U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 341

ESTIMATION AND REMOVAL OF FORECAST CYCLE STRAY

H. JEAN THIÉBAUX AND LAUREN L. MORONE

JULY 1988

ESTIMATION AND REMOVAL OF FORECAST CYCLE STRAY

H. Jean Thiébaux and Lauren L. Morone

Development Division
National Meteorological Center
NWS, NOAA, Washington, DC

ABSTRACT

We have found spatial and temporal coherencies in the departures of short-range-forecasts from observed and analyzed geopotential fields, which can be used diagnostically and as the basis for adjustment of forecast products. We identify characteristics of the differences between forecast and observed, and forecast and analyzed values; and we demonstrate the significant impact on forecast accuracy which may be gained through the routine estimation and removal of these discrepancies. In addition, we devised and tested a mechanism for estimating and removing these discrepancies, globally, in the data assimilation step of the forecast cycle.

Section 1. Introduction

Research in numerical weather prediction clearly has as its principal objective bringing forecast model output into agreement with the signal component of the fields that are observed and analyzed. Increasing the sensitivity of objective analysis algorithms, refining physical parameterizations of prediction models, and improving resolution of numerical integrations, all contribute to progress toward this goal; and measures of forecast skill provide evidence of continuing progress. Nonetheless, global comparisons of current, observed and state-of-the-art forecast fields show space/time coherent discrepancies. We demonstrate here that the coherency itself offers both a diagnostic mechanism for identification of sources of error, and an opportunity for statistical adjustment of forecast products to bring them into closer agreement with the atmosphere as it is observed and analyzed.

Research on objective analysis of forecast errors, at the U.S. National Meteorological Center (NMC) led us to an examination of the extents of spatial and temporal coherencies of discrepancies between forecast geopotential values and North American radiosonde reports (RAOBs). In turn, this led to investigations of discrepancies between forecast, analyzed, and initialized fields, for North America and then for other regions, in both hemispheres. Because, initially, our focus was on objective analysis, the forecasts were 6- and 12-hour global forecasts, verifying at 00 or 12 GMT, with results relevant to the possibility of

adjusting the forecast cycle during data assimilation. We have since repeated similar studies with 24-hour forecasts, whose results are relevant to forecast-product correction. Both are reported.

The literature on forecast model error is quite voluminous but very few published papers have appeared on methods of correcting for systematic error. Emphasis has been on model-dominated error by focusing on forecasts in excess of 48-hours. For example, systematic error and error growth in the ECMWF model has been the subject of papers by Arpe and Klinker (1986) and Dalcher and Kalnay (1987). Papers that have suggested correction procedures are Bennett and Leslie (1981, 1982, 1983) and Saha and Alpert (1988).

Bennett and Leslie, in their series of papers, have described a scheme for correcting regional forecasts for the Australian primitive equation model. Their procedure corrects sea-level pressure fields by applying statistically-based adjustments of the SLP gradients. The adjustments at each grid point are obtained by regression techniques involving the coefficients of a truncated series of empirical orthogonal functions of the forecast-analysis structure determined from a dependent sample. The time coherent structure of the forecast error field does not seem to be explicitly accommodated in this scheme, but only a seasonal mean error.

Saha and Alpert (1988) determine a correction field to be applied to the forecast by averaging spectral coefficients of initialized analysis and forecast differences over various periods on the order of 20 to 75 days. By using spectral coefficients various correction scales were investigated.

As motivation and rationale for our formulation of a continuously-updating estimation scheme, we present in Section 2, the evidence for spatial coherence of mean errors which alerted us to their presence and persistence. By repeated partitioning of a season's data, we demonstrate the evolving character of departures of forecast from observed fields. Results of the latter evoked the designation "forecast cycle stray" for the space-time coherent discrepancies between forecast and signal fields.* The evident time scale of the persistence in the evolutionary component of this forecast error (i.e., neither long-term mean nor incoherent 'noise') suggested the formulation and tests of the stray estimation scheme which we describe in this section.

In addition, Section 2 provides results of initial regional tests of the impact of estimating and removing the stray, using RAOB-minus-forecast differences for 6- and 12- hour forecasts, together with results of an

---

* This is an obvious adaptation of Webster's definition of the verb "stray": "to wander from a direct course, to deviate or to err, as to wander from a path of duty". The dictionary relates the corresponding noun to the mobile behavior of some domestic animals; and we make this a bit more specific to our context as: "wandering about by spatially and temporally coherent path which is not predictable by physical theory". Thus, we have a characterization which aptly describes our subject.

investigation of the possibilities that it enters the cycle through the analysis algorithm or via the initialization procedure.

In Sections 3 and 4, impact test reports show analysis-minus-forecast mean differences (MNERRs) and root-mean-square differences (RMSDs) for sequential daily verifications, which contrast scores for unadjusted forecasts with those for forecasts from which estimates of the stray have been removed. We summarize contrasting verification sequences with histograms indicative of the level of improvements in skill which may be brought by daily adjustments based on the most recent history of analysis/forecast discrepancies.

Section 5 describes experiments which explore and contrast results of two avenues which suggest themselves for accommodation of coherent forecast errors in a global, operational forecast cycle. Section 6 presents an overview of our findings and their evident implications.

Section 2. Evidence and Estimation of Time-Coherent Forecast Errors Defined by RAOBs

In our study of the statistical structure of the differences between the 6-hour forecast geopotential values of NMC's global spectral model and verifying radiosonde observations, we produced maps of seasonal mean values for selected mandatory levels. Figure 1 is an example which presents average values, from the 1984-85 winter data base, for 00 and 12 GMT

combined (and separately). What we noted was not only that they are larger in magnitude than anticipated, but that by and large they are spatially consistent in sign. Other mandatory levels we examined showed similar behavior.

We present the evidence for spatial coherence of mean errors, which alerted us to their significant presence, and a demonstration of the oscillatory character of the departures of forecast from observed fields, by repeated partitioning of a season's data, as motivation and rationale for our formulation of a continuously updating estimation scheme. We examined in detail the behavior of time-averaged forecast error (f.e.) values at individual station locations: by treating 00 and 12 GMT separately, and by averaging over shorter, consecutive, non-overlapping time intervals. Table 1 gives the results for seven stations, selected to cover the North American region. It presents both 10-day mean f.e.'s and corresponding 5-day mean values. Comparisons of these with each other, and with Figure 1, indicate that the departures of 6-hour forecast geopotentials from the signal in the atmosphere tend to have oscillatory behavior. Accordingly, for any verification time, the average value for the entire season is not representative of the time specific "stray".

Looking at a pyramid array for the first station in Table 1, as shown in Table 2, we see that successive shortening of the averaging interval reveals increasing detail in the time-varying structure of the errors - - to seven days. By comparison, the sequence of five-day averages appeared

to be noisy, undoubtedly from observation error contamination. Accordingly, we selected 7 days as the base averaging interval for stray estimation and removal.** The course of our thinking, the evidence, and our final objective, may be represented schematically, as in Figure 2. In 2a. the schematic represents time evolution of geopotential observations at a fixed point of the atmosphere. 2b. represents the ideal situation with RAOBs varying about a "bulk state signal" closely anticipated by forecast values - - for which time averaged forecast errors would be near zero for all averaging intervals and with little, if any, evidence of spatial coherency. In 2c the schematic represents a different situation, one which would produce the statistical features illustrated in Figure 1, and Tables 1 and 2. Specifically, the forecast is shown as departing from the signal field by an increment which is semi-persistent in time, with a significant interval scale. This time-coherent discrepancy, denoted by 's', is the forecast-cycle stray. Our objective is to generate time-specific estimates of s which may be used to adjust forecast values, thus bringing signal, forecast, and observation values into closer agreement, as represented in 2d.

For an initial diagnostic study we accumulated files of 6- and 12-hour geopotential forecasts, and the verifying (radiosonde-)observed, analyzed, and initialized fields. With 90 days of North American winter data, sequences of root-mean-square differences (RMSDs) between forecast values

** Some of the impact tests described below were repeated with both longer and shorter intervals, with results which confirmed the original choice.

and RAOBs were computed. This was done for five mandatory levels (1000, 850, 500, 250, and 100 mb), separately for 00 and 12 GMT. Corresponding values were computed with forecasts which had been adjusted by estimates of the stray. Specifically, at each location and verification time, a tapered mean of the previous seven days' discrepancies between forecast and observed values was subtracted from the current forecast, prior to verification. Thus the more recent f.e.'s were weighted more heavily; and observation errors could be expected to average close to zero in the correcting values. The results are illustrated in Figure 3, for both 6- and 12-hour forecasts of 850 mb geopotential, verifying at 12 GMT. At each verification time, the value plotted is a RMSD in which the summation is over mid-latitude North American radiosonde stations (specifically, those within 25-55°N latitude and 80-120°W longitude).

The information presented by Figure 3 makes it evident that the average forecast-error reductions brought by adjusting each forecast according to recent error histories, are about 40% for both forecast periods. As anticipated, the averaged squared discrepancies of the 6-hour forecast and RAOB geopotentials tend to be smaller than for the 12-hour forecast, although the differences between 6-hour and 12-hour RMSDs are very much reduced when the forecasts have been adjusted for their time-coherent errors.***

*** In fact, not only do the adjusted 12-hour forecast geopotentials have significantly greater skill than the values of the unadjusted 6-hour forecast, but nearly the same degree of improvement in skill can be achieved by 'correcting' the 12-hour forecast as is bought making and correcting a 6-hour forecast.

This unequivocal demonstration of semi-persistent error maintained by the forecast cycle, begs questions of identification with elements of the cycle:

⇨ post-processor ⇨ analysis ⇨ initialization ⇨ numerical integration ···

any of which might be suspected of contributing, and, questions of ultimate correctability. We addressed the question of identification with the simple expedient of running an estimation-and-removal impact test identical to that described above, first with the analysis as input and then with the initialization, in place of the forecast. In both cases RAOBs provided the verification. For comparison of results, we constructed relative frequency histograms which describe the distributions of the differences in skill of the adjusted and unadjusted forecasts (analyzed/initialized fields). These are shown in Figure 4, where 4c is the histogram for the 6-hour forecast scores of Figure 3. The arrow at the base of each histogram marks 0 difference, which would be the point of symmetry in the distribution if the adjustment procedure had no effect, thus degrading the field as much and as frequently as improving it. We note, first, that the error whose estimates are used to adjust forecast values is the cumulative error of all forecast cycle components; and 4c shows that the ensemble impact of correcting for it, at the end of the cycle, is strongly positive. Evidently the analysis is not a contributor, as shown in 4a: whatever discrepancies there are between observed and analyzed values cannot be improved upon by calculating and subtracting a mean of recent analysis-minus-RAOB discrepancies. Figure 4b demonstrates that the initialization is a significant contributor,

although comparison of 4b with 4c makes it clear that this step in the forecast cycle is not the sole source of the stray.

Section 3. Stray Removal and Forecast Verification, Using Analysis-minus-Forecast Differences

In consequence of the results reported in Section 2, attention has been directed to expanding the scale of investigations of the significance of time-coherent errors and the impact of making adjustments for them, and to developing a technology suitable for routine correction of forecasts. To study the stray in the forecast cycle on a global scale and to make corrections operationally with a global model, requires the replacement of observations in their forecast verification role, with either analyzed or initialized field values. Because of the implicit exoneration of analyses and indictment of initializations as contributors to the stray, which was demonstrated in Section 2, we chose analyzed fields: a choice validated by repetition of the impact test which produced Figure 3, using analyzed geopotential values in place of RAOBs, with very similar results. (In Figure 4, compare d with c.) All subsequent work is based on analysis verifications.

Comparisons between pressure levels and latitude bands, of the impact of estimating the stray with forecast-minus-analysis differences and correcting for it prior to verification, are provided by Table 3 and Figure 5. Each entry refers to a pair of sequences of daily RMSD verification

scores for North America: one for uncorrected forecast/analysis discrepancies and one for discrepancies in which each forecast value has been adjusted by the corresponding estimate of the stray. The first entries of Table 3 are average differences between the values of the pair of daily verification plots, while the second entries give the smallest and largest differences obtained during the winter period. We note that there tends to be greater impact of the correction, in average magnitude of error reduction, at 00 GMT for 1000 and 850 mb; but that this reduction in average RMSD is larger at 12 GMT for pressure levels at and above 500 mb. Figure 5 illustrates the impact at 12 GMT for three levels, in greater detail, with histograms constructed from the daily differences. These show clearly the benefit of removing our estimate of the time-coherent discrepancy between forecast and analyzed fields, since large positive values are indicative of consistently better verifications of stray-adjusted forecasts.

To this point, the work we have reported was based on files of 6- and 12-hour forecast geopotential values and corresponding initialized, analyzed, and observed fields at the forecast verification times, at North American radiosonde station locations. Furthermore, the weights for the 7-day tapered-average stray estimate were determined by evaluating a gaussian-shape function at seven, evenly spaced points and then normalizing them to sum to 1. The decision to examine the stray as a global phenomena, with the possibility of correcting for it operationally either during data

assimilation (with 6- or 12-hour forecasts) or during post-processing (with 12- and 24-hour forecasts) required some restructuring. We began accumulation of files of 40-wave coefficients for analyzed, initialized, and forecast fields, from NMC operational forecast runs; and we replaced the gaussian taper with an exponential taper.

The choice of an exponential taper in computation of the weighted average of recent discrepancies between forecast and analyzed fields was a practical one anticipating requirements for retention of global scale files in an operational correction mode. The recursive property of an exponentially weighted average significantly reduces the volume of data which must be retained from one verification time to that 24 hours hence, because only the current file of corrections need be retained. The parameter for the exponential weighting function was adapted from the earlier gaussian-shape weighting function, with a nearly imperceptible change in impact. This parameter and the weights generated by it, are given in the Appendix, where we present details of stray estimation and correction methods.

The use of the spectral coefficients themselves for estimation and correction is a computational, time and space saving expedient. Our code produces two sets of coefficients for global, forecast-minus-analysis differences: one for the uncorrected forecast and one with the estimated stray removed. These are transformed to mandatory-level grid fields for computation of mean and RMSD performance statistics. We have replaced

verification at radiosonde station locations with grid-point verification which can be focused on any region designated by latitude and longitude coordinates. The impacts of estimating and applying corrections are then evaluated in two ways:

1) For any designated region (such as North America, Asia, Australia, or an entire hemisphere) a mean error and a root-mean-square difference are computed for each day, where the averaging is over the grid-points within the region. Thus, a pair of sequences of daily forecast verification summaries, characteristic of the region, is produced: one for the uncorrected forecast and one for the forecast with a stray estimate removed. These sequences will be denoted by (MNERR, RMSD) and (MNERR-SR, RMSD-SR), respectively.

ii) For a designated time period, on the order of a month, a mean error and a RMS analysis-minus-forecast difference is computed at each gridpoint, with averages over the days of the selected interval. This is done on a hemispheric scale (either Northern or Southern) and the resulting time-averaged verification scores are contoured from grid values. Again this is done for both the uncorrected and the stray-corrected forecast. Hemispheric maps of forecast verification scores (which we will

distinguish as ERRMAP and ERRMAP-SR) may then be compared for identification of areas of intense response to the removal of our estimate of the time coherent error.

We present the results together in the next section.

Section 4. Regional and Global Impacts of Corrections

We have created and examined daily verification sequences for five mandatory pressure levels (1000, 850, 500, 250, and 100 mb) for both hemispheres and, individually, for several regions for which uniform, high quality radiosonde coverages contribute to the verifying analyses; and we have obtained impact test results for different seasons of the year. The degree of impact of correcting for a semi-persistent error, as it is estimated by a tapered 7-day mean of the most recent discrepancies between gridpoint forecast and analysis values, varies markedly with level, region, and time of year. Nonetheless, the consensus of results is that the impact is everywhere positive, i.e., there is ensemble improvement in verifications of forecast geopotential fields, with few cases showing (small) negative impact.

a. For the North American region (80-120°W, 30-50°N), impacts of stray-removal on 6-, 12-, and 24-hour forecasts are illustrated in Figure 6, for 850 mb. This figure presents contrasting RMSD:RMSD-SR and MNERR:MNERR-SR sequences, for the same 30-day interval, early in 1988, for

each forecast period. Positive impacts on root-mean-square difference scores are virtually consistent across this time interval for 6-, 12-, and 24-hour forecasts. The latter shows the only significant negative impact (over a two-day period), although it also registers the largest improvements in verification. All three, uncorrected forecast-minus-analysis mean difference sequences (MNERRs) demonstrate significant biases. The stray-correction algorithm has removed these, on average, and the MNERR-SR sequences oscillate about zero, with relatively small excursions.

Results for other pressure levels for this North American region showed overall positive impact from stray-correction. Compared with the 850 mb results presented here, the magnitudes of differences between RMSD and RMSD-SR are similar at 1000 mb, smaller at 500 and 250 mb, and greater at 100 mb. Changes in verification scores relative to RMSD values for uncorrected forecasts are close to being the same at 1000, 850, and 100 mb, and are less at 500 and 250 mb.

b. An individual case comparison of Northern Hemisphere 24-hour-forecast errors is shown in Figure 7. Here, in a polar-stereographic projection are contour maps of f.e.s, separately for the uncorrected forecast and for the stray-corrected forecast. We see very significant areas of improvement with the correction. There are some areas of small degradation of the agreement with the verifying analysis; however they are difficult to spot in these pictures where the largest error magnitude have been rather profoundly reduced.

c. Hemispheric, 24-hour-forecast mean error comparisons, for both North and South, are shown in the next two figures. Figure 8 presents comparisons of the time-mean analysis-minus-forecast differences, with and without correction for the stray, for 850 mb; and Figure 9 shows the same comparisons for 250 mb. The figures were created with data for the same period of 31 days during October-November 1987; and the contour intervals are the same throughout.

At both pressure levels, in Northern and Southern hemispheres, comparisons between ERRMAP and ERRMAP-SR show strong reductions in mean analysis/forecast discrepancies. These changes in verification with removal of the stray, are most evident and profound in the areas of greatest (original) error intensity.

d. The only region spanning the equator with sufficient radiosonde coverage to give credence to tropical analysis verifications (if, indeed, any may be believed) is Indonesia. We defined this region as (210-270°W, 15°S-15°N), and computed daily verification scores for 24-hour forecasts, over the same Fall interval as in Section 4.c, above. Figure 10 illustrates the impacts on 500 mb scores, with plots of both root-mean-square differences and mean differences. Inspection shows that the very significant mean bias has been removed by the correction algorithm and the RMSD-SR is less than half the RMSD, throughout this October-November period.

e. Hemispheric forecast "skill scores" as they are routinely calculated and exhibited at the NMC, for monitoring 1000 and 500 mb forecast performance, have been adapted to put some of our comparisons in a familiar format. Thus the regions submitted for sequential scoring of forecast verifications were the hemispheres: 20 to 80°N and 20-80°S. Admittedly, the results for these regions may suffer from the inclusion of substantial areas in which there is no radiosonde contribution to the analyses used for verification. (This is by far the greatest portion of the Southern Hemisphere.) Nonetheless, the analyses are our best measure of 'truth'; and our results indicate that the information in them has a highly significant "corrective" impact on 6-, 12-, and 24-hour forecasts. Hemispheric comparisons of RMS differences and mean errors, for 00 GMT 24-hour-forecasts of 500 mb geopotential, are provided in Figure 11. In both hemispheres the RMSD-SR values are consistently and significantly less than corresponding RMSD verification scores, and the overall biases with respect to the analyses have been removed by the adjustments for the stray.

f. Results of impact tests using analysis verifications of 24-hour forecasts for the Australian region (210-250°W, 10-35°S), registered the greatest impact of stray-correction at 500 mb; although, here again, it was almost consistently positive throughout the interval from 24 March to 24 April 1988, on the five levels of our tests. Figure 12 illustrates this, and shows that some verification score reductions are as great as 50%.

