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

OFFICE NOTE 119

Analysis Error as a Function of Observation Density for Satellite Temperature Soundings with Spatially Correlated Errors

Kenneth H. Bergman
and
William D. Bonner
Development Division

SEPTEMBER 1975

1. Introduction

An important question in planning the observing system for the GARP Global Experiment is the horizontal resolution required for satellite-derived temperatures. Accuracy and resolution requirements for temperature measurements have been clearly established as ± 1 deg C and 500 km (GARP Publication 11). It seems unlikely at this time that the satellite derived temperatures will be this accurate, and suggestions have been made that the failure to achieve 1 deg C accuracy can be compensated for by increasing the yield of satellite-derived temperatures. The purpose of this note is to point out that our ability to do this may be severely limited by spatial correlations in the errors of the temperature retrievals.

In order to examine the effect of spatial error correlations on the usefulness of observational data, we have carried out a simple numerical experiment, similar in design to one by Alaka and Elvander (1972a, 1972b). The significant departures from their experiment are that calculations are for temperatures rather than winds, that the analysis is considered to be in terms of deviations from a forecast rather than from some climatological state and, most important, that the errors in observations are not assumed to be randomly distributed in space. Systematic errors in the satellite temperature soundings may be expected to arise from the use of a numerical forecast as a first guess for the temperature retrieval (McMillin et al, 1973) and/or from the effects of large-scale cloud patterns on the radiance measurements.

2. The Experiment

Figure 1 shows the basic design of the experiment. Twelve observations are arrayed in a rectangular grid of horizontal spacing h. Suppose an interpolated value of the meteorological variable is desired at the central point X. (This point might be a grid point for a numerical prediction model, for instance.) The statistically most reliable way of doing this is by optimum interpolation, an application of linear regression theory to spatial interpolation of data.

Suppose temperature is the variable being analyzed. Following Gandin (1963), let

$$\hat{T} = \tilde{T} + \hat{t} \ , \tag{1}$$

where $\hat{T}$ is the observed value of the temperature at a given location, $\tilde{T}$ is some "guess" value--provided by a forecast or climatology--of the temperature at the location, and $\hat{t}$ is the deviation of the observed temperature from the guess value. This may be re-expressed as

$$\hat{T} = \tilde{T} + (t + \varepsilon) \ , \tag{2}$$

where t is the "true" temperature deviation at the location and ε is the observational error.

Then the "analyzed" value $\hat{T}_a$ of the temperature field at an arbitrary location is given by a weighted linear combination of the deviational temperatures of the neighboring observations:

$$\hat{T}_a = \tilde{T}_a + \sum_{i=1}^{n} c_i \hat{t}_i = \tilde{T}_a + \sum_{i=1}^{n} c_i (t_i + \varepsilon_i) \ , \tag{3}$$

where $\tilde{T}_a$ is the guess value at the analysis location, the c_i are the weights to be assigned to the deviational temperature observations $\hat{t}_i$, and n is the number

of observations used in the analysis. In the example of Figure 1, n is 12 if all of the observations are used in the analysis for the central point.

In general, the right-hand side of (3) will not give the true value T_a of the temperature at the analysis point, $\hat{T}_a$ will differ from T_a by an amount called the analysis error E. It is generally not possible to eliminate this source of error exactly, no matter how the weights c_i are chosen, but we can require that the statistical mean-square error of equation (3),

$$\overline{E^2} = \overline{(T_a - \tilde{T}_a)^2} = \overline{[T_a - \tilde{T}_a - \sum_{i=1}^{n} c_i(t_i + \varepsilon_i)]^2}, \qquad (4)$$

be a minimum for a large ensemble of interpolations and use this requirement to determine the weights. The result is a set of linear equations,

$$\sum_{j=1}^{n} \overline{(t_i + \varepsilon_i)(t_j + \varepsilon_j)} c_j = \overline{t_a(t_i + \varepsilon_i)}, \quad i=1,2,\ldots n \qquad (5)$$

where $t_a \equiv T_a - \tilde{T}_a$,

which may be solved for the weights c_i provided the statistical covariances of (5) are known or approximated.

