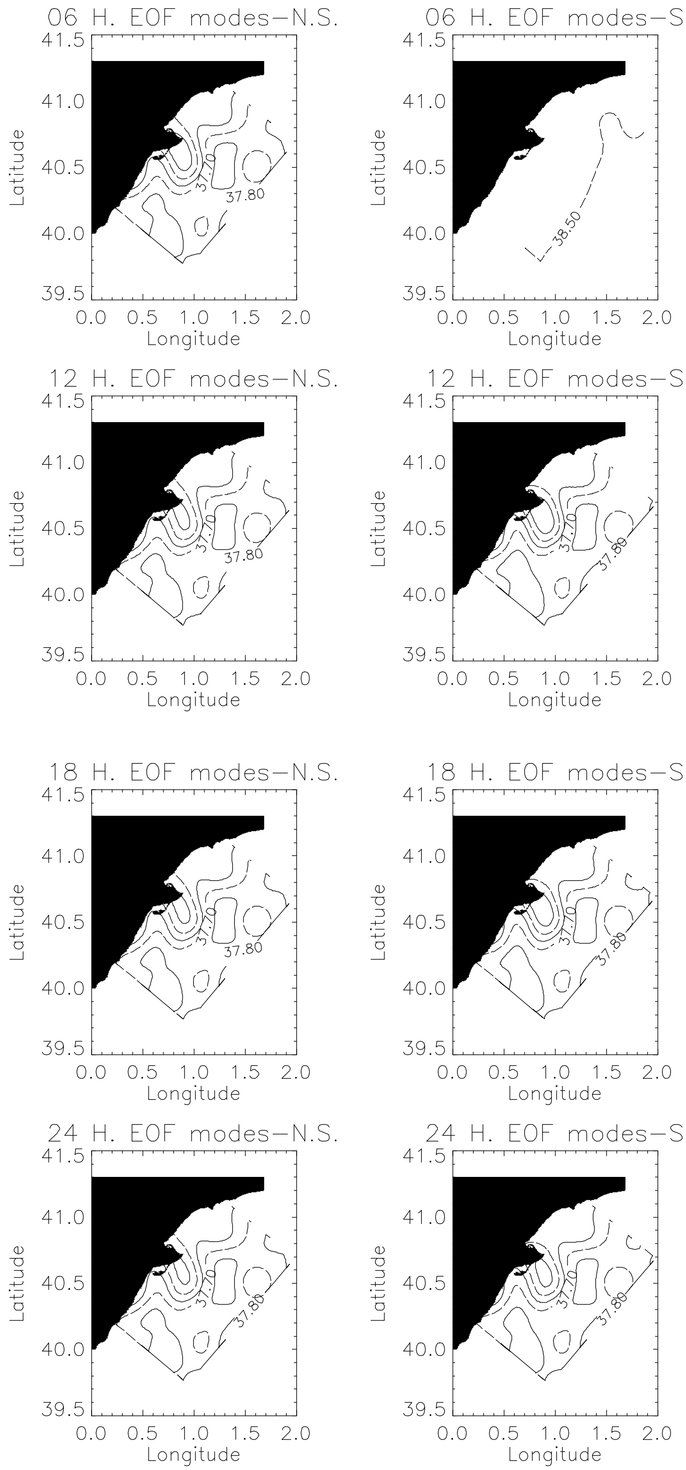


Figure 6-82 FANS III salinity contours at 10 m considering 6, 12, 18 and 24 modes, for the non-standardized (left) and standardized (right) analyses.