Section 5. Experiments with Correction Procedures

The work discussed thus far has compared forecast geopotential values with radiosonde reports, with analyzed values, and with initialized values. Mean differences and root-mean-square differences of forecast values, with and without the stray removed, from observed, analyzed, and initialized geopotential values, indicate that significant benefit will accrue from making the correction routinely, provided the analysis is used as the global representation for the atmospheric "bulk state signal".

We examined evidence of the impact of removing the geopotential stray at five mandatory pressure levels (from the 1000 mb surface to 100 mb), for 6-, 12-, and 24-hour forecasts; and we demonstrated that a short-term history can produce a current-time estimate which may be utilized to tune forecast model output more closely to geopotential analyses. Our investigation did not lend insight to the sources of time-coherent error within the forecast cycle, except to the limited extent that we have shown a portion to be associated with the initialization and no evidence for any contribution from the analysis step.

Our purpose in this work was to determine whether there is a significant, correctable 'stray', and to consider possible mechanisms for its routine removal. The evidence is unequivocal concerning the presence of this semi-persistent error; and two correction mechanisms suggest themselves. The first is making a correction during data assimilation, by subtracting a

current error-estimate which can be carried as a single field and updated as each forecast field is differenced with its verifying analysis. (See the Appendix for details.) Our discovery that the initialization process is party to the generation of time coherent errors, suggests that removing the stray from an analysis, which is then submitted to the initialization, may not be fully effective. The second is making the correction at the end of the forecast cycle. NMC's parallel global forecast capability made it possible to apply and verify the impacts of the two mechanisms, over a single time period.

The parallel forecast cycle is designed to be identical to the operational cycle in all aspects except the one being tested. Thus, during the May-June 1988 period of our tests, the operational version of the global data assimilation system (GDAS) and its experimental version (GDAX) were the same, except that a current estimate of the stray was maintained and used to correct all GDAX 6-hour forecast fields during every data assimilation step of the parallel forecast cycle. In fact we had three distinct test periods, each with a 7-day start-up to re-initialize the stray estimate. The three periods differed in that during:

May 5-May 26 - full spectral coefficient fields of 6-hour forecast-minus-analysis differences were used to estimate and correct each forecast as it was brought into the data assimilation suite;

May 27-June 10 – on each of the 12-mandatory levels a 10x10 coefficient field, corresponding to the longest waves, was used to estimate and correct for the forecast-cycle stray;

June 11-June 25 – differences between initialized and time-coincident 6-hour forecast fields were truncated by including only coefficients for the 15 longest waves, to define the cumulative elements of corrections for the stray.

Routine comparisons of the performances of parallel and operational forecasts include means, root-mean-squares, and anomally correlations of the differences between forecasts and verifying analyses at 1000 and 500 mb, with particular emphasis on five-day forecasts. For our tests we supplemented these with RAOB verifications of 24-hour forecasts of wind and geopotential, in both hemispheres, for 850, 500, 250, and 100 mb levels.

During the first test period, maps of analyzed wind and geopotential fields produced by the two systems were compared daily, and their discrepancies were observed to increase, particularly for winds in the Southern Hemisphere. Except at 100 mb, verifications of 24-hour forecasts from the parallel forecast cycle tend to deteriorate relative to those of the operational forecast cycle, through the test period. The performance comparison based on anomally correlations for the 5-day forecast from the

two systems gives a somewhat different impression, as shown in Figure 13. Towards the end of the period the GDAX seems to improve the forecast, although the differences are not great by this measure and the overall evidence is that the GDAX leads to an inferior 5-day forecast. Thus the results do not recommend this mode of correction.

The change made, going into the second test period was suggested by the "noisy" appearance of the later analyses of the first period, and by a conclusion of Saha and Albert (1988) that only the longest waves are useful in correcting systematic forecast error. Accordingly, on each mandatory level, only a 10x10 coefficient component of the correction estimate was maintained and applied to the "first guess used" in each GDAX analysis. The impact measure provided by comparative 5-day forecast anomally correlations for 500 mb geopotential shows the parallel model to have a slight edge on the operational. However RAOB verifications of 24-hour forecasts of geopotential and vector winds indicate overall deterioration of the parallel forecast products, in both hemispheres, during this short period (June 5-11). As seen in Table 4a and b, there are some exceptions - notably at 100 mb - where some operational/parallel verification scores show remarkable discrepancies, in both directions. However, once again, the results do not constitute a basis for recommending the parallel test mode for routine correction of the stray in the forecast cycle.

During the final test period the "first guess" entering the data assimilation suite of GDAX was corrected prior to post-processing by a tapered average of previous forecast-minus-initialized field differences, using the first 15 wave numbers. The rationale for the choice of the limit on wavenumbers derives from information displayed in Figure 14. Here, for each of five verification levels, is a plot of Northern Hemisphere average RMS error improvement for 24-hour forecast-analysis differences, achieved by correcting each operational forecast by an estimated stray component, versus the number of waves used in computing the estimate. We see that a 15-wave rhomboidal truncation includes about 85% of total achievable improvement, while excluding the shorter noisier waves. The rationale for using the initialized analysis to define the stray derives from the fact that the initialization procedure balances the geopotential and wind fields. Accordingly, stray estimates generated from previous forecast-minus-initialized field differences could be expected to be balanced and not to create undesirable features in the resulting analyzed fields. Nonetheless, during this final test, operational and parallel forecast performance scores were very similar. In the Northern Hemisphere, RAOB verification scores for 100 mb geopotential forecasts, from Table 5a, show the parallel out-performing the operational cycle. However, this is not true for the vector winds at that level; the results are mixed at other levels; and the scores in Table 5b show the parallel forecasts to be generally worse than those of the operational cycle, throughout the period, for all levels in the Southern Hemisphere.

The fact that the forecasts from the GDAX, during the three test periods, did not register overall improvement relative to forecasts from the GDAS, in part confirms earlier findings. Analyzed geopotential fields are not systematically biased with respect to radiosonde observations, although their initialized fields and forecasts made from them, are. Thus applying a systematic-error-correction to the forecast going into the construction of an objectively analyzed field is redundant: the objective analysis makes its own correction, by interpolation of discrepancies between forecast and observed values. Although the interplay of components of the forecast cycle is not completely in our view, we would have been surprised by any other results than those we obtained.

Although perhaps not as aesthetic as applying corrections to the "first guess" during data assimilation, the ultimate point for correction is during post-processing of longer period forecast products. For contrast with the May-June test results reported above, we carried out 24-hour forecast correction impact tests over a thirty-day interval during May-June 1988. Our forecast performance measure is a root-mean-square difference (RMSD) of forecast and verifying analysis values, where the averaging is over the grid points of the region used for impact assessment. Day-by-day plots of RMSDs for corrected and for uncorrected geopotential forecasts demonstrate the impact of adjusting with (full field) estimates of the stray. Figure 15 illustrates very consistent impacts, at 100 mb in both hemispheres: regions which we have bounded by latitudes 20 and 80°. For the Northern Hemisphere the 500 mb mean error comparison, for uncorrected

versus corrected geopotential forecasts, is shown in Figure 16; and Figure 17 illustrates the corresponding stray-correction-impact with RMSD and mean (24-hour forecast minus verifying analysis) difference plots for the North American region. We see that the mean error is profoundly reduced by daily adjustment for the stray in the 24-hour forecast product, for the entire hemisphere. With three minor exceptions, North American daily RMS differences show very significant reductions over the 30-day period; and the mean errors over this continent have been brought down to oscillate about zero.

Section 6. Overview

At the U.S. National Meteorological Center we have established that temporally coherent departures of 6-, 12-, and 24-hour forecasts of the global spectral model from observed and analyzed geopotential fields can be estimated, and the estimates used beneficially for adjustment of the forecast fields at the time of issue. Our diagnostic techniques have brought together evidence that:

a. the forecast cycle stray is an oscillatory phenomenon with large variations over a season, which is most effectively estimated with short-interval averages;

b. it is a global phenomenon apparently present in all seasons; and

c. although the analysis component of the forecast cycle does not contribute, the initialization step is a major source of these systematic errors - which involve a large range of wave numbers.

We have evaluated the effects, on forecast performance scores, of making corrections for these systematic errors. Short-term (7-day) tapered averages of the most recent discrepancies between forecast and observed, or forecast and analyzed values provide current estimates of the stray, and adjustment of forecast products by these increments has been shown to have positive impact in all three seasons of our study. With occasional exceptions of relatively small negative impact, forecast verification scores were significantly improved by this mechanism, for both hemispheres and from the surface well into the stratosphere.

Parallel forecast cycle tests were carried out to evaluate the impact on forecast product accuracy of applying a correction to the "first guess" going into the data assimilation suite. However the results did not establish this method of adjusting for the stray in the forecast cycle as a generally beneficial alternative. By contrast, correction of the 24-hour geopotential forecasts made from the global data assimilation system, at the time of issue, had an overwhelming clear impact on the agreement between forecasts and verifying analyses. Thus, the technique of correcting the forecast product during post-processing, rather than attempting to make statistical adjustments in the analysis step, is

identified as an effective control for the time- and space-coherent errors brought through the forecast cycle.

APPENDIX: Details of the Technology of Stray Estimation Removal

A. The Technique

We require a notation to designate elements of the problem. Because present work concerns only one atmospheric variable, the letters f, $\mathcal{O}$ and Q will suffice to designate forecast, observed, and analyzed values of this variable. With the following elements distinguished as:

$\mu \equiv$ the signal component (or "bulk state signal") of the atmospheric field,

$X \equiv$ small scale departures of the atmosphere from its own bulk state, together with contributions to observed values of measurement and data-transmission errors, assumed to have zero mean,

$S \equiv$ the stray of the forecast cycle, as illustrated in Figure 4d,

$Y \equiv$ departures of the forecast from $\mu + S$, which are O mean errors,

observed and forecast fields will be represented in these terms as

$$\mathcal{O} = \mu + X \qquad f = \mu + S + Y$$

Thus, observation reports will be treated as a combination of the signal information we wish to use and unavoidable contaminations which we assume will average to zero, and forecast values treated as combinations of the same signal plus zero mean forecast errors plus forecast cycle stray.

The difference between a forecast and its verifying observed field:

$$f - o = (\mu + Y + S) - (\mu + X) = (Y - X) + S$$

gives us the basis for estimating and correcting for the stray, by virtue of the statistical properties of X and Y, and the time coherence of S. To make this explicit we introduce an index for analysis (forecast verification) times, to distinguish the present, t, from analysis times at 24-hour intervals preceding the present: $t-k$ for $k = 1,2,\ldots$. Because of the persistence in S, a tapered mean of the most recent values, namely,

$$\sum_1^K \alpha_k S_{t-k} \quad \text{for} \quad \alpha_1, \ldots, \alpha_K \quad \text{and} \quad \sum_1^K \alpha_k = 1.$$

will approximate the value of the present stray, S_t. Furthermore, since X and Y represent small scale variability plus observation error and forecast error, with zero means, then

$$\sum_1^K \alpha_k X_{t-k} \approx 0 \quad \text{and} \quad \sum_1^K \alpha_k Y_{t-k} \approx 0.$$

Consequently, the tapered mean of recent forecast errors is

$$\sum_{1}^{K} \alpha_k \left(f_{t-k} - \theta_{t-k} \right) = \sum_{1}^{K} \alpha_k \left[\left(Y_{t-k} - X_{t-k} \right) + S_{t-k} \right]$$

$$\approx 0 + 0 + S_t .$$

or, turning this around,

$$\hat{S}_t = \sum_{1}^{K} \alpha_k \left(f_{t-k} - \theta_{t-k} \right).$$

We write the present, stray-corrected forecast as

$$f_t^c = f_t - \hat{S}_t .$$

In this notation it may be regarded as a full, tensor field, or as the scalar value at a specific location, which we have neglected to index. We make use of both: the first in the context of correcting a global forecast product and the second in the context of correcting analyses during data assimilation.

Very simply, correction of the analysis during each data assimilation is achieved by constructing and maintaining a single, current, full-field tapered mean of the most recent discrepancies between forecast and analyzed fields, with which the forecast (which serves as the guess for the current analysis) is corrected. Here we have substituted the analyzed field for the observed field, to provide verification and correction everywhere the forecast is required.

B. The Taper

Early in our investigations we made use of a taper with $\underset{\sim}{\alpha} = (.22, .21, .19, .15, .11, .07, .04)$ obtained by evaluating a guassian curve at even intervals, beginning at 0. This was the basis for estimation and removal of the stray as it was done in the production of the results in Section 2 of this paper. For use in an operational context, with attendant concerns for computer space and computational time, the gaussian taper was replaced with a recursive exponential taper. Specifically, we write $\alpha_k = \beta(1-\beta)^{k-1}$ for $k = 1, 2, \ldots$, with β selected to give the first seven values close to those of the array $\underset{\sim}{\alpha}$, above. The practical savings, in requirements for data retention and computation, are possible because once we have the stray estimate

$$\hat{S}_t = \sum_1 \beta(1-\beta)^{k-1} \Delta_{t-k} \qquad \text{where} \quad \Delta_{t-k} = f_{t-k} - a_{t-k},$$

the subsequent estimate is computed from it, together with the next forecast error $\Delta_t = f_t - a_t$:

$$\begin{aligned} \hat{S}_{t+1} &= \sum_1 \beta(1-\beta)^{k-1} \Delta_{(t+1)-k} \\ &= \beta \Delta_t + \sum_2 \beta(1-\beta)^{k-1} \Delta_{t-(k-1)} \\ &= \beta \Delta_t + (1-\beta) \sum_1 \beta(1-\beta)^{k-1} \Delta_{t-k} \\ &= \beta \Delta_t + (1-\beta) \hat{S}_t . \end{aligned}$$

Thus, the stray-correction file is updated with each forecast verification.

Histories of forecast-minus-analysis discrepancies more than seven days removed from the present, receive negligible weight, because of our choice of $\beta = .23$ to closely match $\underset{\sim}{\alpha}$.

Figure 1. Forecast-minus-RAOB average values, for 6-hour forecasts verifying at 00 and 12 GMT. Averaging is over December 1984 through February 1985.

Figure 2. Schematic representation of time evolution of geopotential observations (+), the atmosphere's "bulk state signal" (——), ideal forecast (....), actual forecast values (- -), corrected forecast (—.—.).

Figure 3. Daily, North American root-mean-square differences between observed and forecast 850 mb geopotential values, for 12 GMT. 0 denotes RMSD's for 12-hour forecast values, * denotes RMSD's for stray-corrected 12-hour forecasts, ... denotes RMSD's for 6-hour forecast values, __ denotes RMSD's for stray-corrected 6-hour forecasts.

Figure 4. Frequency histograms for the differences between RMS observation-minus-computed values, without and with stray-correction, for (a) analysis, (b) initialization, and (c) 6-hour forecast. 4d is the corresponding histogram for analysis-minus-forecast differences, without and with stray-correction.

Figure 5. Frequency histograms for the differences between RMS discrepancies of analyses and 6-hour forecasts, without and with stray-correction, at 12 GMT for 850, 500, and 250 mb, and North American latitude bands 0-30°N, 25-55°N, and 50-90°N.

Figure 6. Verification scores for 850 mb North American 00 GMT, analysis-minus-forecast values without (__) and with (...) stray removal, for (a) 6-hour forecasts, (b) 12-hour forecasts, (c) 24-hour forecasts.

Figure 7. (a) Difference between the 00 GMT March 4 verifying analysis and the 24-hour, 500 mb geopotential forecast for the Northern Hemisphere. (b) Same as (a) with the forecast corrected by subtracting a 7-day, tapered mean estimate of the stray. Contour interval is 10 m.

Figure 8. Time-mean analysis-minus-forecast differences, for 24-hour forecasts of 850 mb geopotential, October 27 through November 26, 1987. (a/b) Southern Hemisphere without/with stray-correction.

Figure 9. Time-mean analysis-minus-forecast differences, for 24-hour forecasts of 250 mb geopotential, October 27 through November 26, 1987. (a/b) Southern Hemisphere without/with stray-correction.

Figure 10. (a) RMSD and RMSD-SR verification scores for 24-hour forecasts of 500 mb geopotential verifying at 00 GMT, October-November 1987; (b) MNERR and MNERR-SR verification scores, corresponding to 9b.; (c) relative frequency histogram of RMSD/RMSD-SR discrepancies of 9b; and (d) Indonesian Region radiosonde locations.

Figure 11. Verification scores for 24-hour forecasts of 500 mb geopotential, for Northern and Southern Hemispheres: top and bottom, respectively.

Figure 12. RMSD and RMSD-SR verification scores for 24-hour forecasts of 500 mb geopotential, for the Australian region, during the interval March 24 through April 24, 1988.

Figure 13. Northern Hemisphere (20°-80°) anomally correlations for 5-day forecasts made from the operational global data assimilation system (GDAS), indicated by values connected with solid lines, and for forecasts made from the experimental system (GDAX), indicated with dotted lines. On the abscissa are the forecast verification dates, with breaks between valid verification periods.

Figure 14. Average differences between daily RMSD curves for uncorrected 24-hour geopotential forecasts and forecasts adjusted with stray estimates computed from rhomboidal truncations (to n waves) of recent forecast-minus-analysis discrepancies. Impact assessments are for the 30-day period April 2 to May 2, 1988, for North America, for five levels: 1000 mb —+—; 850 --△--; 500 —⊖—; 250 ···+···; and 100 —·▲·— .

Figure 15a. Daily, Northern Hemisphere root-mean-square differences between analyzed and 24-hour forecast 100 mb geopotential values, for 00 GMT, -- denotes the uncorrected forecast and .... denotes the forecast corrected with an exponentially tapered mean of prior analysis-forecast discrepancies.

Figure 15b. Same, for the Southern Hemisphere.