For radiosonde temperatures, whose errors may safely be assumed to be random and mutually independent, the set of equations (5) simplifies to

$$\sum_{j=1}^{n} \overline{t_i t_j} c_j + \varepsilon_i^2 c_i = \overline{t_a t_i}, \quad i=1,2,\ldots n \qquad (6)$$

Here ε_i^2 is simply the error variance of the i^{th} observation. This is essentially the expression used for determining the weights in Alaka and Elvander's numerical experiment.

For satellite-derived temperatures with spatially correlated errors, $\overline{\varepsilon_i \varepsilon_j}$ does not vanish when $i \neq j$; therefore the set of equations (5) becomes in this case

$$\sum_{j=1}^{n} (\overline{t_i t_j} + \overline{\varepsilon_i \varepsilon_j}) c_j = \overline{t_a t_i} , \quad i=1,2,\ldots n \tag{7}$$

Equation (7) is strictly correct only if the observational errors of the measurement system are not correlated with the true values of temperature. In an unpublished study, we have shown that such correlations do exist for satellite-derived temperatures--as they will for any conservative observing system which tends to underestimate extremes; however the variances ($\overline{t_i \varepsilon_j}$, etc.) are probably small compared to $\overline{\varepsilon_i \varepsilon_j}$.

It is convenient for computational purposes to normalize equations (7) by dividing by the deviational variance $\overline{t_a}^2$. Equations (7) then take the form

$$\sum_{j=1}^{n} (\mu_{ij} + \rho_{ij}\, \sigma_{\varepsilon i}\, \sigma_{\varepsilon j}) c_j = \mu_{ai}, \quad i=1,2,\ldots n, \tag{8}$$

where

$$\mu_{ij} \equiv \frac{\overline{t_i t_j}}{\overline{t_a}^2} \tag{9a}$$

is the spatial correlation of the i^{th} and j^{th} temperature deviations,

$$\rho_{ij} \equiv \frac{\overline{\varepsilon_i \varepsilon_j}}{\left(\overline{\varepsilon_i}^2\, \overline{\varepsilon_j}^2\right)^{1/2}} \tag{9b}$$

is the spatial correlation of the observational errors, and

$$\sigma_{\varepsilon i} \equiv \left(\overline{\varepsilon_i}^2 / \overline{t_a}^2\right)^{1/2} \tag{9c}$$

is the normalized standard deviation of observational error. It should be noted that the normalizing factor $\overline{t_a}^2$ is just the error variance of the

"guess" temperature at the analysis point, and that the μ_{ij} are in fact the spatial correlations of the error in the guess temperature field.

In the numerical experiment based on Figure 1, the initial guess is assumed to be provided by a 12-hour forecast. Spatial correlations of forecast errors have not been determined for the NMC prediction models; however, such correlations have been published by Bengtsson and Gustavsson (1971) for 6-, 12-, and 24-hr forecasts of the 500 mb geopotential height using a quasi-geostrophic model. Their spatial correlations are approximately independent of location or directional orientation; that is, they are approximately homogeneous and isotropic in character. Thus, the forecast error correlation is a function only of separation distance s between the i^{th} and j^{th} points. Bengtsson and Gustavsson's experimentally determined curve for 12-hour forecast error correlations as a function of s is displayed in Figure 2. This curve is well approximated by the analytic function

$$\mu_{ij}(s) = e^{-k_\mu s^2} \tag{10}$$

where $k_\mu = 1.56 \times 10^{-6} km^{-2}$. We will assume that (10) with the same value of k_μ applies to the spatial correlation of temperature forecast errors as well.