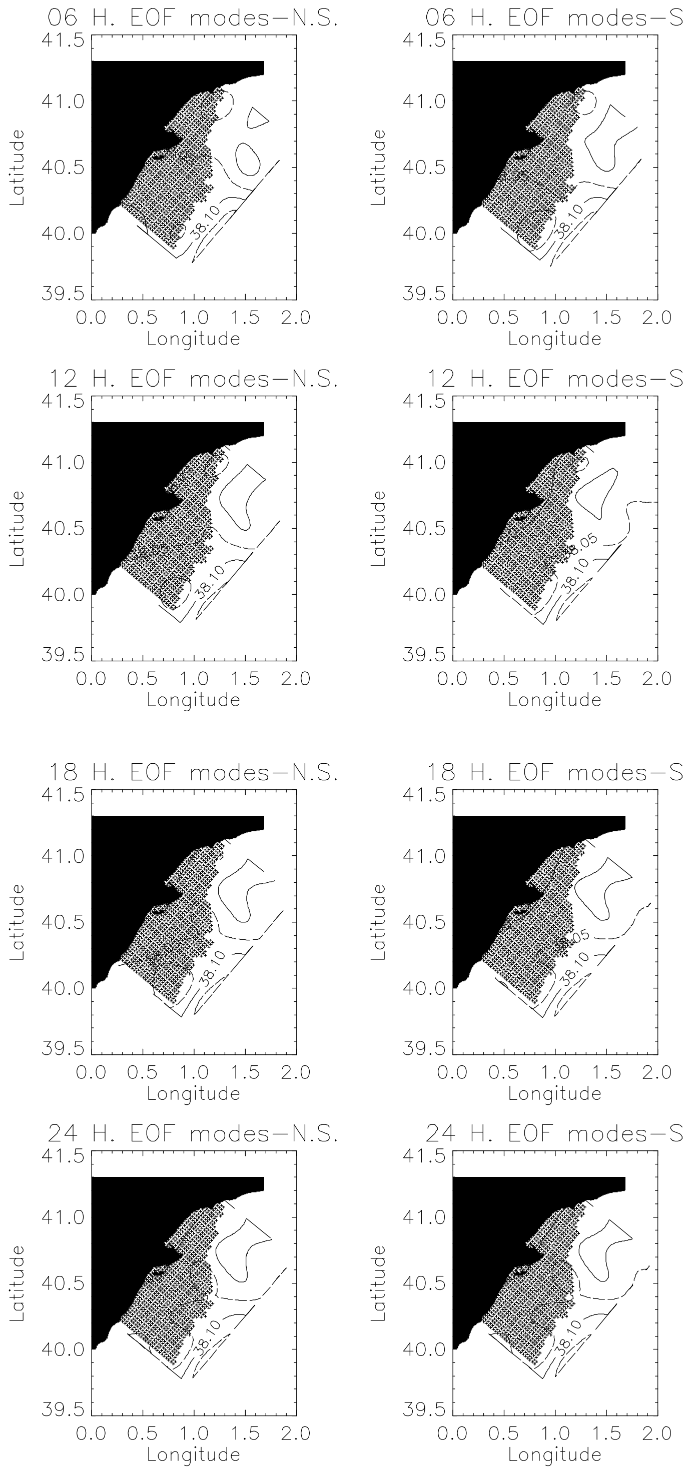


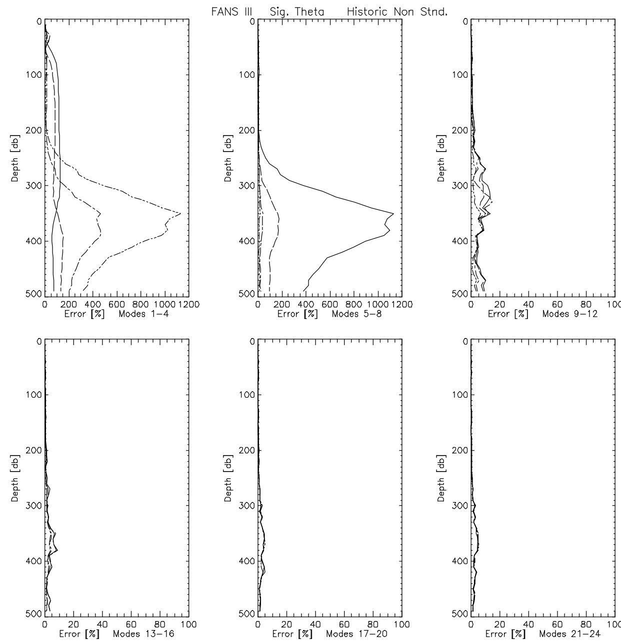
Figure 6-83 FANS III salinity contours at 100 m considering 6, 12, 18 and 24 modes, for the non-standardized (left) and standardized (right) analyses.

For the density field of FANS III, error profiles (Figure 6-84) clearly signal out critical regions with the first modes (x-axis well above 100). For the non-standardized analysis (frame a) the error profile with the first and second modes is lower than 200%, but it then increases to much higher values below 200 m. When the sixth mode is considered the error starts to decrease again. Nonetheless, while the error grows below 200 m, above this depth it decreases as the number of modes increases. When the eighth mode is considered, there are no error values above 60%.