Figure 16. Time-mean analysis-minus-forecast differences, for 24-hour forecasts of 500 mb geopotential, May 26 through June 23, 1988. (a/b) Northern Hemisphere without/with stray correction.

Figure 17. (a) RMSD and RMSD-SR verification scores for 24-hour forecasts of 500 mb geopotential verifying at 00 GMT, May-June 1988. (b) MNERR and MNERR-SR verification scores, corresponding to (a).

REFERENCES

Arpe, K. and E. Klinker (1986): "Systematic errors of the ECMWF operational forecasting model in mid-latitudes". Quart. J. R. Met. Soc., 112, 181-202.

Bennett, A. F. and L. M. Leslie (1981): "Statistical correction of the Australian region primitive equation model. Mon. Wea. Rev., 109, 453-462.

_____, and _____ (1982): "Statistical correction of thickness height prognoses". Aust. Meteor. Mag., 30, 201-203.

_____. and _____ (1983): "Statistical correction of dynamical prognoses: the decision problem". Mon. Wea. Rev., 111, 343-352.

Dalcher, A. and E. Kalnay (1987): "Error growth and predictabliliy in operational ECMWF forecasts". Tellus, 39A, 474-491.

Saha, S. and J. Alpert (1988): "Systematic errors in NMC medium range forecasts and their correction". Eighth Conference on Numerical Weather Prediction. Baltimore, Maryland, February 22-26, pp. 472-477.

ACKNOWLEDGEMENTS

Much of this work was done while Dr. Thiébaux was supported by the UCAR Visiting Scientist Program at NMC.

Table 1. Time-averaged forecast-error values of 500 mb geopotential at 00 GMT for seven North American radiosonde stations. Consecutive means were computed with data from non-overlapping intervals, from an 81-day record. Top values are 10-day means; lower values are 5-day means.

LON/LAT								
45/61	-39	-34	-19	- 6	+ 1	+12	+ 8	+14
	-33 -44	-30 -37	-25 -12	-14 3	- 5 5	16 7	5 14	27 - 1
73/42	- 3	- 5	- 2	+ 5	+ 7	+ 1	+ 7	+ 4
	-12 5	- 5 - 5	7 -13	8 3	9 6	-10 14	1 12	25 -22
11/47	-15	-16	-10	-15	- 1	- 3	10	9
	-25 - 4	-13 -18	-16 - 2	-11 -18	- 1 - 2	6 -13	7 13	14 3
118/47	11	-17	- 6	- 1	14	6	12	3
	10 12	-21 -15	-14 5	- 5 3	20 11	1 11	8 16	13 - 9
124/47	9	-19	-14	7	15	9	7	21
	4 14	-21 -17	-11 -20	- 2 14	15 15	13 3	4 10	31 10
165/64	-13	-22	- 7	7	17	9	-14	- 6
	-13 -13	-16 -28	-14 1	10 4	24 9	8 10	-19 -10	- 6 - 5
79/8	20	14	21	18	11	7	14	1
	17 22	20 9	16 27	14 21	18 6	0 14	18 1	14 -15

Table 2. 500 mb, winter, forecast-error averages, for consecutive non-overlapping intervals of decreasing length, corresponding to the first line of Table 1.

No.s of day in averages											
81						−8					
30				−31			+1				
15			−37	−25		−6	+9	+16			
10		−39	−34	−19	−6		+1	+12	+8	+14	
7	−28	−45	−46	−2	−19	−2	4	15	9	5	13

Table 3. Average differences between RMS discrepancies of analyses and 6-hour forecasts, without and with stray-correction, and (minimum,maximum) values of these differences, for North America during the winter of 1986-87.

P Latitude Band	00 GMT Average Difference	(min, max)	12 GMT Average Difference	(min, max)
1000 mb				
0 - 30	6.83	(- 4., +13.)	3.63	(- 1., + 9.)
25 - 55	6.53	(- 4., +13.)	2.55	(- 4., + 9.)
50 - 90	1.68	(- 4., + 8.)	2.73	(- 4., +11.)
850 mb				
0 - 30	6.04	(- 7., +12.)	4.98	(- 4., +11.)
25 - 55	6.49	(- 6., +14.)	6.46	(- 4., +13.)
50 - 90	2.48	(- 6., + 8.)	3.35	(- 5., + 8.)
500 mb				
0 - 30	2.23	(- 9., + 8.)	4.67	(- 6., +13.)
25 - 55	0.96	(- 4., + 7.)	3.93	(- 7., +14.)
50 - 90	1.88	(- 6., + 9.)	2.45	(- 5., +11.)
250 mb				
0 - 30	1.13	(- 8., +10.)	6.33	(-11., +17.)
25 - 55	1.36	(- 7., +11.)	7.03	(- 7., +22.)
50 - 90	1.49	(- 7., + 7.)	3.05	(- 6., +11.)
100 mb				
0 - 30	2.32	(-13., +17.)	4.48	(-19., +21.)
25 - 55	3.29	(- 5., +17.)	10.20	(- 8., +28.)
50 - 90	5.41	(- 7., +19.)	5.25	(- 8., +19.)

Table 4. RMSD's for 24-hour forecasts and verifying radiosonde observations of geopotential and vector winds, for the operational (parallel) forecast cycles, during the second test period.

a. Northern Hemisphere (20-80°N)

		850 MB		500 MB		250 MB		100 MB	
		Z	V	Z	V	Z	V	Z	V
June	5	15.1 (15.1)	5.3 (5.3)	22.8 (24.0)	6.2 (6.2)	36.6 (38.2)	9.0 (9.0)	51.3 (51.4)	5.6 (5.5)
	6	19.9 (18.8)	4.9 (4.9)	27.9 (27.3)	6.1 (6.1)	41.8 (40.3)	10.6 (10.5)	63.4 (63.0)	5.0 (5.6)
	7	17.9 (18.7)	4.9 (5.0)	23.7 (23.9)	5.8 (5.9)	35.8 (35.4)	8.0 (7.8)	51.8 (51.0)	4.7 (4.9)
	8	16.3 (16.2)	4.8 (4.8)	23.0 (22.8)	5.6 (5.5)	35.5 (35.3)	7.8 (7.8)	51.7 (50.3)	5.1 (5.1)
	9	18.0 (18.1)	5.2 (5.2)	23.1 (22.4)	7.2 (7.2)	39.0 (38.4)	8.6 (8.7)	52.4 (50.4)	5.1 (5.4)
	10	18.3 (17.6)	5.2 (5.1)	26.1 (26.0)	5.8 (5.9)	40.6 (41.2)	9.5 (9.5)	51.2 (51.4)	5.1 (5.0)
	11	14.7 (14.6)	4.9 (4.9)	22.4 (23.0)	6.5 (6.7)	36.0 (35.8)	8.5 (8.5)	55.7 (55.0)	6.5 (6.4)

b. Southern Hemisphere (20-80°S)

June	5	25.8 (33.0)	6.3 (5.3)	24.1 (29.8)	6.0 (6.8)	49.9 (60.0)	9.3 (9.0)	93.1 (66.8)	11.6 (11.3)
	6	25.5 (36.7)	10.7 (12.1)	37.7 (51.8)	11.3 (11.8)	53.9 (61.5)	14.5 (14.4)	42.4 (34.3)	10.7 (10.4)
	7	32.8 (39.6)	8.2 (7.9)	46.2 (59.6)	8.4 (8.7)	63.8 (80.3)	11.4 (11.6)	78.9 (74.3)	10.2 (9.3)
	8	26.3 (26.1)	5.3 (6.8)	38.9 (50.6)	8.7 (10.7)	52.0 (71.2)	12.8 (13.7)	68.3 (65.6)	13.9 (12.8)
	9	21.7 (23.0)	5.7 (6.3)	22.2 (24.8)	6.7 (7.0)	35.5 (47.7)	9.9 (9.4)	23.7 (34.4)	7.8 (5.6)
	10	25.7 (30.5)	4.6 (5.6)	22.9 (32.9)	7.5 (8.2)	28.0 (38.9)	11.6 (13.2)	29.0 (38.2)	9.1 (7.6)
	11	23.6 (28.1)	4.7 (4.5)	34.9 (34.5)	8.5 (8.2)	69.0 (71.4)	13.0 (13.0)	46.4 (73.9)	9.8 (8.6)

Table 5. RMSD's for 24-hour forecasts and verifying radiosonde observations of geopotential and vector winds, for the operational (parallel) forecast cycles, during the third test period.

a. Northern Hemisphere (20-80°N)

	850 MB		500 MB		250 MB		100 MB	
	Z	V	Z	V	Z	V	Z	V
June 16	15.6 (16.3)	5.7 (5.8)	23.9 (24.2)	6.6 (6.7)	43.0 (44.5)	8.9 (9.1)	54.8 (54.9)	5.2 (5.2)
17	15.3 (15.2)	5.6 (5.5)	20.2 (20.0)	5.8 (5.8)	31.7 (32.9)	8.8 (8.9)	45.0 (43.9)	5.6 (5.4)
18	13.5 (13.8)	4.1 (4.1)	22.0 (21.9)	5.6 (5.7)	40.0 (40.2)	8.1 (8.3)	58.5 (57.9)	5.2 (5.4)
19	16.1 (16.4)	4.7 (4.9)	25.2 (24.9)	6.1 (6.1)	42.1 (42.5)	9.6 (9.8)	56.5 (54.9)	5.7 (5.8)
20	13.7 (13.9)	4.9 (4.9)	21.8 (21.0)	5.4 (5.4)	40.2 (38.6)	8.3 (8.2)	50.5 (49.3)	4.9 (5.2)
21	15.1 (15.6)	4.8 (4.8)	23.5 (23.1)	6.1 (6.2)	42.9 (42.4)	9.4 (9.5)	59.3 (56.9)	5.1 (5.2)
22	16.0 (16.8)	5.2 (5.3)	24.0 (24.9)	6.1 (6.4)	35.8 (35.7)	8.6 (8.8)	50.6 (49.2)	5.3 (5.3)

b. Southern Hemisphere (20-80°S)

June 16	24.3 (26.2)	6.1 (5.8)	29.7 (29.7)	7.4 (7.2)	41.3 (40.8)	15.5 (15.4)	41.6 (42.1)	7.0 (7.1)
17	24.7 (27.9)	5.0 (5.2)	25.5 (26.5)	8.1 (8.6)	36.0 (38.1)	10.6 (10.8)	38.9 (39.2)	7.8 (7.7)
18	15.7 (24.3)	6.5 (6.9)	28.2 (30.8)	7.7 (8.7)	43.0 (45.9)	12.9 (13.9)	35.1 (41.6)	14.2 (14.4)
19	10.9 (12.3)	5.1 (5.4)	20.1 (19.3)	6.3 (6.1)	29.7 (36.4)	11.7 (10.8)	34.6 (51.7)	8.0 (7.7)
20	23.9 (20.5)	6.8 (6.9)	20.8 (20.5)	9.2 (9.5)	37.3 (31.3)	11.4 (12.0)	28.6 (28.8)	7.6 (8.1)
21	19.7 (24.8)	5.1 (5.7)	26.6 (28.9)	10.7 (10.5)	53.3 (48.0)	10.6 (11.3)	33.9 (36.8)	10.9 (8.9)
22	29.9 (29.3)	7.0 (7.2)	41.8 (44.8)	8.2 (8.7)	59.7 (62.7)	14.6 (15.8)	58.3 (58.3)	11.3 (11.2)

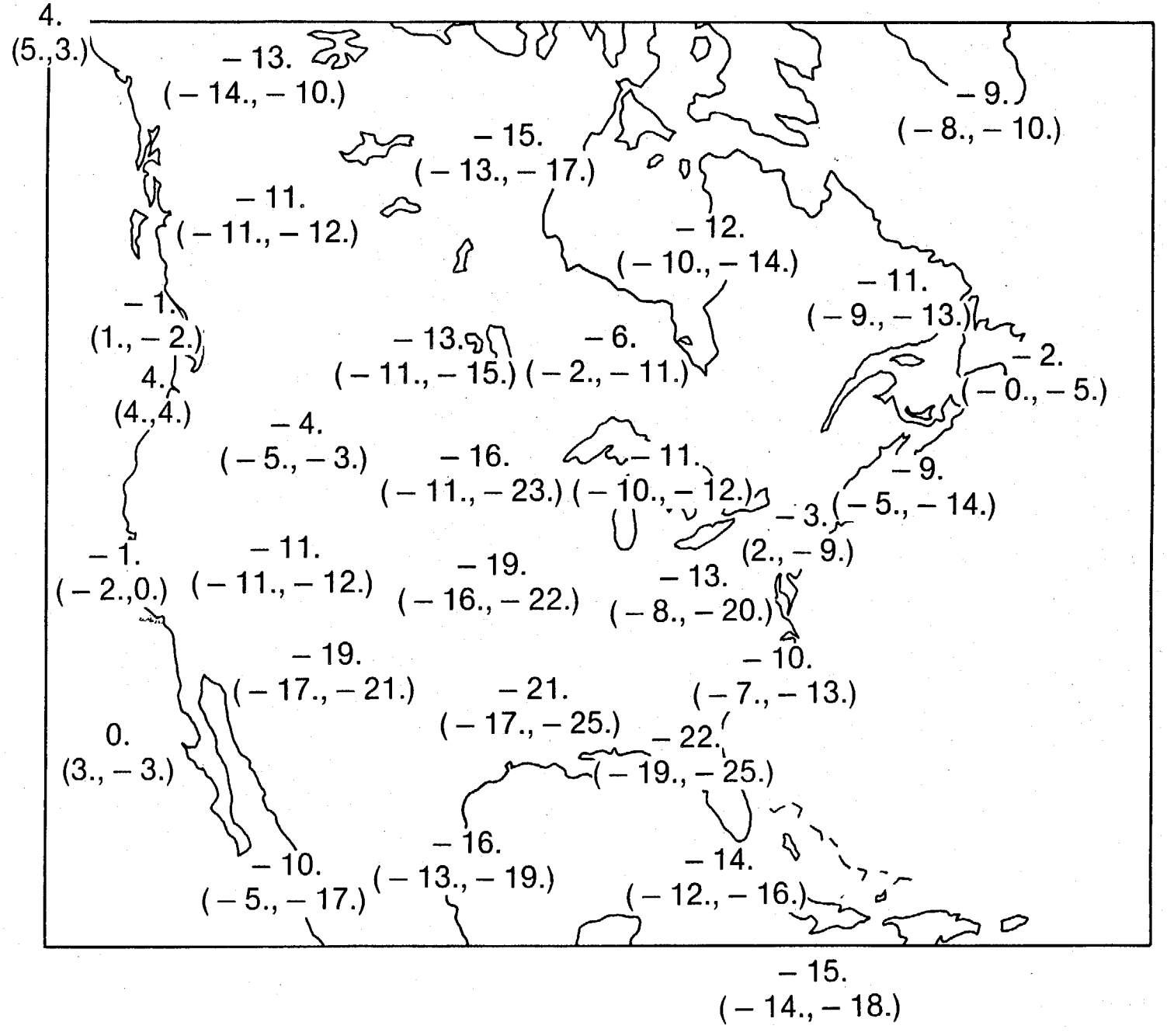
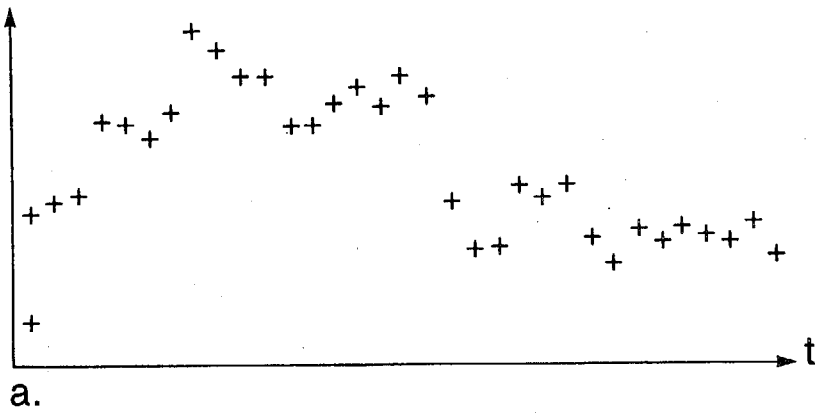
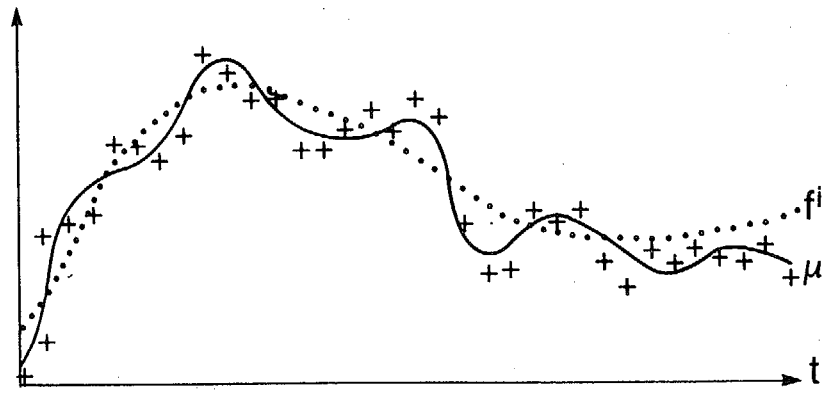

4.
(5.,3.)
– 13.
(– 14.,– 10.)
– 15.
(– 13.,– 17.)
– 9.
(– 8.,– 10.)
– 11.
(– 11.,– 12.)
– 12.
(– 10.,– 14.)
– 11.
(– 9.,– 13.)
– 1.
(1.,– 2.)
– 13.
(– 11.,– 15.)
– 6.
(– 2.,– 11.)
– 2.
(– 0.,– 5.)
4.
(4.,4.)
– 4.
(– 5.,– 3.)
– 16.
(– 11.,– 23.)
– 11.
(– 10.,– 12.)
– 9.
(– 5.,– 14.)
– 3.
(2.,– 9.)
– 1.
(– 2.,0.)
– 11.
(– 11.,– 12.)
– 19.
(– 16.,– 22.)
– 13.
(– 8.,– 20.)
– 19.
(– 17.,– 21.)
– 10.
(– 7.,– 13.)
– 21.
(– 17.,– 25.)
0.
(3.,– 3.)
– 22.
(– 19.,– 25.)
– 16.
(– 13.,– 19.)
– 10.
(– 5.,– 17.)
– 14.
(– 12.,– 16.)
– 15.
(– 14.,– 18.)