A reliable determination of the spatial correlation of satellite errors has not yet been made. Preliminary calculations based upon comparisons between temperature cross-sections constructed from radiosonde and from regression-derived soundings show spatial correlations of temperature differences that are comparable to correlations in Figure 2 but that fall off somewhat more quickly with increasing separation distance. With this in mind, we decided to

"model" ρ_{ij} with the functional form

$$\rho_{ij} = e^{-k_\rho s^2} \tag{11}$$

assigning two different values to k_ρ: $k_\rho = k_\mu$, $k_\rho = 4\ k_\mu$. The latter value gives a spatial correlation curve (11) which agrees more closely with the radiosonde - satellite temperature comparisons, decreasing from 1 to 0.37 (1/e) in about 400 km (compare with Figure 2). A third series of calculations was made with $k_\rho = \infty$ ($\rho_{ij} = 0$ when $i \neq j$) to compare results from the first two experiments with those for uncorrelated data.

Once the weights c_i are determined, the mean square analysis error is given by

$$\overline{E^2} = \overline{t_a^2}\left[1 - \sum_{i=1}^{n} \mu_{ai}\ c_i\right]. \tag{12}$$

Thus, the mean square analysis error is a fraction of the error variance of the 12-hr forecast. Although the observational error variance and the spatial correlations do not appear explicitly in this expression, their influence enters through the weights c_i determined by solving equations (8).

The effect of observational density on the resulting analysis error was simulated by varying the observational spacing h in Figure 1 from 100 to 1600 km. One might argue that, when h is small, additional observations beyond the 12 shown in Figure 1 should be included in order to correctly simulate increasing density. However, a trial run with h = 100 km and 20 additional observations extending the grid of Figure 1 resulted in very little difference in the analysis error, as the outer observations received virtually no weight in the analysis. Besides, any practical real-time analysis method must limit predictors to a relatively small number which is less than 12 in existing optimum interpolation analysis schemes.

For the limit h = 0 (all observations coincide), the mean square analysis error may be expressed as an explicit function of $\bar{\rho}$ and σ_ε:

$$\overline{E^2} = \overline{t_a^2}\,\frac{\sigma_\varepsilon^2[1 + (n-1)\bar{\rho}]}{n + \sigma_\varepsilon^2[1 + (n-1)\bar{\rho}]} \tag{13}$$

Here $\bar{\rho}$ denotes the mean value of the inter-observational error correlation, and it is assumed that all observations are of the same type, hence $\sigma_{\varepsilon i} = \sigma_{\varepsilon j} \equiv \sigma_\varepsilon$. When $\bar{\rho} = 0$, (13) reduces to

$$\left(\overline{E^2}/\overline{t_a^2}\right)_{h=0,\bar{\rho}=0} = \frac{\sigma_\varepsilon^2}{n + \sigma_\varepsilon^2}\,, \tag{14a}$$

an error expression appropriate to radiosonde data.

On the other hand, if observational errors are spatially correlated according to (11), then $\bar{\rho} \to 1$ as $h \to 0$ and

$$\left(\overline{E^2}/\overline{t_a^2}\right)_{h=0,\bar{\rho}=1} = \frac{\sigma_\varepsilon^2}{1 + \sigma_\varepsilon^2}\,, \tag{14b}$$

the form appropriate for satellite observations. It is evident that the former decreases without limit as n becomes larger, whereas the latter reaches a limiting value which does not depend on the number of observations.

3. Results

The set of equations (8) with n = 12 was solved numerically for the weights c_i using the iterative method of conjugate gradients (Beckman, 1960). It was assumed that all 12 observations are of the same type and hence have the same statistical error level. The solutions were obtained for normalized observational errors σ_ε of 0, 0.25, 0.50, and 1.00. (The normalized observational error is the ratio of the standard deviation of the observational error to that of the forecast error.) The correlation coefficients μ_{ij} and ρ_{ij} were determined

from the geometry of Figure 1 and from (10) and (11). In (10), k_μ was assigned the value $1.56 \times 10^6 km^{-2}$. In (11), k_ρ was assigned the values k_μ, $4\ k_\mu$, and ∞. The observational spacing h varied from 100 to 1600 km.