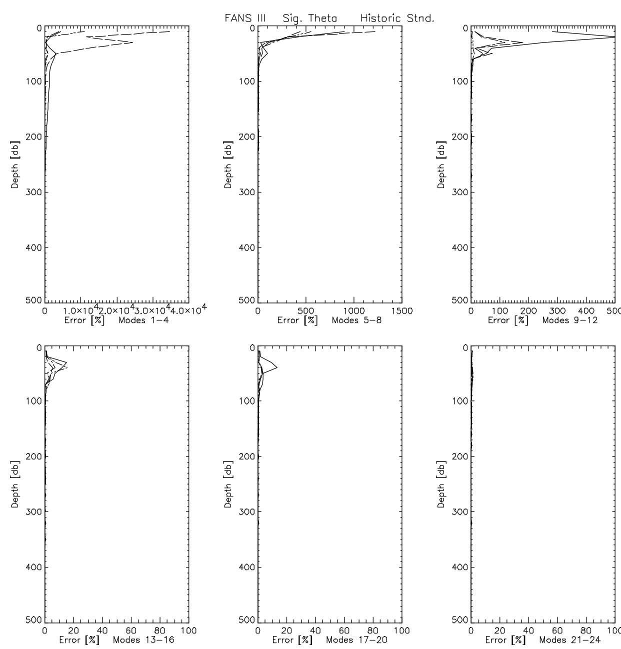
On the other hand, the error for the standardized analysis (frame b) reaches very high values in the upper 60 m, with alternating higher and lower errors as the number of modes increases. It is not until the twelfth mode that the pick values decrease to around 20%.

The contour plots at 10 m (Figure 6-85) show a clear evolution of the resulting distributions. The non-standardized analysis (left column) with 18 modes has a distribution that really approaches the data all over the domain, with the only exception of the area adjacent to the southern Ebro bar [error less than 3%]. As for the standardized analysis (right column) the changes are more significant, beginning with a field (six modes) that does not resemble at all the data, up to 24 modes, when it resolves reasonably well the domain but fails to reproduce the gradients adjacent to the coast in the southern half. In this particular case, even 24 modes can not bring the interpolated fields error to values lower than 25%.

At 100 m (Figure 6-86) the data field is well reproduced with 18 modes in the non-standardized analysis [error less than 1%], while with the standardized analysis the error considering 24 modes is 12%.



(a)



(b)

Figure 6-84 FANS III sigma theta error profiles for the non-standardized (a) and standardized (b) analyses, considering the addition of successive modes, from 1 to 24.