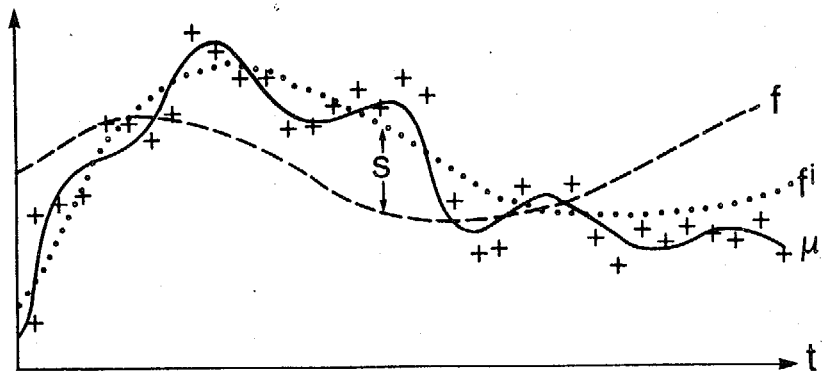
fig 1.



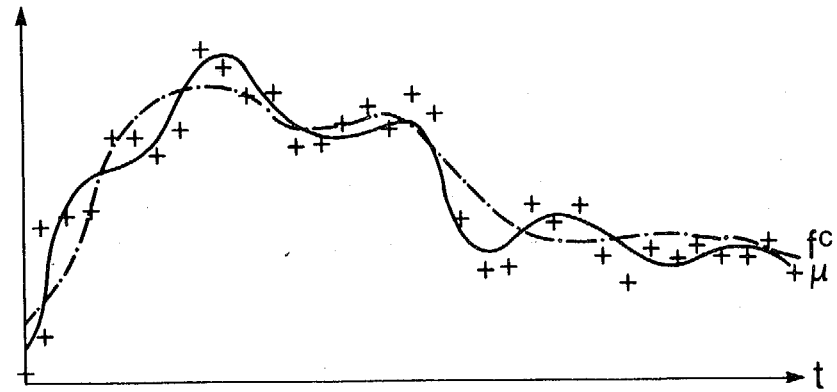
a.



b.



c.



d.

fig. 2.

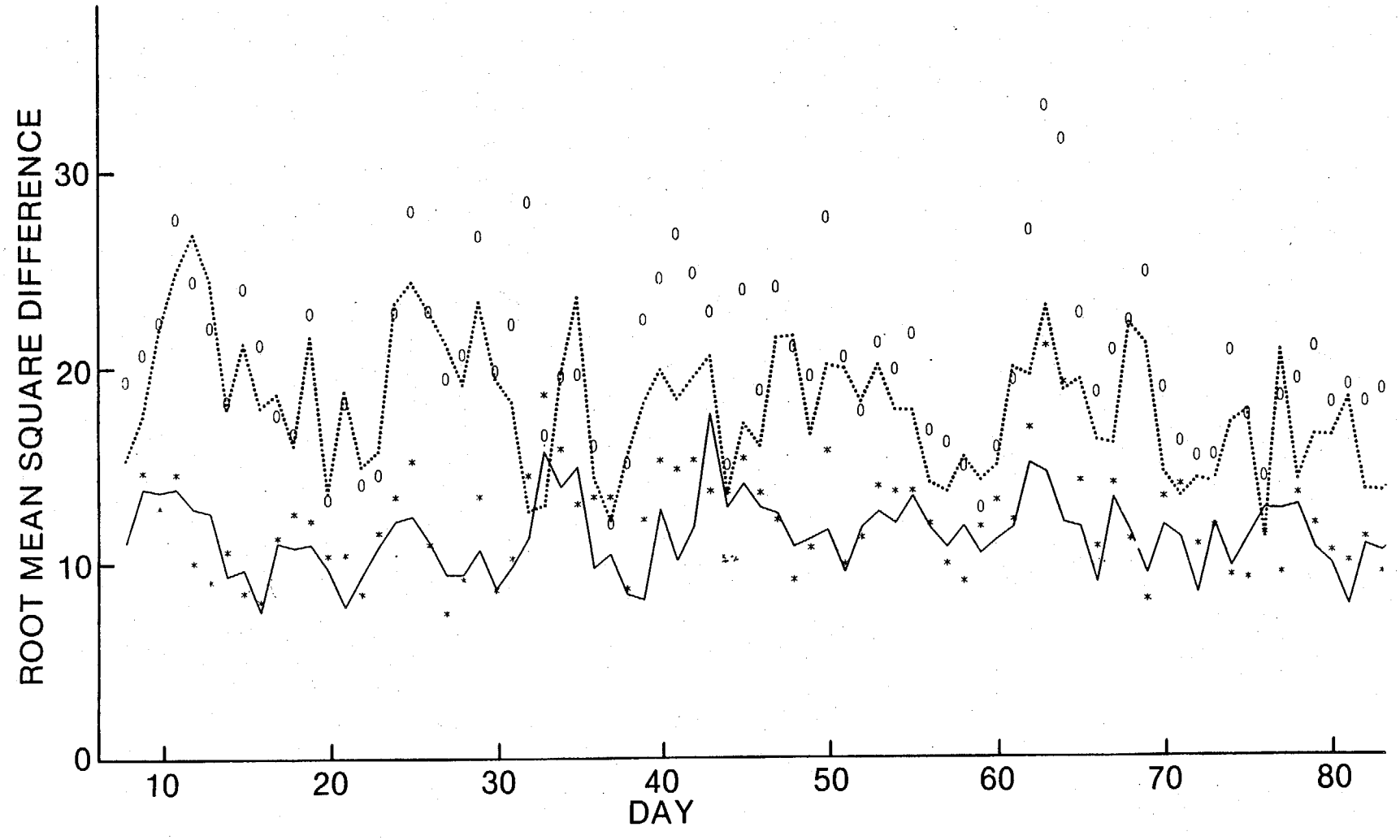
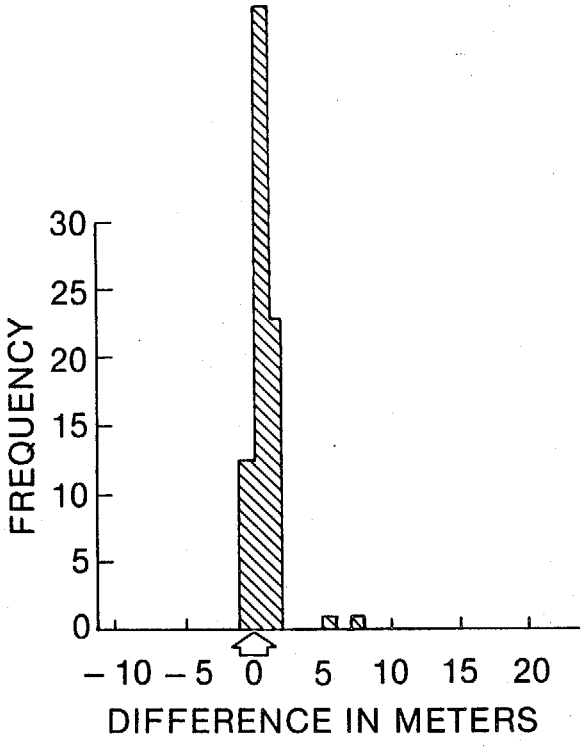
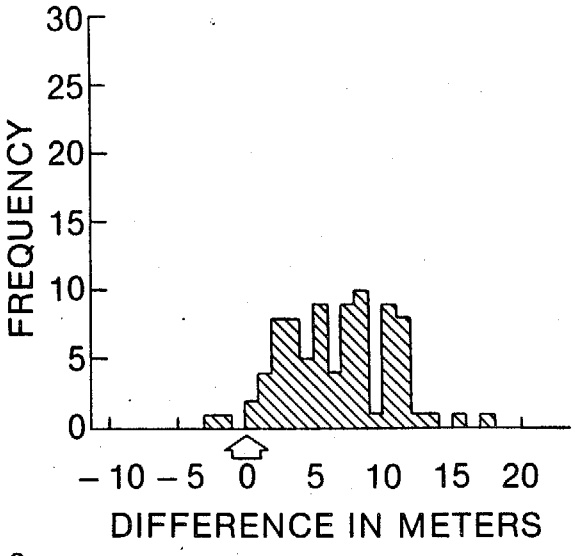

ROOT MEAN SQUARE DIFFERENCE
0
10
20
30
10
20
30
40
50
60
70
80
DAY

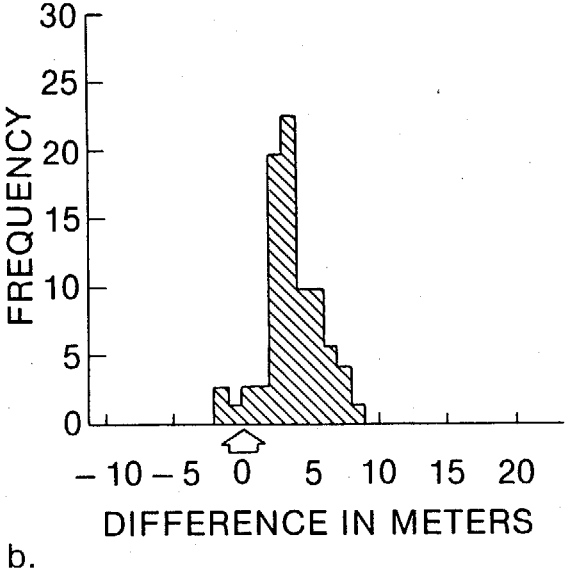
Fig. 3


30
25
20
15
10
5
0
FREQUENCY
– 10 – 5 0 5 10 15 20
DIFFERENCE IN METERS

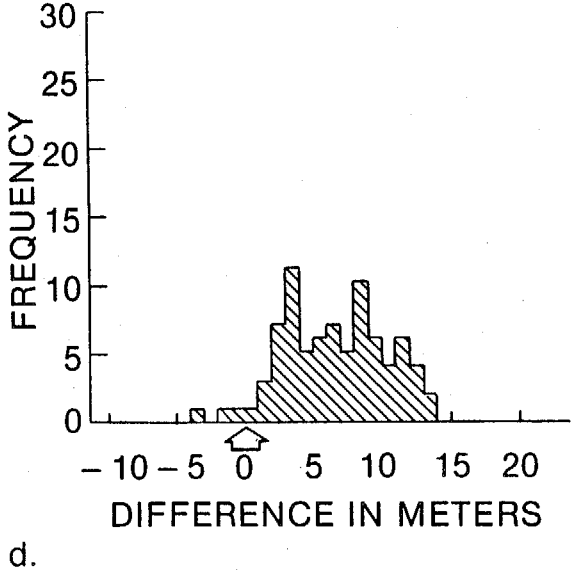
a.


30
25
20
15
10
5
0
FREQUENCY
– 10 – 5 0 5 10 15 20
DIFFERENCE IN METERS

c.


30
25
20
15
10
5
0
FREQUENCY
– 10 – 5 0 5 10 15 20
DIFFERENCE IN METERS

b.


30
25
20
15
10
5
0
FREQUENCY
– 10 – 5 0 5 10 15 20
DIFFERENCE IN METERS

d.

fig 4.

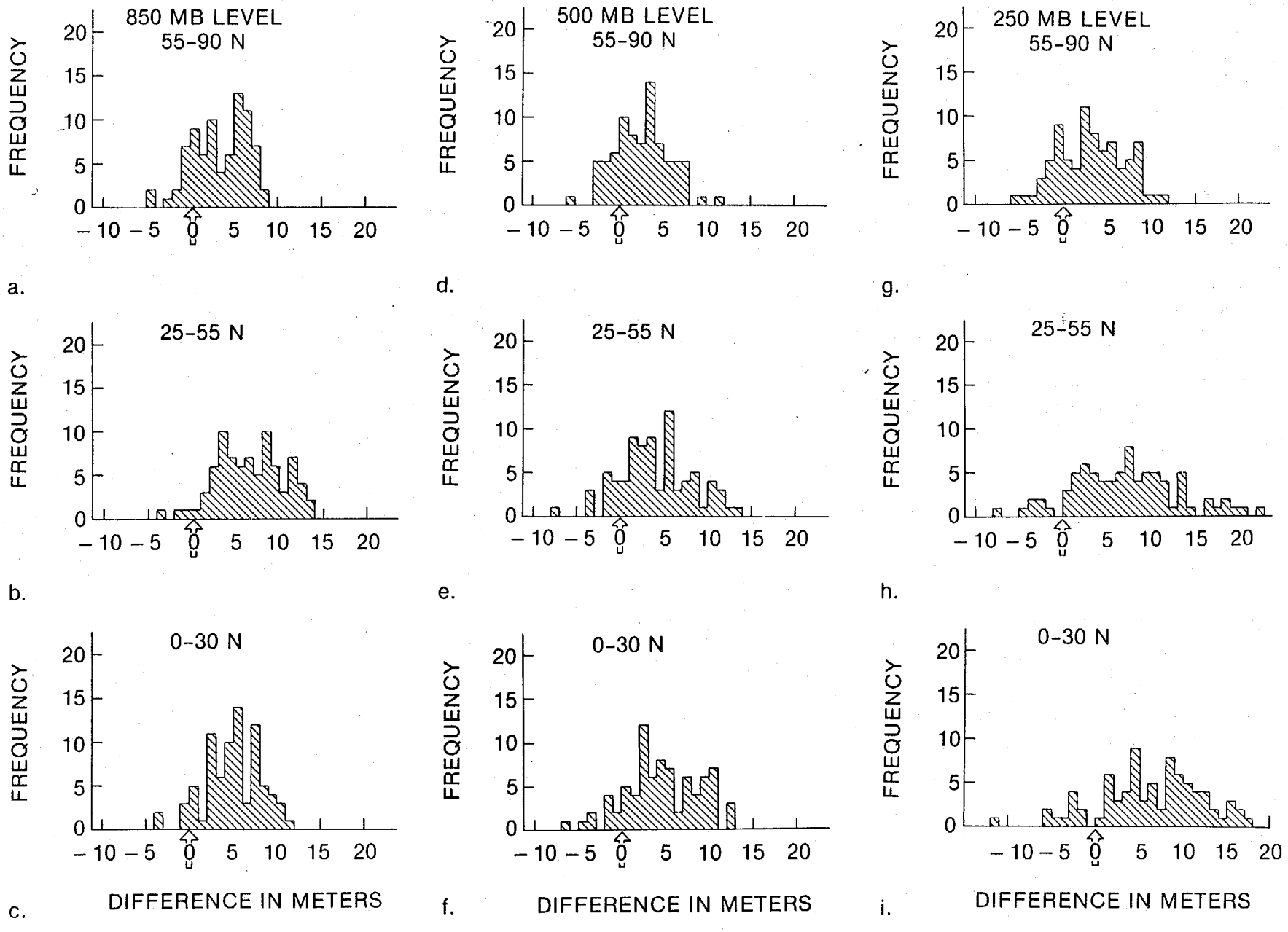
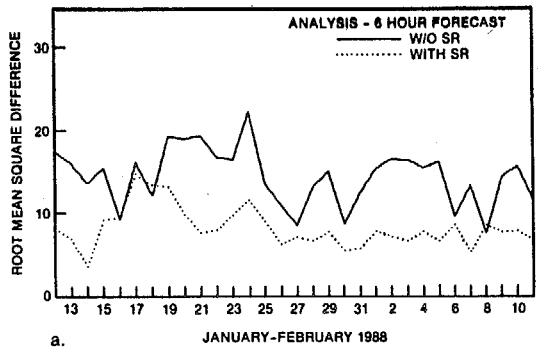
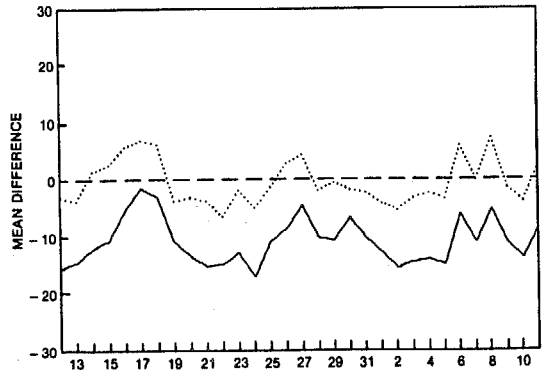

850 MB LEVEL
55-90 N
500 MB LEVEL
55-90 N
250 MB LEVEL
55-90 N
25-55 N
0-30 N
FREQUENCY
0 5 10 15 20
−10 −5 0 5 10 15 20
DIFFERENCE IN METERS
a.
b.
c.
d.
e.
f.
g.
h.
i.

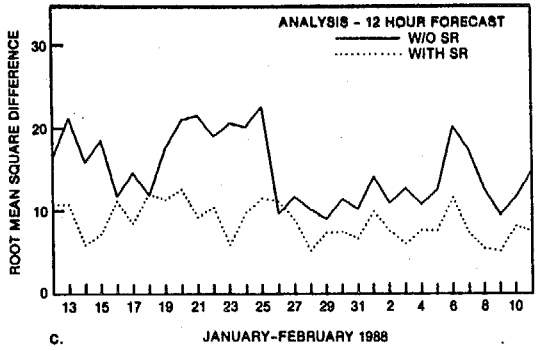
Fig. 5


ANALYSIS - 6 HOUR FORECAST
W/O SR
WITH SR
ROOT MEAN SQUARE DIFFERENCE
30
20
10
0
13 15 17 19 21 23 25 27 29 31 2 4 6 8 10
JANUARY-FEBRUARY 1988

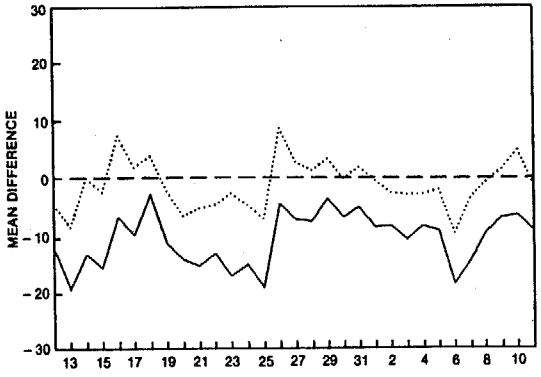
a.


MEAN DIFFERENCE
30
20
10
0
-10
-20
-30
13 15 17 19 21 23 25 27 29 31 2 4 6 8 10
JANUARY-FEBRUARY 1988

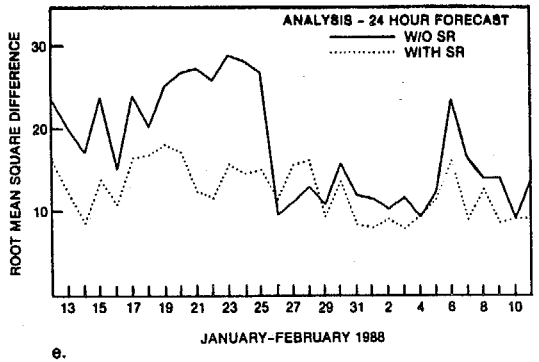
b.