Once the weights c_i were obtained for a particular set of conditions, the normalized analysis error

$$\sigma_a \equiv \left(\overline{E^2}/\overline{t_a^2}\right)^{\frac{1}{2}}$$

was obtained from (12). The results for σ_a as a function of h are plotted in Figure 3 for $\sigma_\varepsilon = .25$, Figure 4 for $\sigma_\varepsilon = .50$, and Figure 5 for $\sigma_\varepsilon = 1.00$. The limiting values of σ_a for $h = 0$ were obtained from (14a,b). In each of the figures, the dotted curve for $\sigma_\varepsilon = 0$ (perfect observations) is shown for comparison. The other three curves are for the three values of k_ρ that are assumed. The solid curve for $k_\rho = \infty$ is appropriate for conventional data with random, independent errors, whereas the dashed curves apply to data with spatially correlated errors such as is assumed for satellite data.

The dotted curve indicates that portion of the analysis error which results from spatial interpolation of the observations to the analysis point; it is seen to be of negligible importance when h is less than about 400 km. For closely spaced observations, then, nearly all of the analysis error is a result of the observational errors. For widely spaced observations, on the other hand, the analysis error is primarily a result of the spatial interpolation.

The figures clearly indicate that the analysis error is less for independent observations ($k_\rho = \infty$) than for observations with spatially correlated errors. Of these, the case with the higher degree of correlation ($k_\rho = k_\mu$) has the larger analysis error. For both correlated cases, however, the analysis error

is almost constant for all values of h less than 500 km. (The increase in σ_a, as h approaches zero, to the value given by (14b) in the $k_\rho = 4\ k_\mu$ case is a consequence of assuming perfect correlation of observational errors in the limit h = 0.) These results suggest that there is little information to be gained from having satellite observations spaced closer than about 500 km. This spacing corresponds to a value of the field correlation μ of about 0.5 its maximum value (Figure 2).

In Figures 6, 7, and 8, the normalized analysis error σ_a is shown as a function of both the normalized observational error σ_ε and the observational spacing h for $k_\rho = \infty$, $4\ k_\mu$ and k_μ respectively. Where the isopleths of σ_a are approximately vertical, the analysis error is primarily dependent on the spacing of observations. Where the isopleths are approximately horizontal, analysis error is a function of the errors in observation and cannot be reduced by increasing the density of observations. The figures show clearly that, for both cases with spatially correlated errors, there is no increase in analysis accuracy associated with an increase in observation density to a spacing that is less than about 500 km. For values of h less than 200 km (Figure 6) or 300 km (Figure 7), analysis errors are at least twice as large as the corresponding errors for the uncorrelated case.

The weights c_i obtained by solving (8) are of secondary interest here. In a real analysis situation, the analyzed value is given by (3) once the weights have been determined. In this experiment, the observed values $\hat{t}_i$ have been left unspecified. Because of the symmetry of the observational array (Figure 1), the four inner observations all receive the same weight, as do the eight outer

ones. Generally, the inner observations receive the bulk of the weight, the outer ones comparatively little. When the observational error level σ_ε is high or when h is large, the weights c_i tend to decrease in magnitude, effectively giving greater importance to the guess value $\tilde{T}_a$ in the final analysis.

4. Summary and Conclusions

A simple numerical experiment has been performed which extends the study of Alaka and Elvander (1972a, 1972b) to include the effects of spatially correlated errors. Results indicate that this spatial error correlation reduces the information content of point observations compared to that for data with random errors. They suggest that increasing the density of observations beyond a certain threshold (a spacing of about 500 km in the experiment) will yield little or no improvement in analyses produced from satellite soundings. This statement assumes that the sounding errors are spatially correlated on a scale that approaches the scale of the baroclinic waves. Although the assumption appears reasonable and has been used in simulation studies (Baumhefner and Julian, 1972), such correlations have not yet been measured. The problem is made difficult by the fact that operational satellite soundings are produced only over the oceans where it is impossible to define a "true" temperature or height field from radiosonde data.