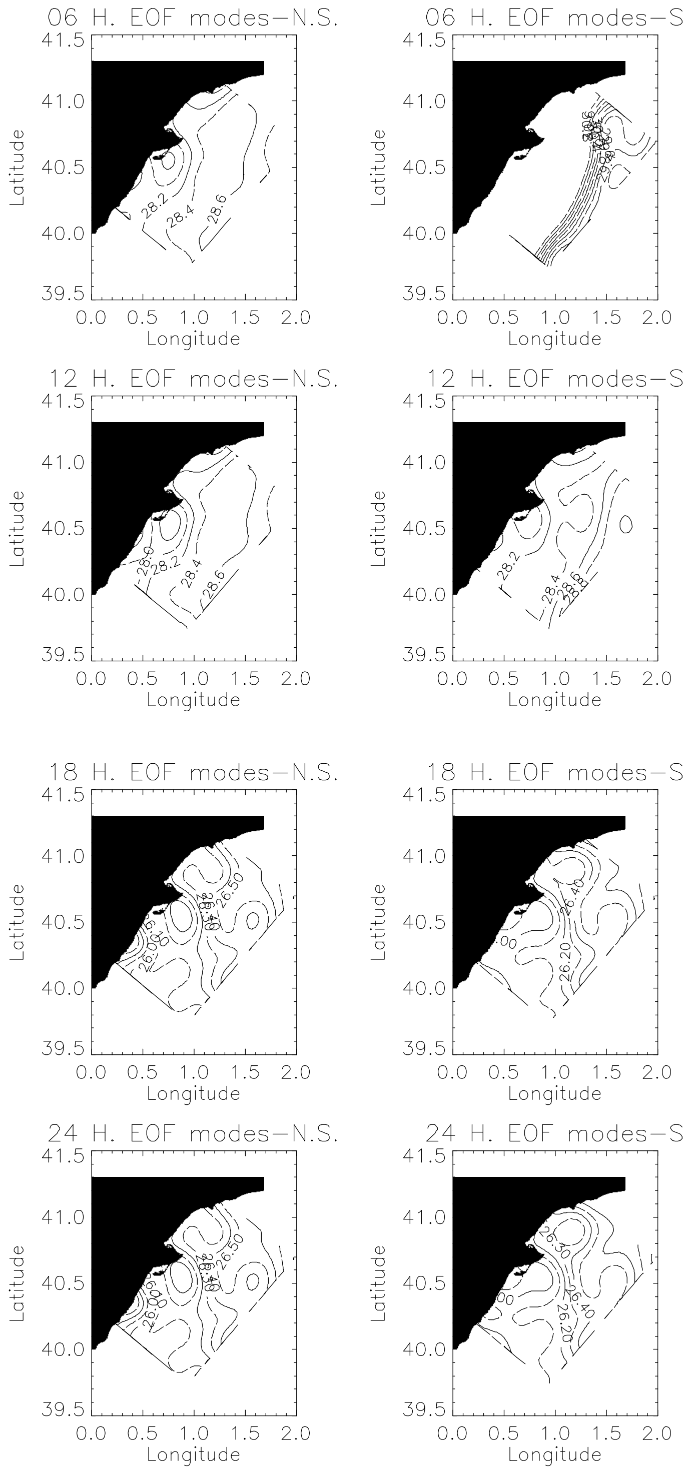


Figure 6-85 FANS III sigma theta contours at 10 m considering 6, 12, 18 and 24 modes, for the non-standardized (left) and standardized (right) analyses.

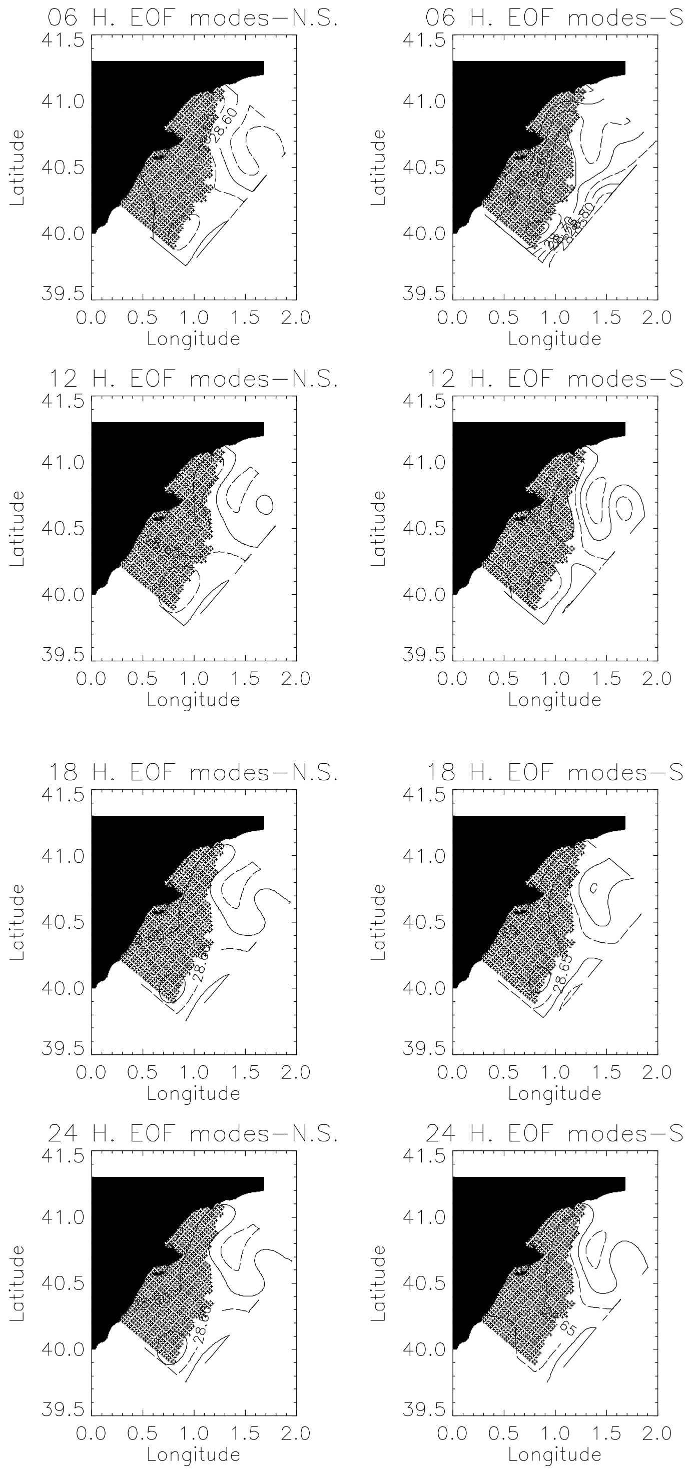


Figure 6-86 FANS III sigma theta contours at 100 m considering 6, 12, 18 and 24 modes, for the non-standardized (left) and standardized (right) analyses.