ANALYSIS - 12 HOUR FORECAST
W/O SR
WITH SR
ROOT MEAN SQUARE DIFFERENCE
30
20
10
0
13 15 17 19 21 23 25 27 29 31 2 4 6 8 10
JANUARY-FEBRUARY 1988

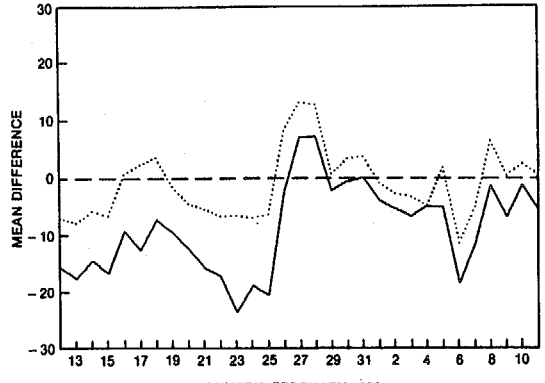
c.


MEAN DIFFERENCE
30
20
10
0
-10
-20
-30
13 15 17 19 21 23 25 27 29 31 2 4 6 8 10
JANUARY-FEBRUARY 1988

d.


ANALYSIS - 24 HOUR FORECAST
W/O SR
WITH SR
ROOT MEAN SQUARE DIFFERENCE
30
20
10
13 15 17 19 21 23 25 27 29 31 2 4 6 8 10
JANUARY-FEBRUARY 1988

e.


MEAN DIFFERENCE
30
20
10
0
-10
-20
-30
13 15 17 19 21 23 25 27 29 31 2 4 6 8 10
JANUARY-FEBRUARY 1988

f.

Fig. 6

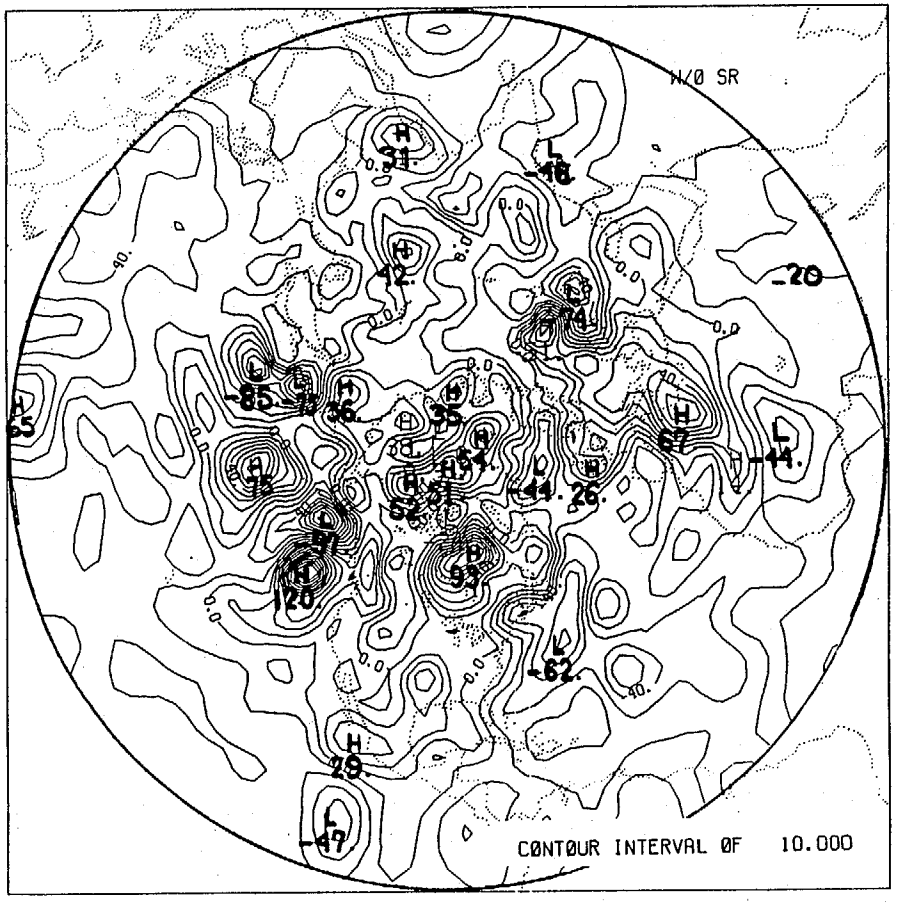

W/Ø SR
CØNTØUR INTERVAL ØF 10.000

fig. 7a.

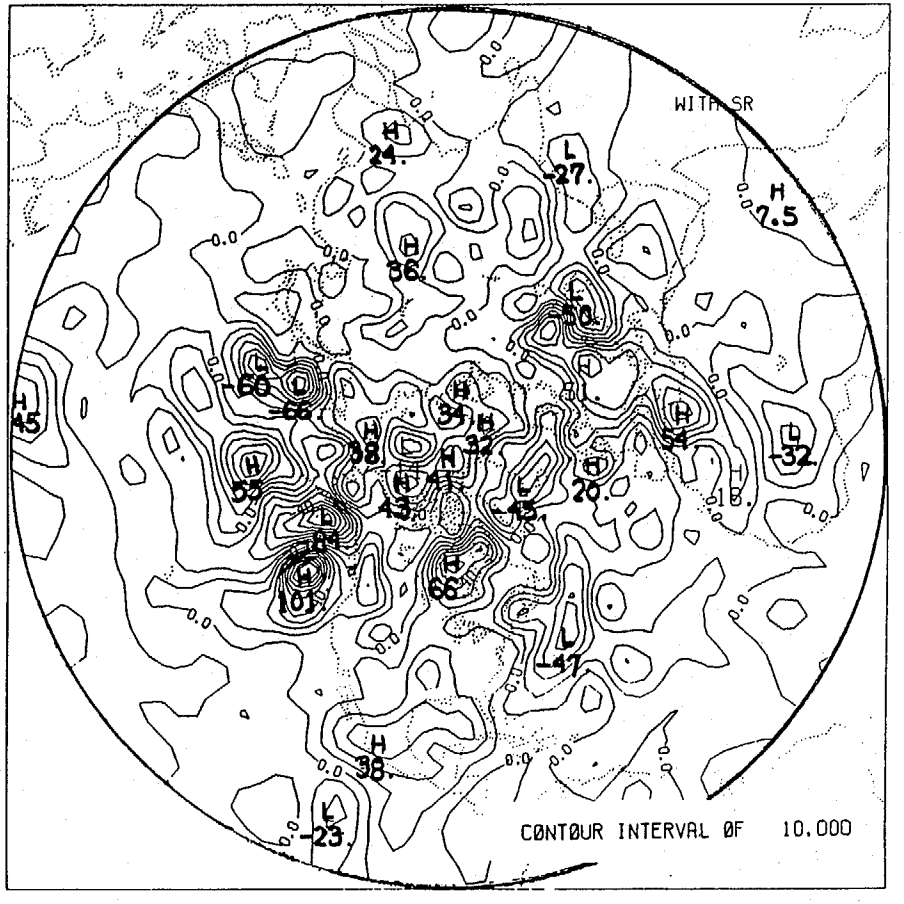

WITH SR
CØNTØUR INTERVAL ØF 10.000

fig. 7b.

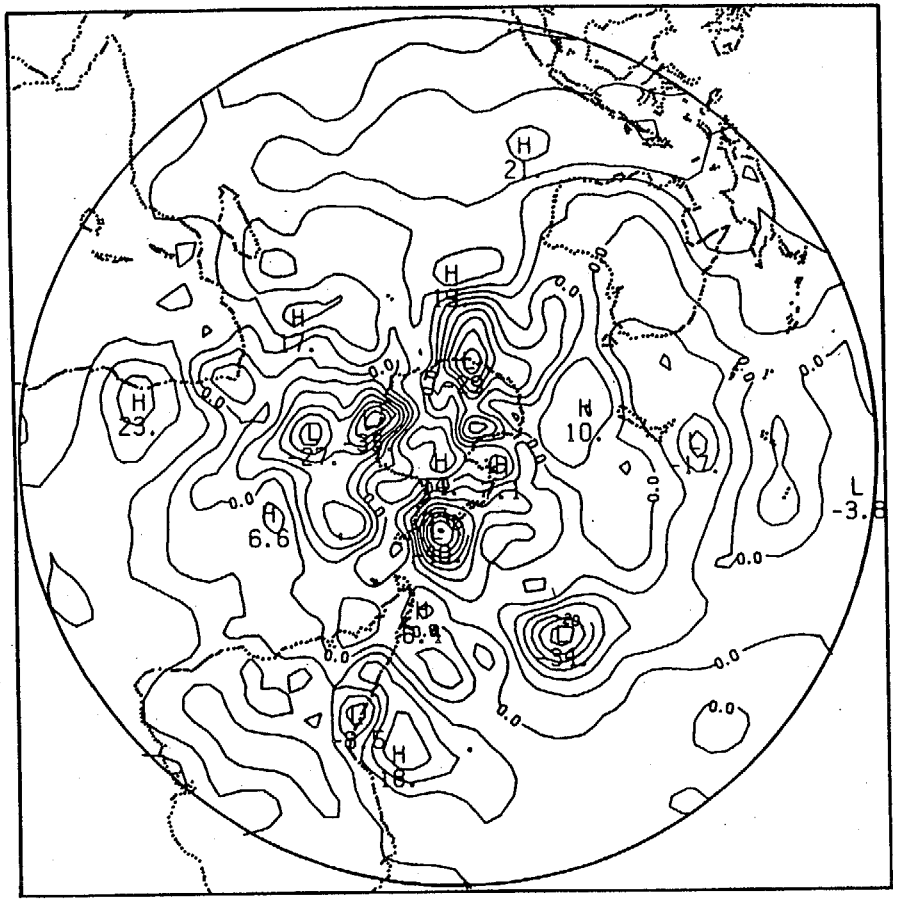


fig. 8a.

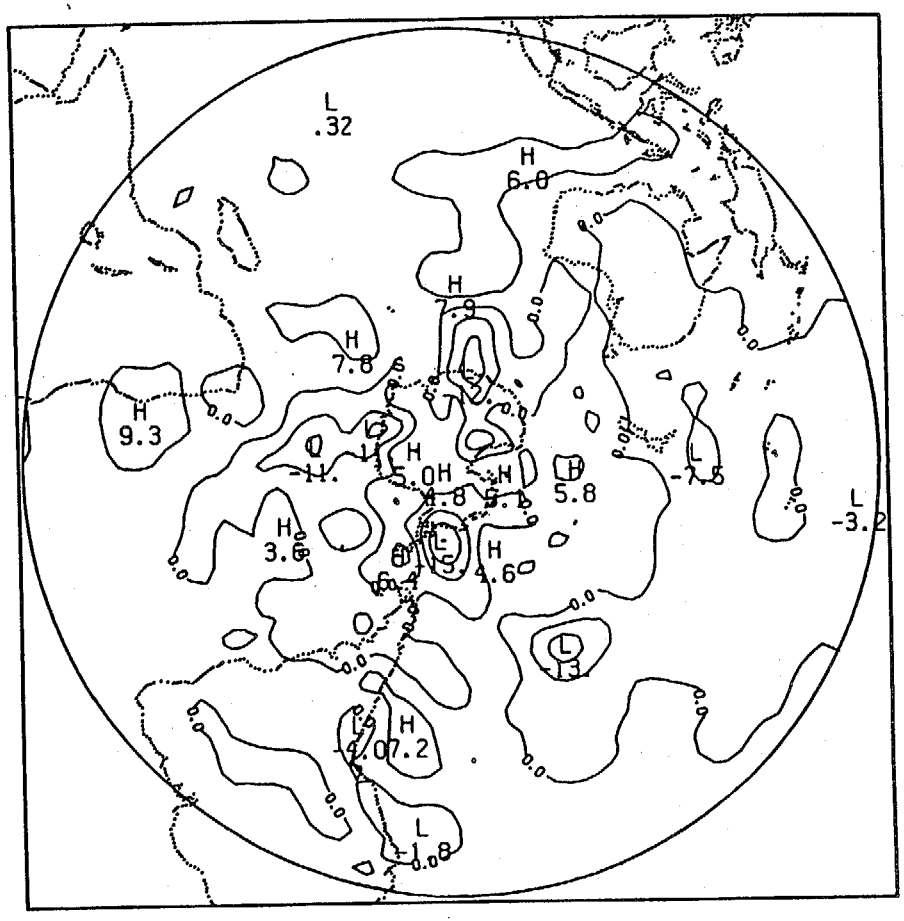


fig. 8b.

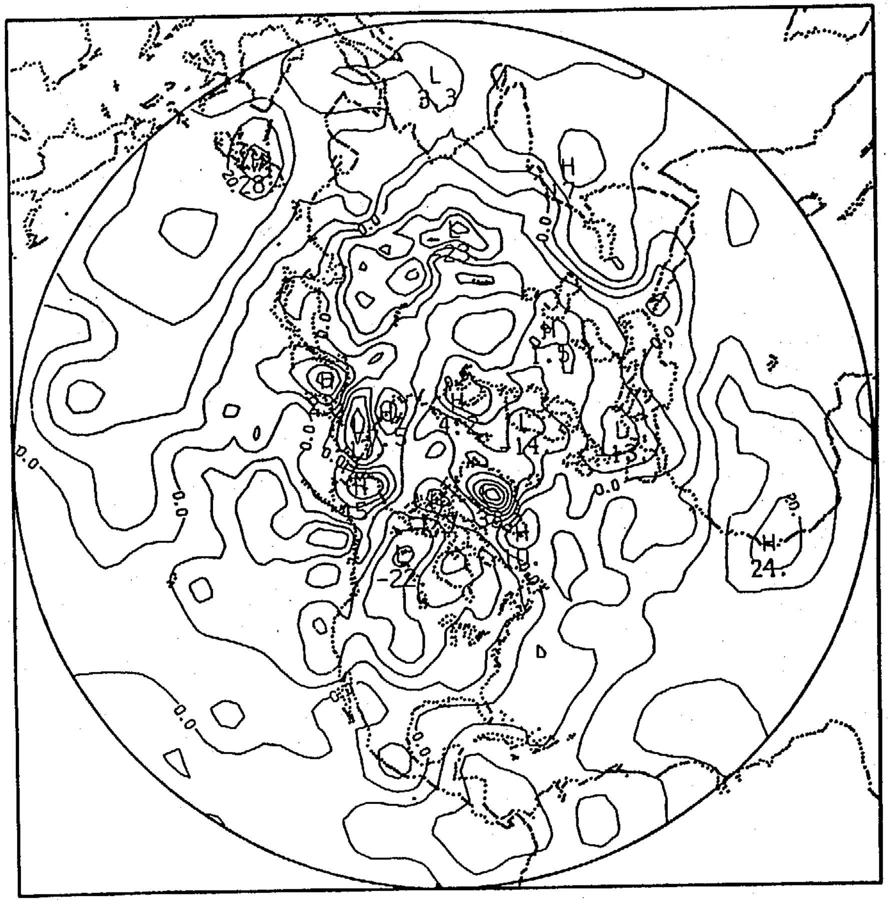


fig. 8c

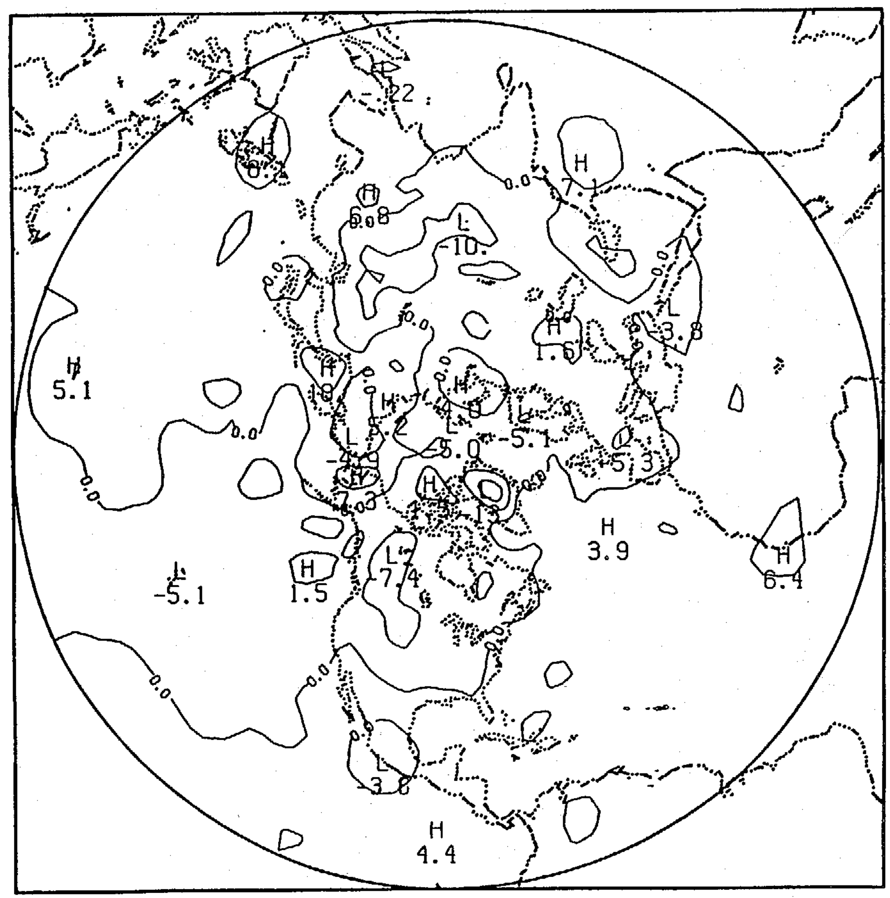


fig. 8d.