It shouldbe emphasized that our conclusions apply only to the case of direct analysis of temperature or height data. Errors in the analysis of gradient quantities, such as thermal wind fields from satellite temperatures, are likely to be lower for spatially correlated than for random observational errors.

Acknowledgments

We would like to acknowledge that similar calculations have been performed but not published by Dr. Lennart Bengtsson of the Swedish Meteorological and Hydrological Institute.

References

Alaka, M. A., and R. C. Elvander, 1972a: Optimum interpolation from observations of mixed quality, Mon. Wea. Rev., 100, 612-624.

_____ 1972b: Matching of observational accuracy and sampling reduction in meteorological data acquisition experiments. J. Appl. Meteor., 11, 567-577.

Baumhefner, D. P., and P. R. Julian, 1972: The reference-level problem: Its location and use in numerical weather prediction. J. Atmos. Sci., 29, 285-299.

Beckman, F. S., 1960: The solution of linear equations by the conjugate gradient method. Mathematical Methods for Digital Computers, Vol. 1, 62-72. (Ed. by A. Ralston and H. S. Wilf. John Wiley and Sons, New York, 293 pp.)

Bengtsson, L., and N. Gustavsson, 1971: An experiment in the assimilation of data in dynamical analysis. Tellus, 23, 328-336.

Gandin, L. S., 1963: Objective Analysis of Meteorological Fields. Gidrometeorologicheskoe Izdatelstvo (Hydrometeorological Institute), Leningrad. Translated by Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.

McMillin, L. M., et al., 1973: Satellite infrared soundings from NOAA spacecraft. NOAA Technical Rept. NESS-65. Washington, D.C., 112 pp.

Captions for Figures

Fig. 1 Array of 12 observations for the numerical analysis experiment. Small circles are observation locations and h is the spacing between them. The analysis point is indicated by X. (Same as Figure 1, Alaka and Elvander, 1972b.)

Fig. 2 Spatial correlation of 12-hour forecast errors of 500 mb geopotential height as a function of separation distance s. (After Bengtsson and Gustavsson, 1971.)

Fig. 3 Normalized standard deviation of analysis error σ_a as a function of observational spacing h for $\sigma_\varepsilon = 0$ and for three cases where $\sigma_\varepsilon = .25$. The curve $k_\rho = \infty$ applies to uncorrelated observational errors.

Fig. 4 Same as Fig. 3, but for $\sigma_\varepsilon = 0$ and .50.

Fig. 5 Same as Fig. 3, but for $\sigma_\varepsilon = 0$ and 1.00.

Fig. 6 Contours of normalized standard deviation of analysis error σ_a as a function of normalized standard deviation of observational error σ_ε and observational spacing h for the case $k_\rho = \infty$.

Fig. 7 Same as Fig. 6, but for the case $k_\rho = 4\ k_\mu$.

Fig. 8 Same as Fig. 6, but for the case $k_\rho = k_\mu$.