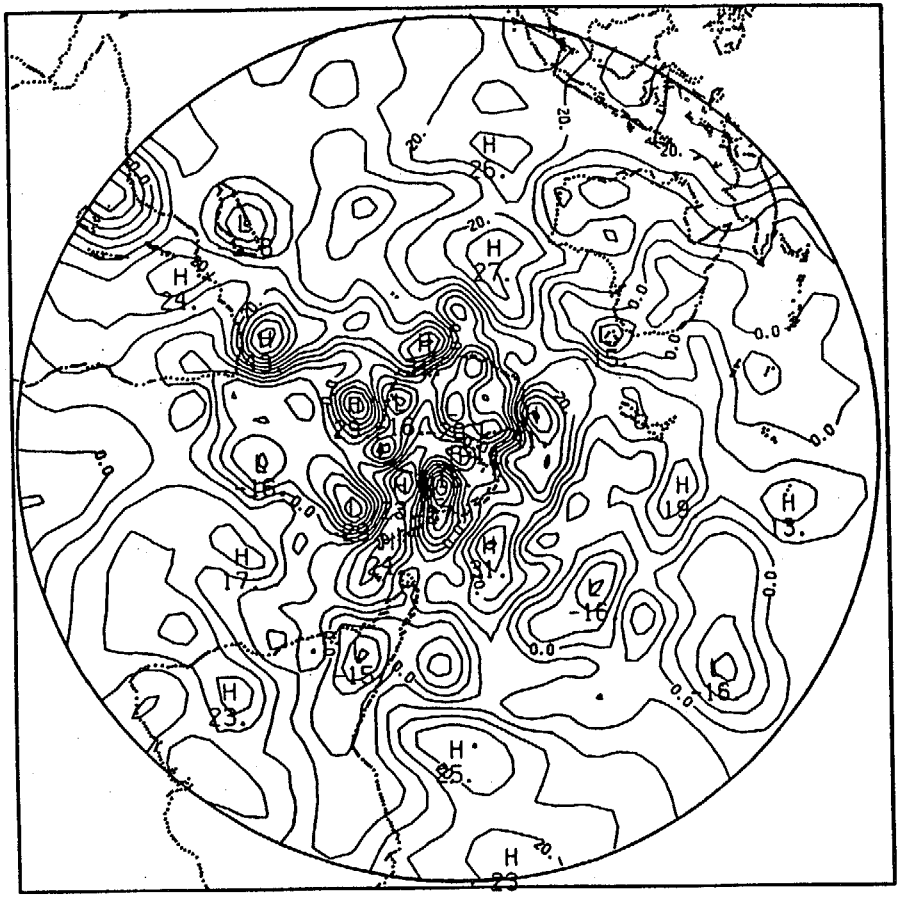


fig. 9a.

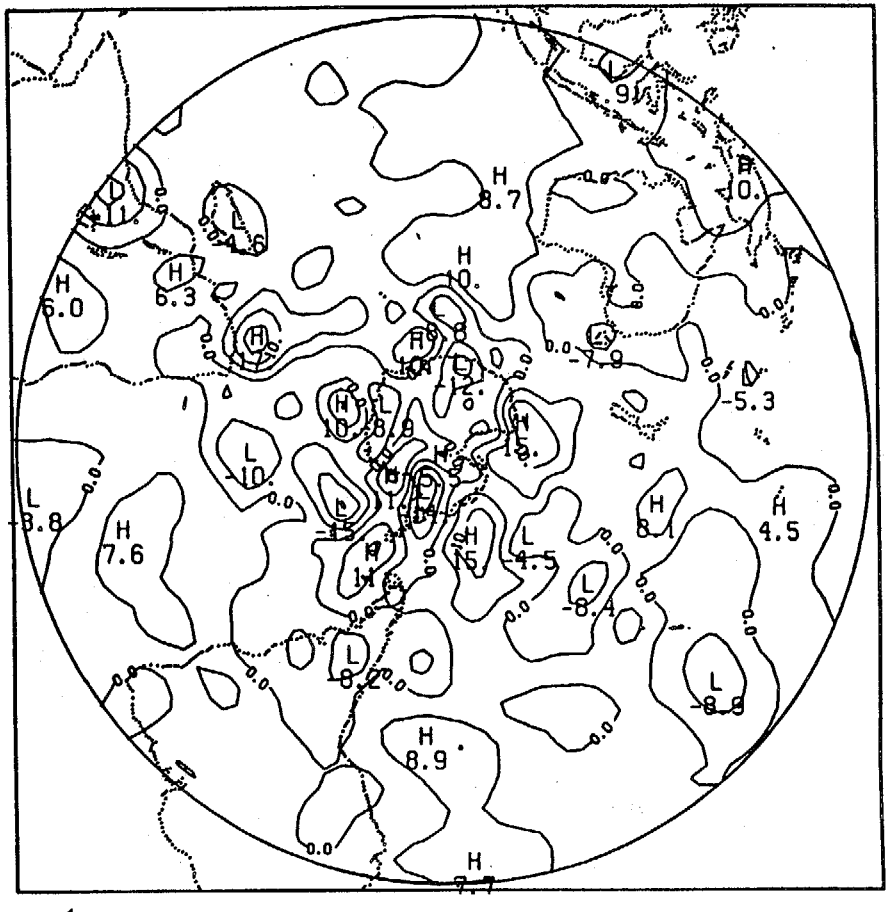


fig. 9b.

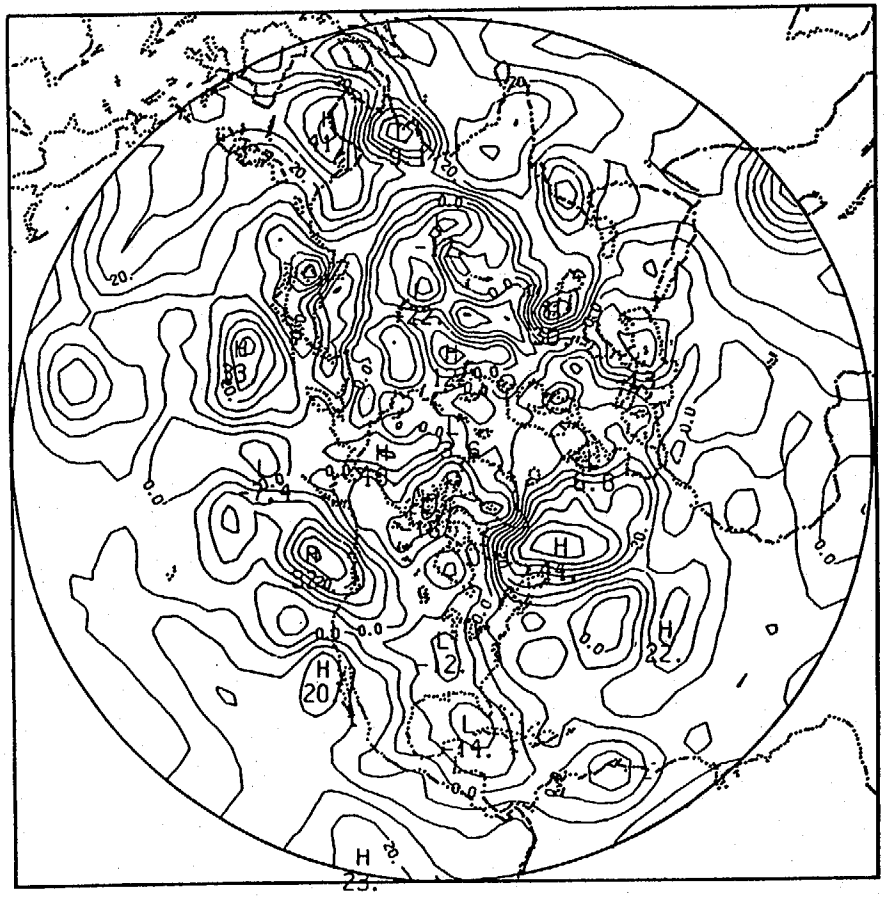


fig. 9c

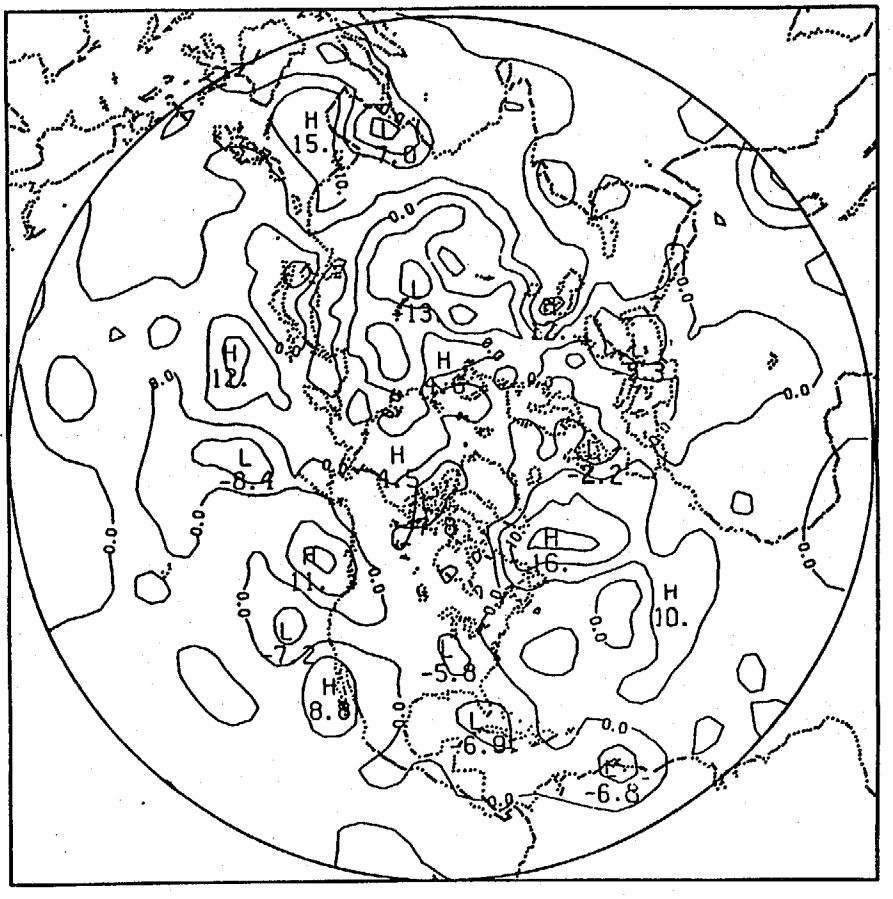
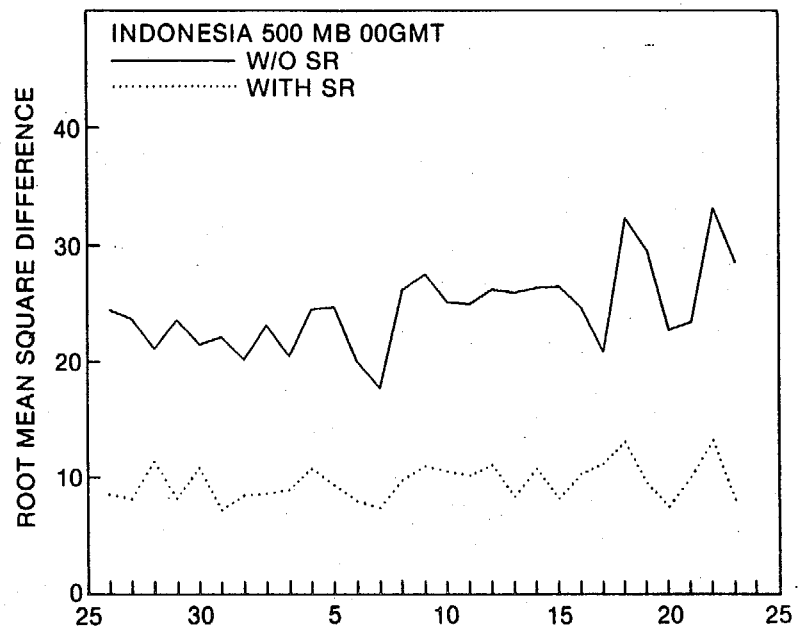
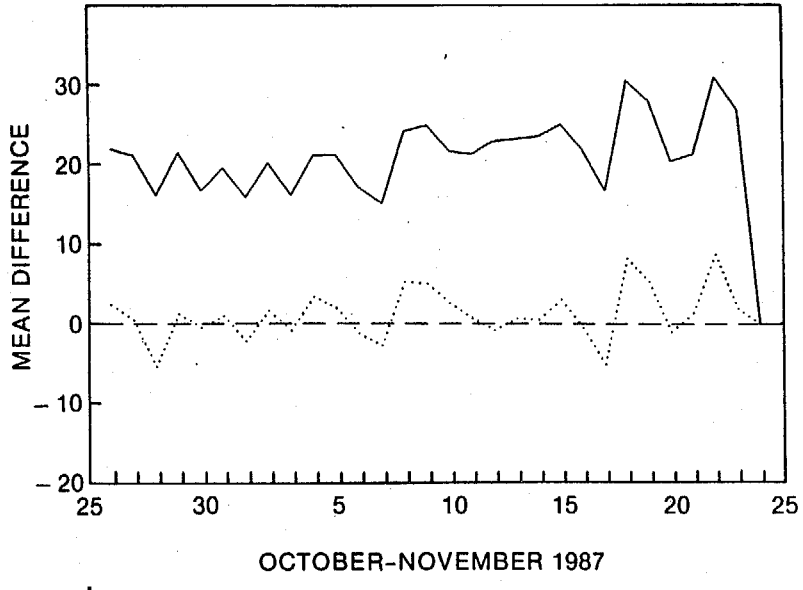


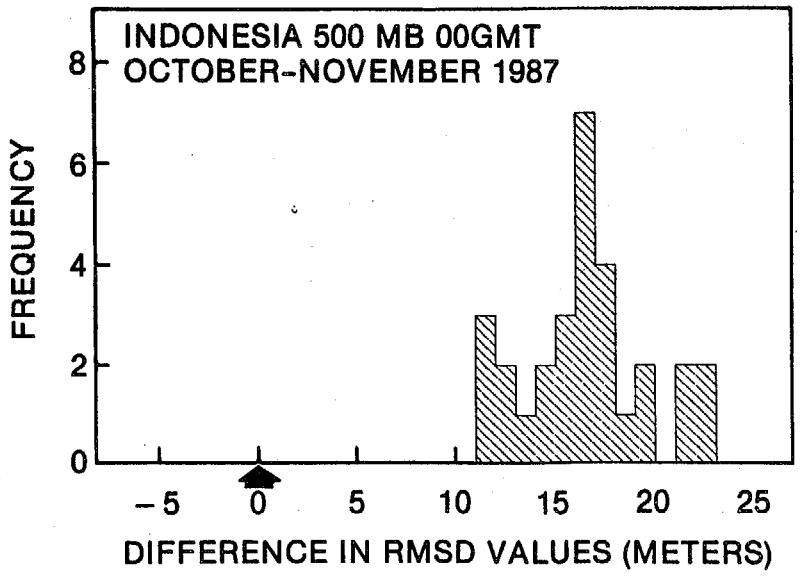
fig. 9d


INDONESIA 500 MB 00GMT
W/O SR
WITH SR
ROOT MEAN SQUARE DIFFERENCE
40
30
20
10
0
25
30
5
10
15
20
25

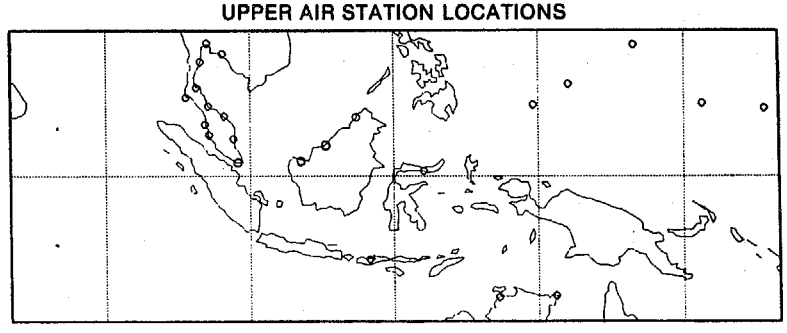
a.


MEAN DIFFERENCE
30
20
10
0
– 10
– 20
25
30
5
10
15
20
25
OCTOBER-NOVEMBER 1987

b.

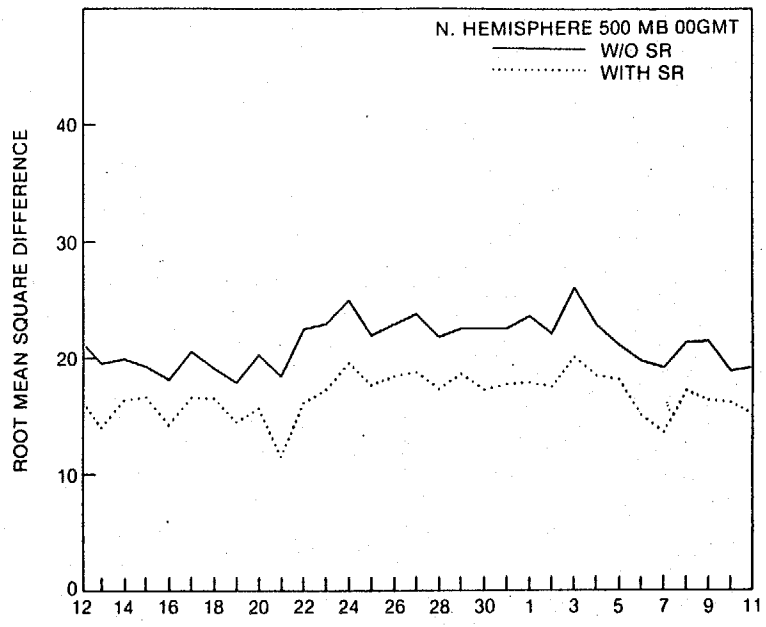

INDONESIA 500 MB 00GMT
OCTOBER-NOVEMBER 1987
FREQUENCY
8
6
4
2
0
– 5
0
5
10
15
20
25
DIFFERENCE IN RMSD VALUES (METERS)

c.

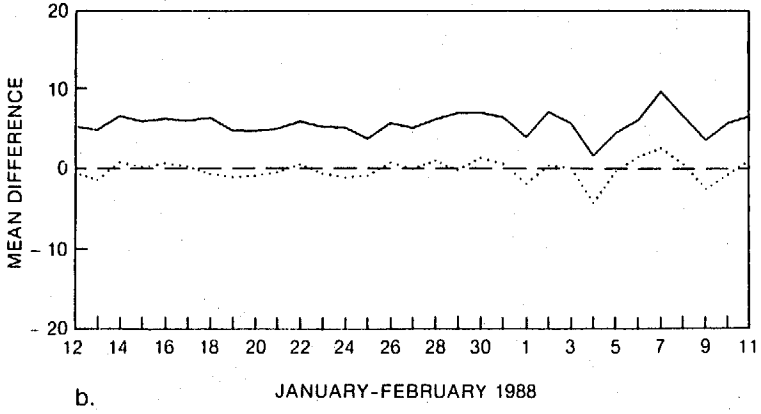

UPPER AIR STATION LOCATIONS

d.

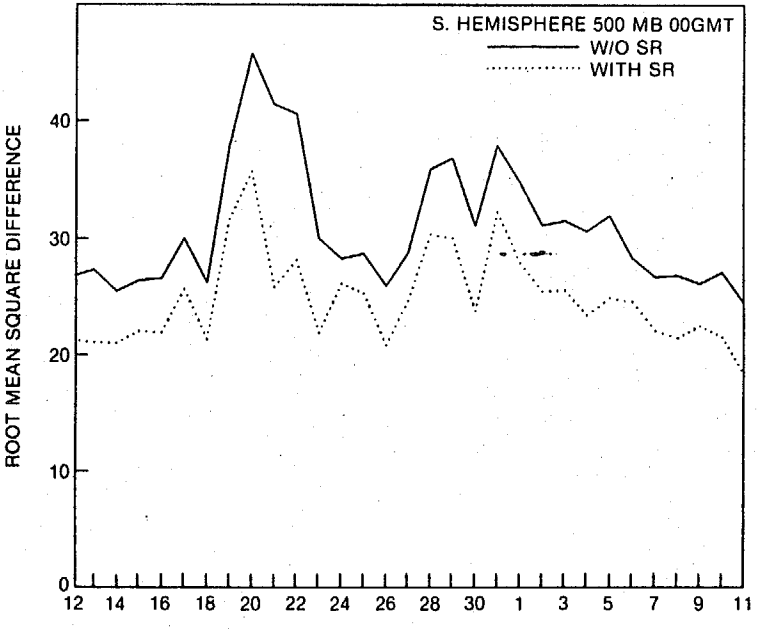
fig. 10


N. HEMISPHERE 500 MB 00GMT
W/O SR
WITH SR
ROOT MEAN SQUARE DIFFERENCE
0
10
20
30
40
12 14 16 18 20 22 24 26 28 30 1 3 5 7 9 11

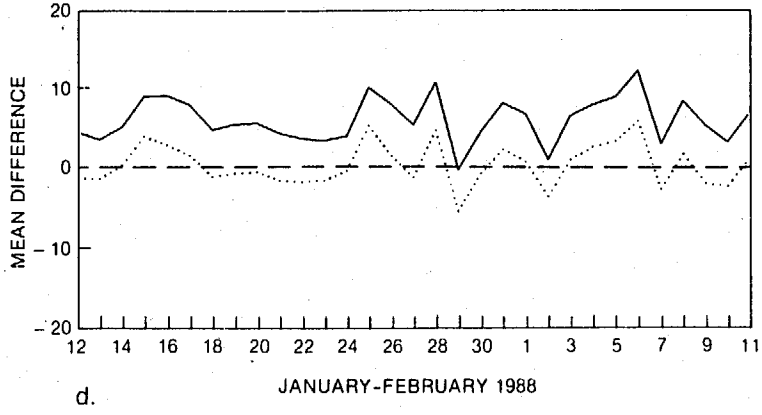
a.


MEAN DIFFERENCE
20
10
0
-10
-20
12 14 16 18 20 22 24 26 28 30 1 3 5 7 9 11
JANUARY-FEBRUARY 1988

b.


S. HEMISPHERE 500 MB 00GMT
W/O SR
WITH SR
ROOT MEAN SQUARE DIFFERENCE
0
10
20
30
40
12 14 16 18 20 22 24 26 28 30 1 3 5 7 9 11

c.


MEAN DIFFERENCE
20
10
0
-10
-20
12 14 16 18 20 22 24 26 28 30 1 3 5 7 9 11
JANUARY-FEBRUARY 1988

d.

fig. 11

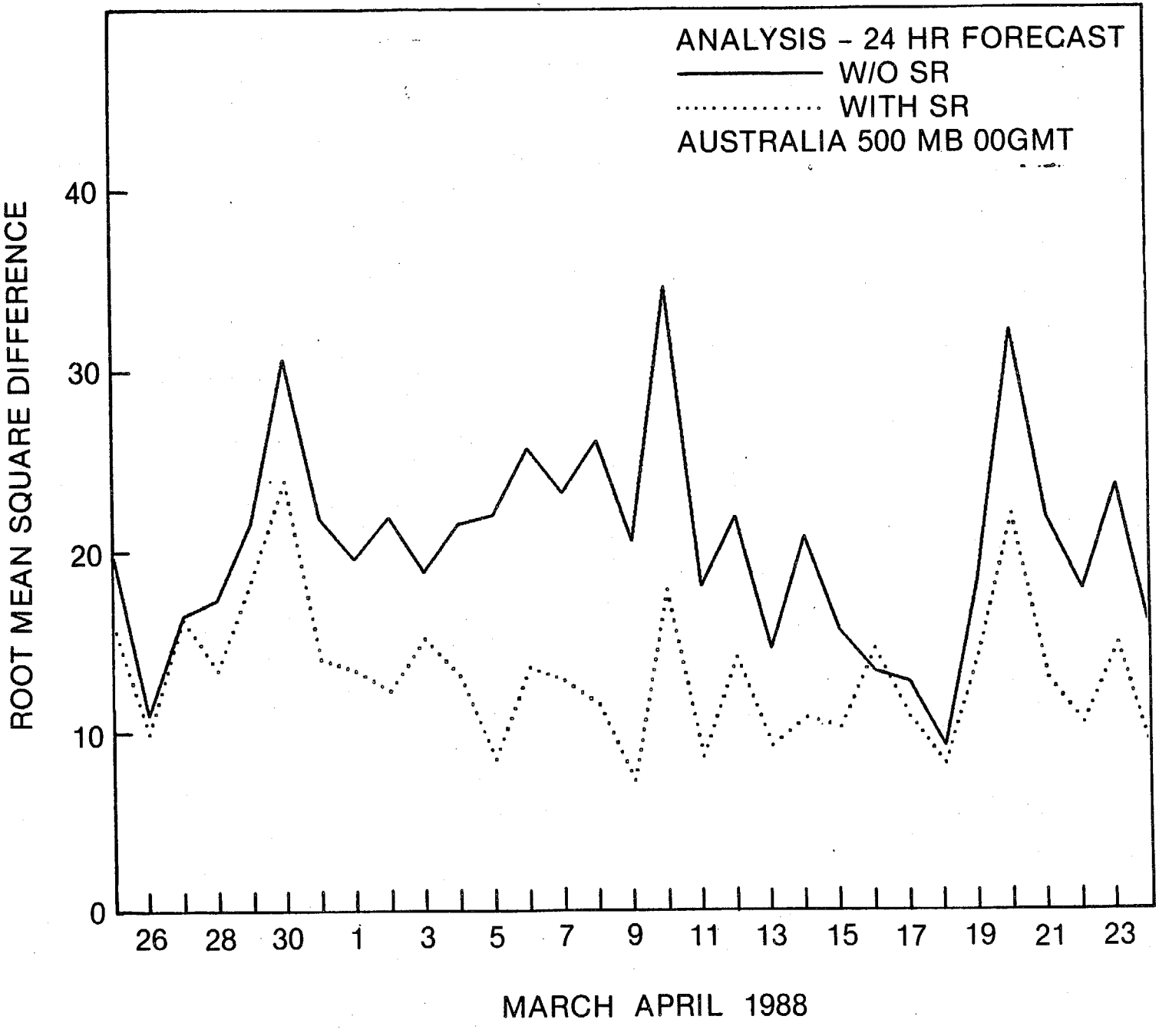

ANALYSIS - 24 HR FORECAST
W/O SR
WITH SR
AUSTRALIA 500 MB 00GMT
ROOT MEAN SQUARE DIFFERENCE
0
10
20
30
40
26
28
30
1
3
5
7
9
11
13
15
17
19
21
23
MARCH APRIL 1988

fig. 12

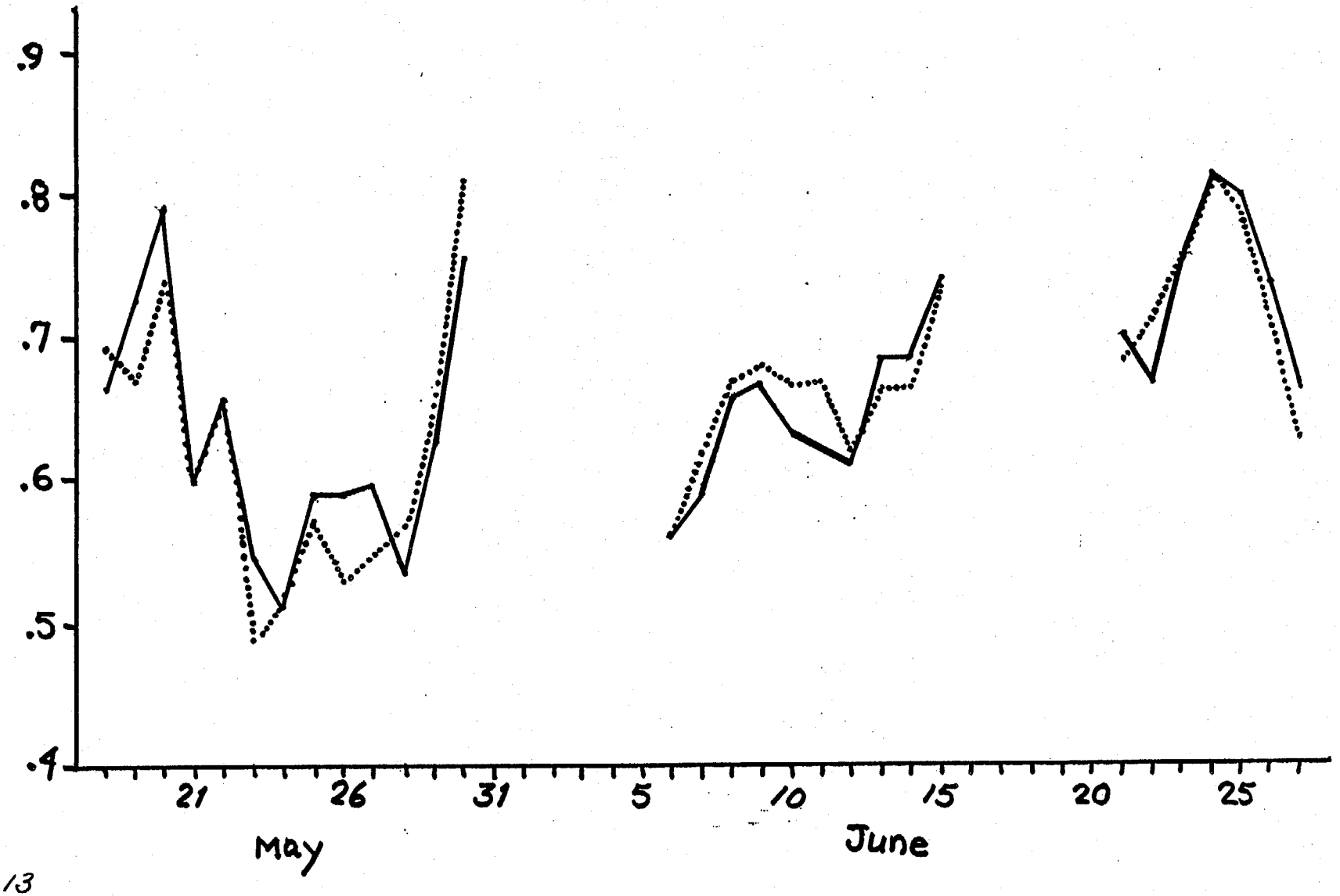


Fig. 13

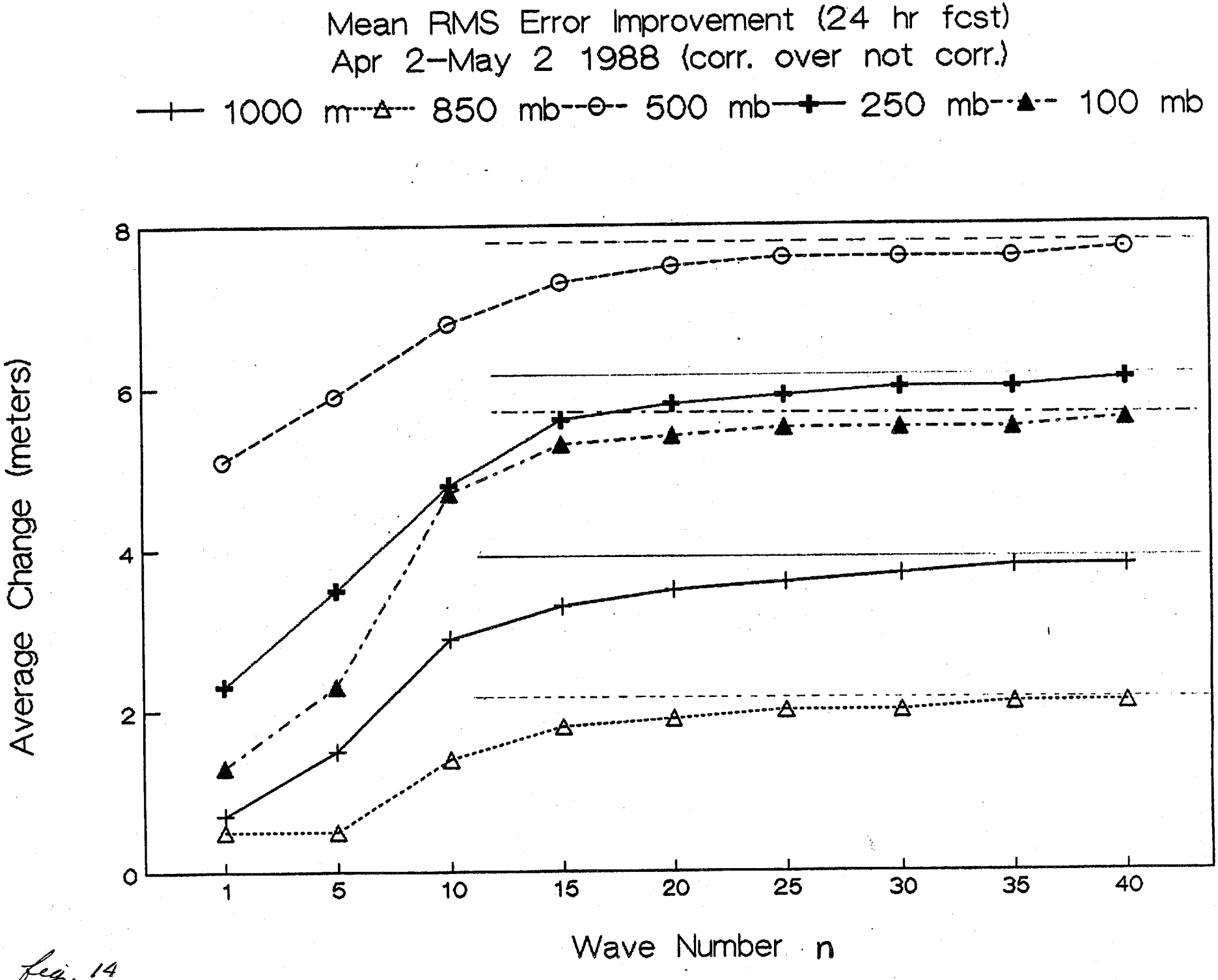

Mean RMS Error Improvement (24 hr fcst)
Apr 2–May 2 1988 (corr. over not corr.)
1000 mb
850 mb
500 mb
250 mb
100 mb
Average Change (meters)
0
2
4
6
8
1
5
10
15
20
25
30
35
40
Wave Number n

fig. 14

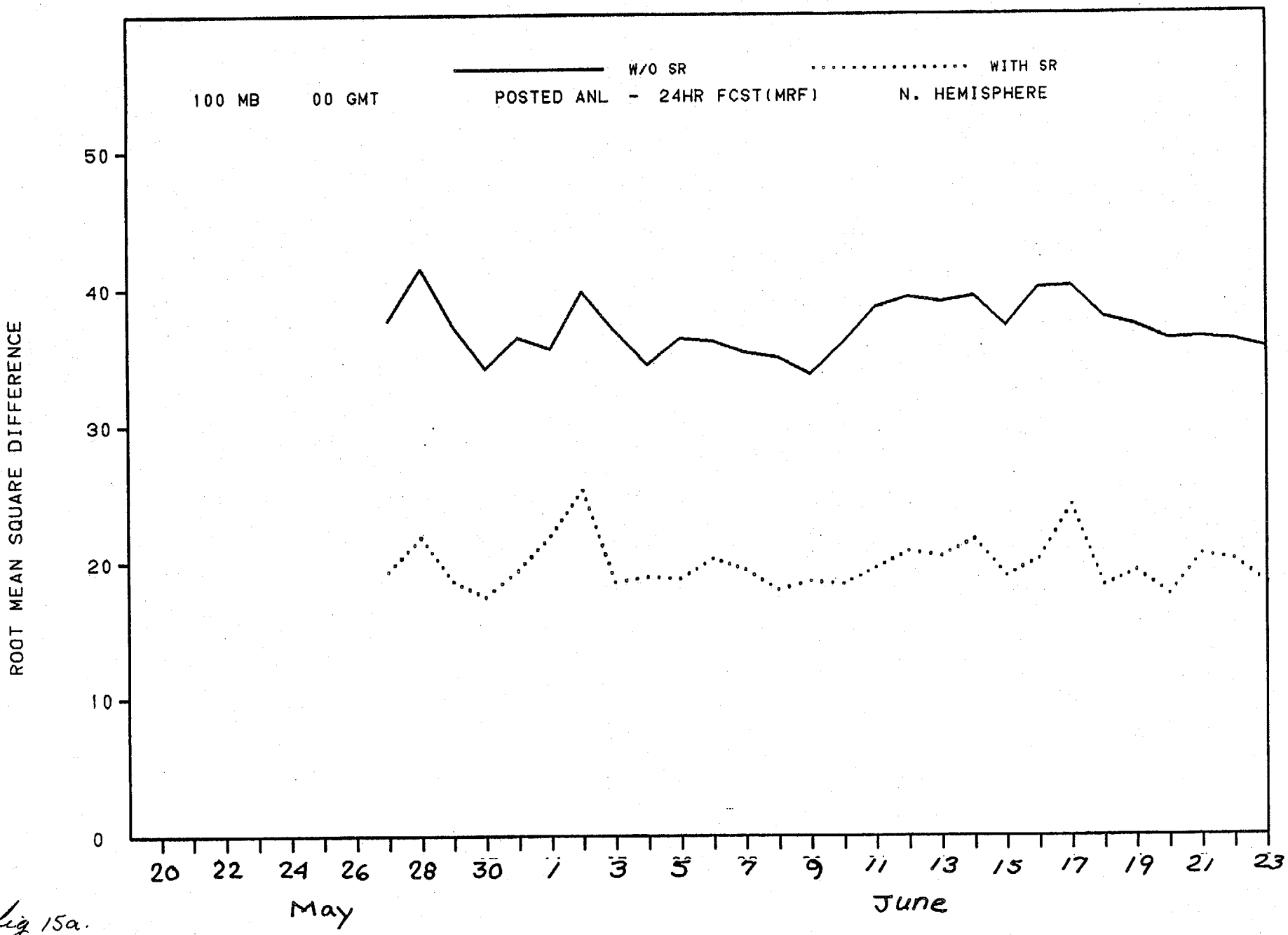

W/O SR
WITH SR
100 MB 00 GMT
POSTED ANL - 24HR FCST(MRF)
N. HEMISPHERE
ROOT MEAN SQUARE DIFFERENCE
50
40
30
20
10
0
20 22 24 26 28 30 1 3 5 7 9 11 13 15 17 19 21 23
May
June

fig 15a.