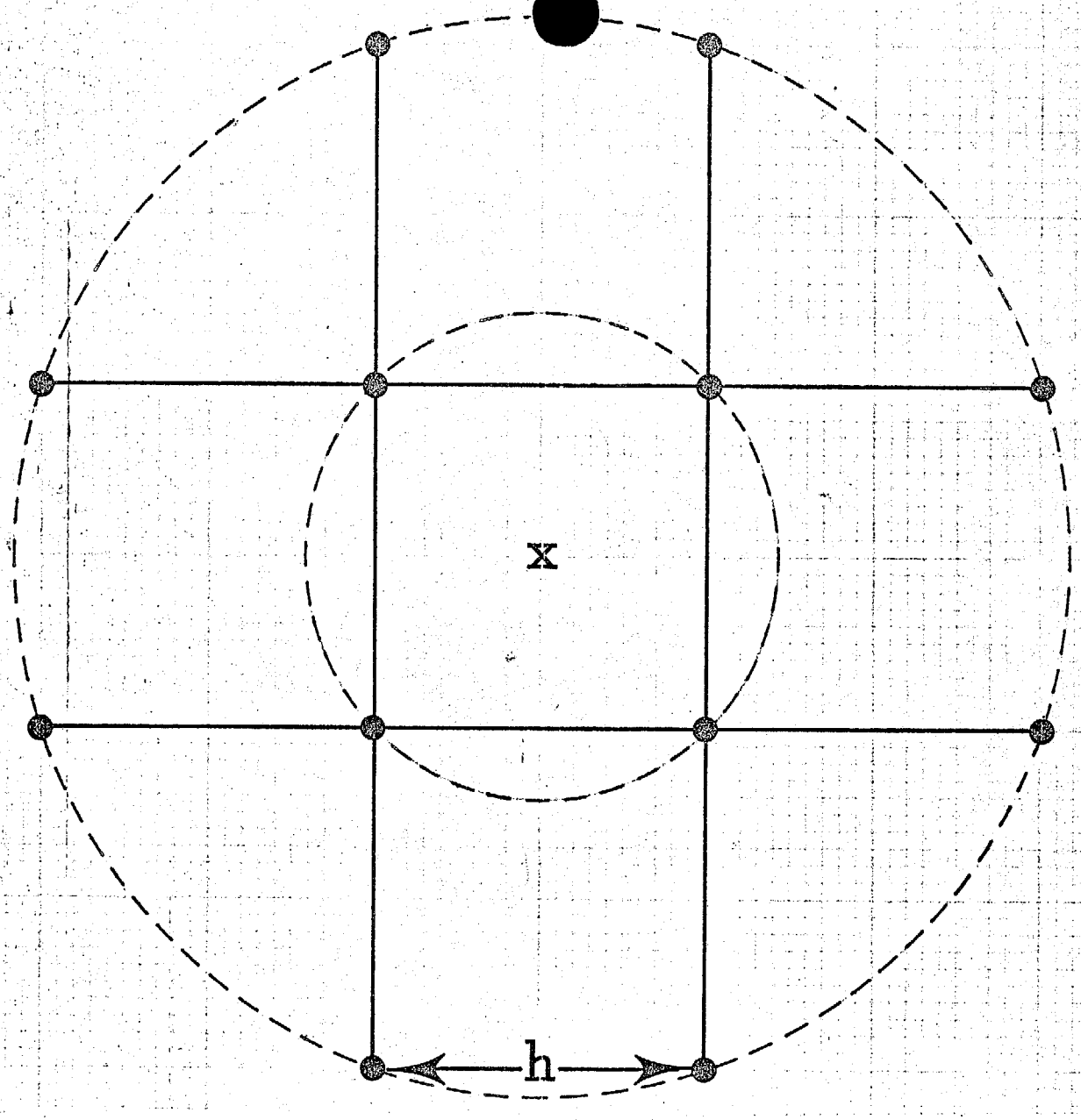

Figure 1 Array of 12 observations for the numerical analysis experiment. Small circles are observation locations and h is the spacing between them. The analysis point is indicated by X. (Same as Figure 1, Alaka and Elvander, 1972b.)