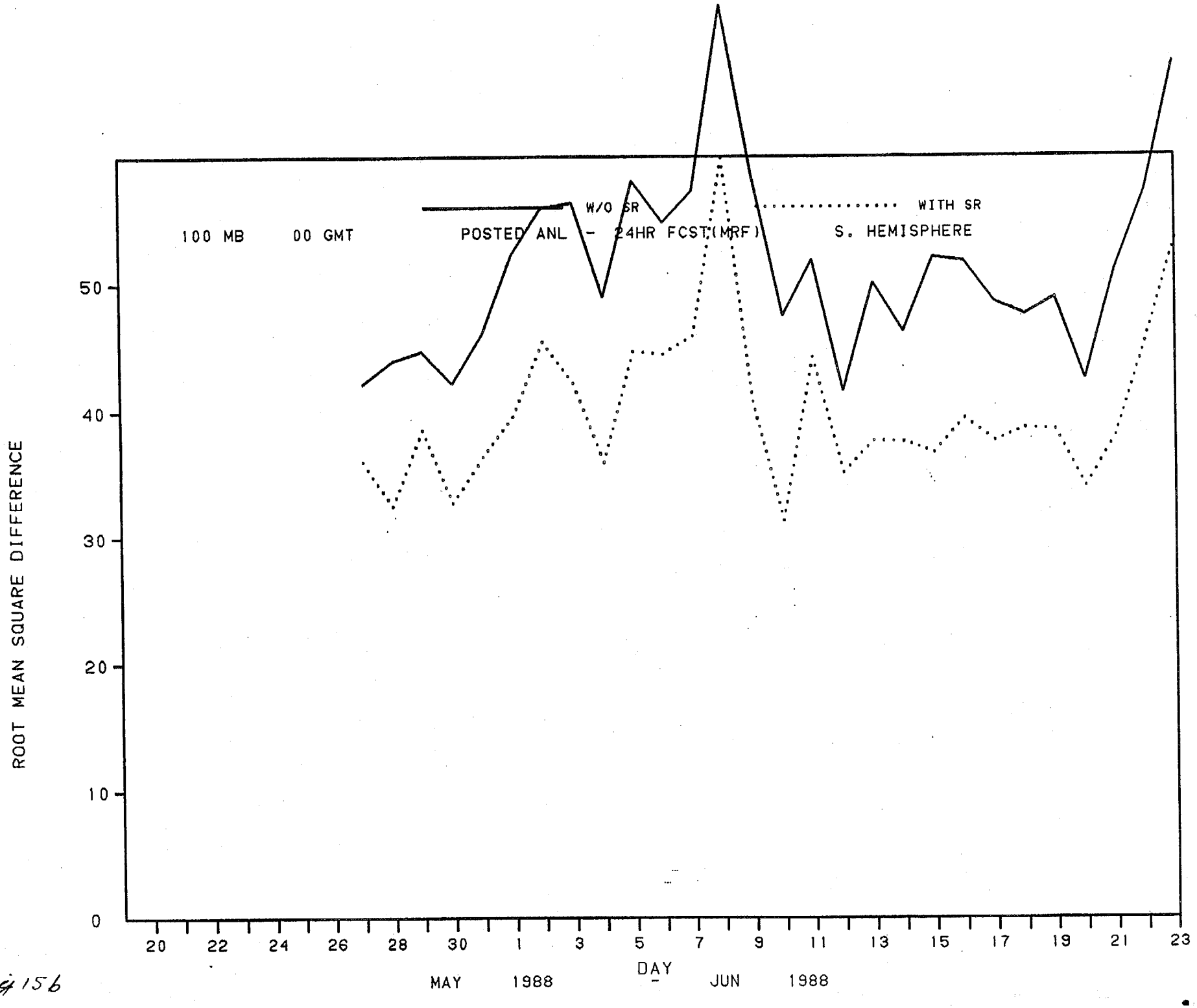

W/O SR
WITH SR
100 MB 00 GMT
POSTED ANL - 24HR FCST(MRF)
S. HEMISPHERE
ROOT MEAN SQUARE DIFFERENCE
0
10
20
30
40
50
20 22 24 26 28 30 1 3 5 7 9 11 13 15 17 19 21 23
MAY 1988
DAY
JUN 1988

fig 156

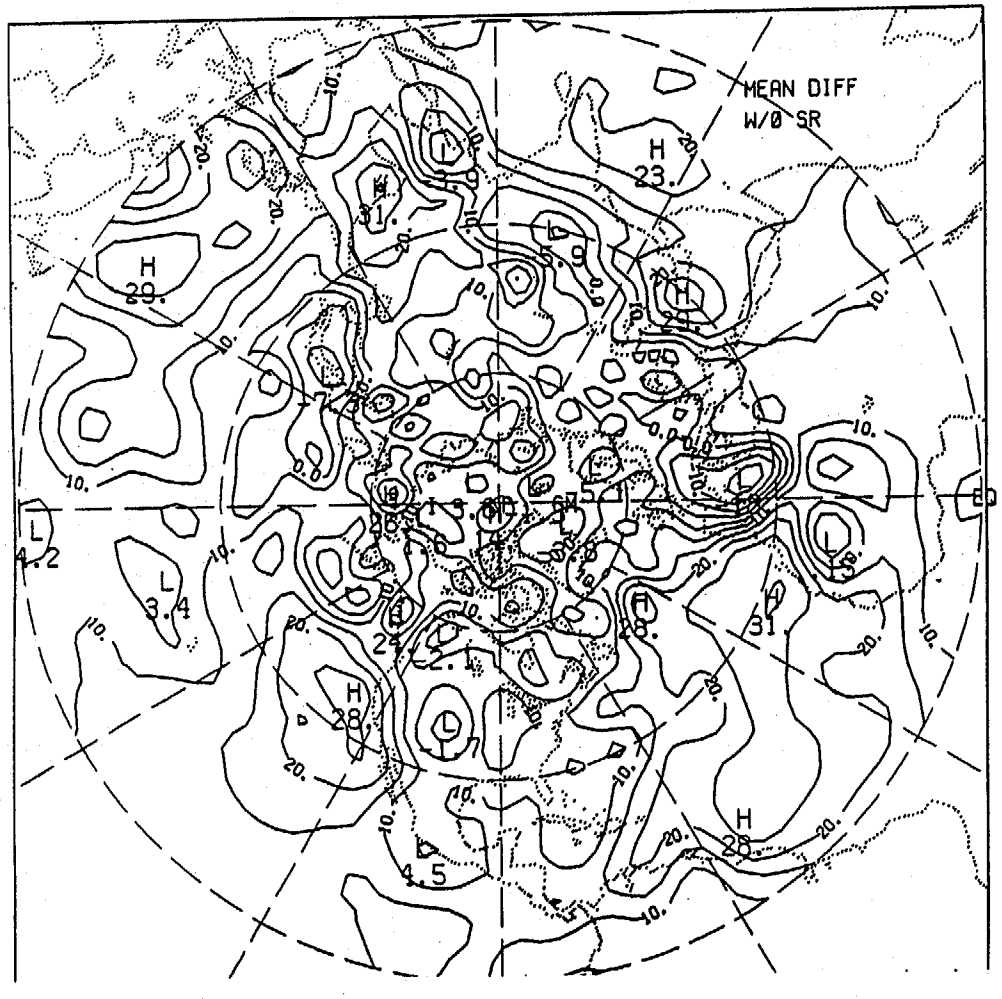


fig. 16a

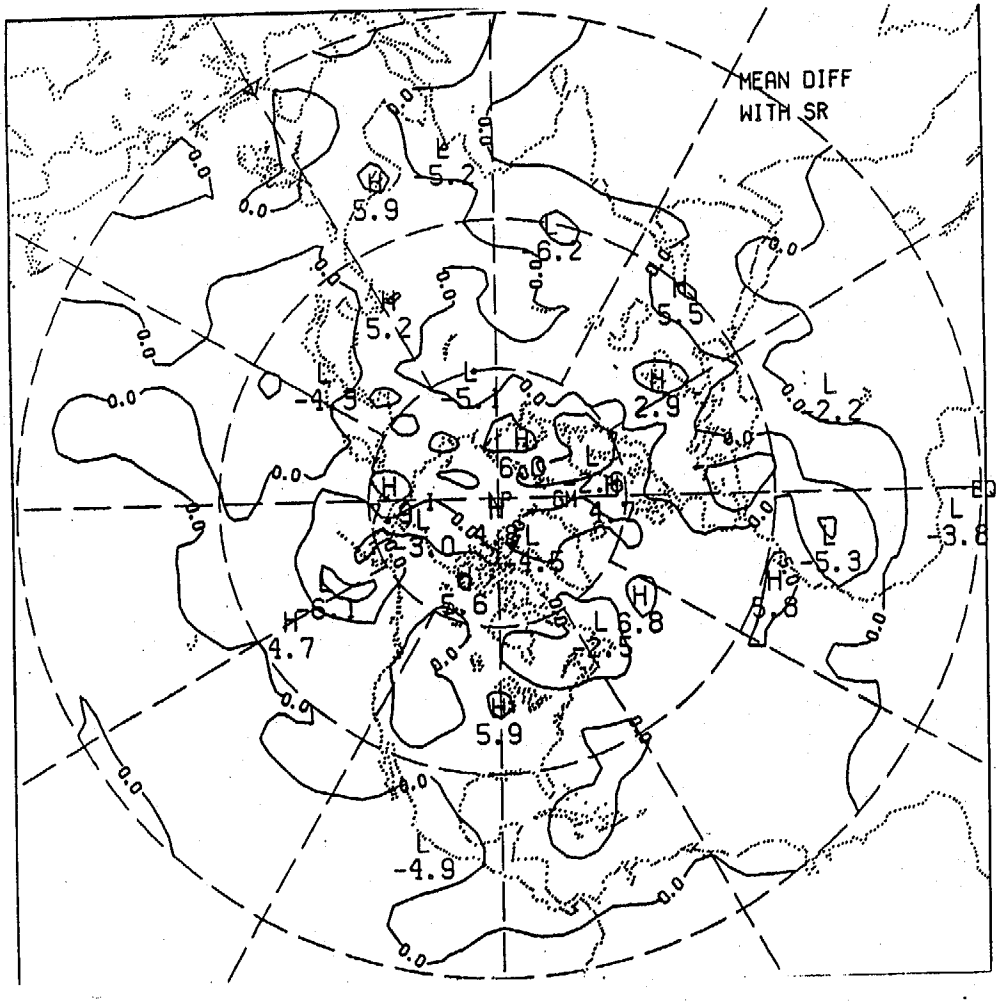


fig. 16b

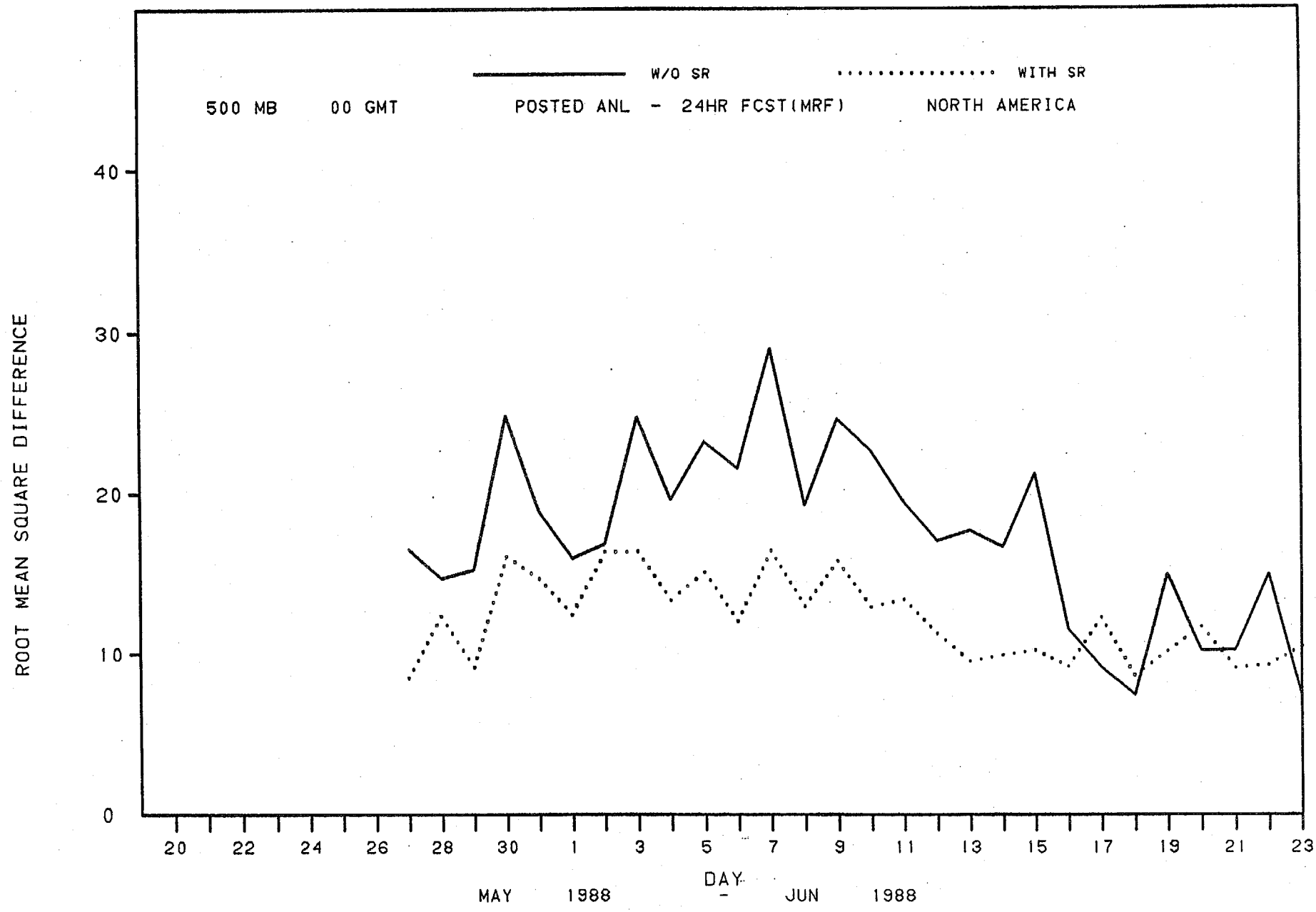

W/O SR
WITH SR
500 MB 00 GMT
POSTED ANL - 24HR FCST(MRF)
NORTH AMERICA
ROOT MEAN SQUARE DIFFERENCE
0
10
20
30
40
20 22 24 26 28 30 1 3 5 7 9 11 13 15 17 19 21 23
DAY
MAY 1988 - JUN 1988

fig. 17a

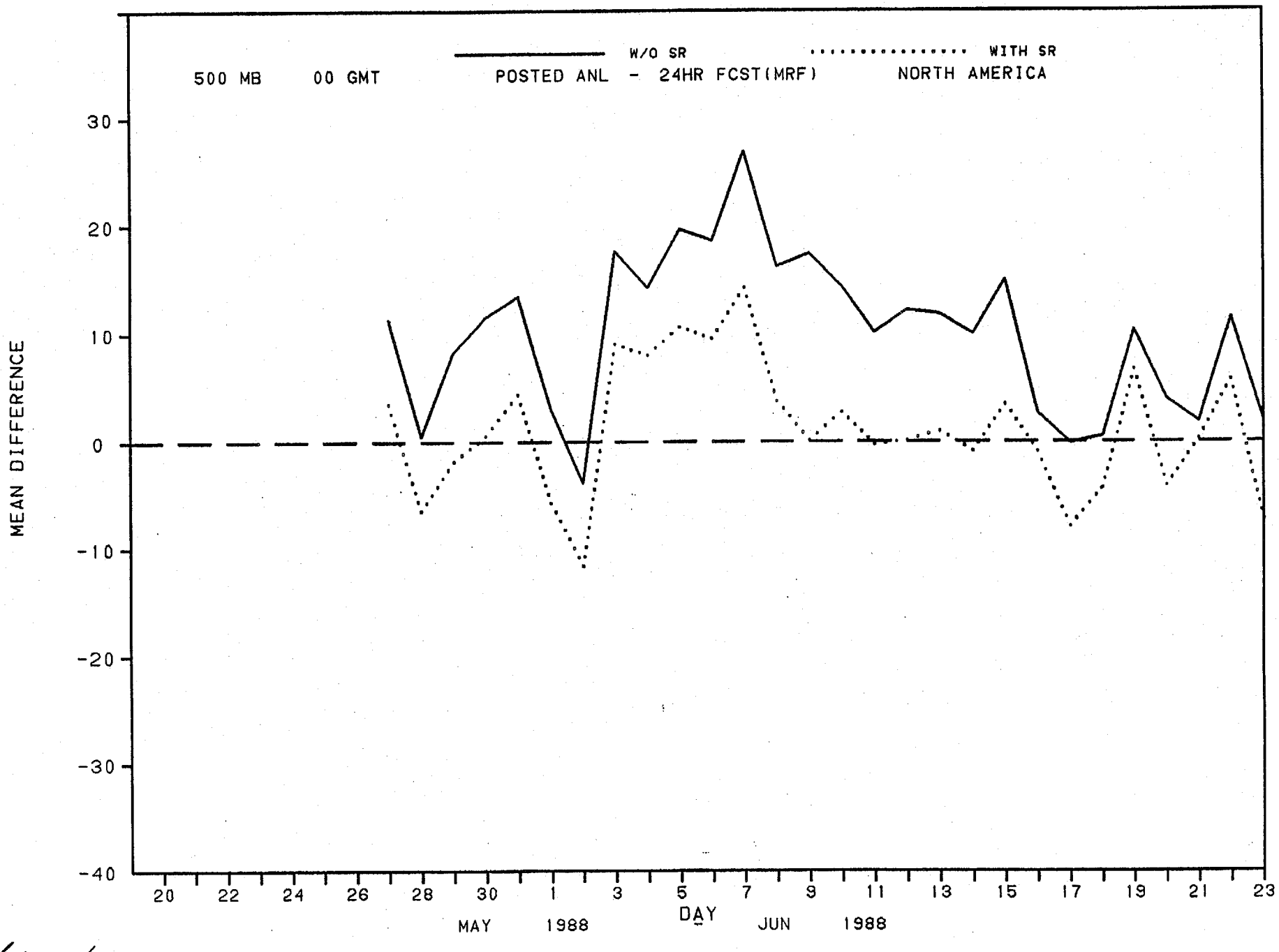

W/O SR
WITH SR
500 MB 00 GMT
POSTED ANL - 24HR FCST(MRF)
NORTH AMERICA
MEAN DIFFERENCE
30
20
10
0
-10
-20
-30
-40
20 22 24 26 28 30 1 3 5 7 9 11 13 15 17 19 21 23
MAY 1988
DAY
JUN 1988

Fig. 176

Stimulating America's Progress
1913-1988