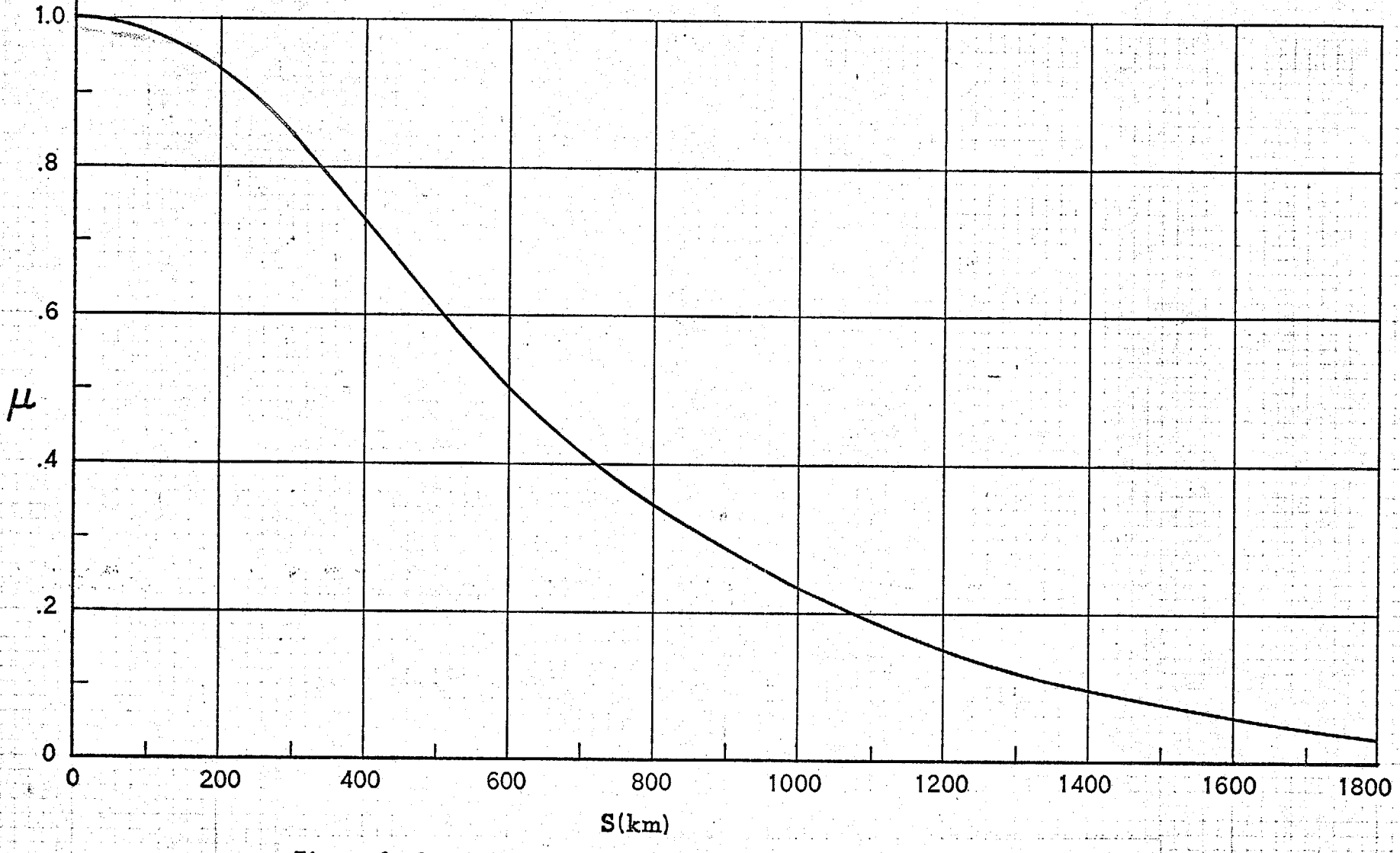

Figure 2 Spatial correlation of 12-hour forecast errors of 500 mb geopotential height as a function of separation distance s. (After Bengtsson and Gustavsson, 1971.)

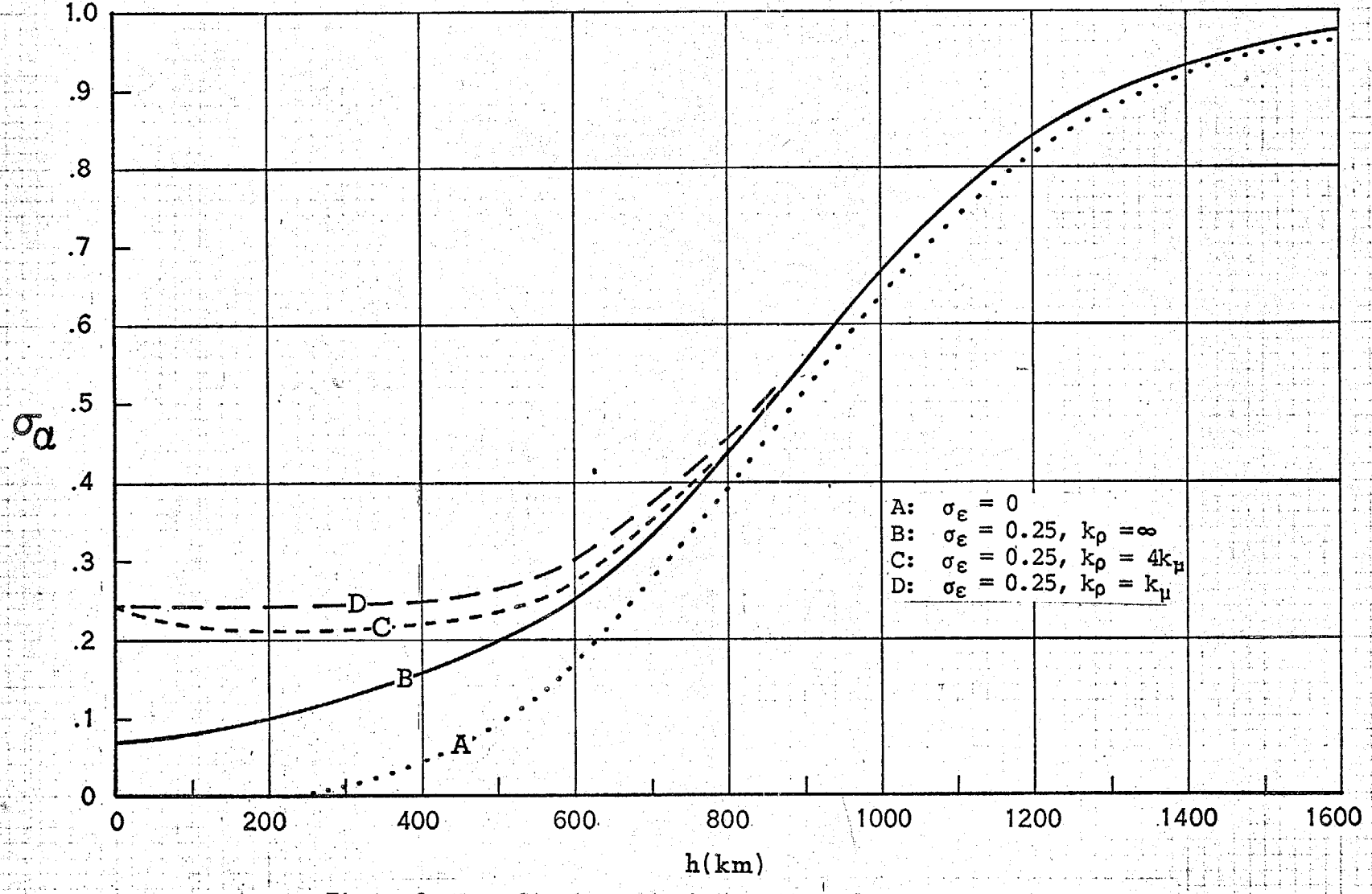

Figure 3 Normalized standard deviation of analysis error σ_a as a function of observational spacing h for $\sigma_\varepsilon = 0$ and for three cases where $\sigma_\varepsilon = .25$. The curve $k_\rho = \infty$ applies to uncorrelated observational errors.

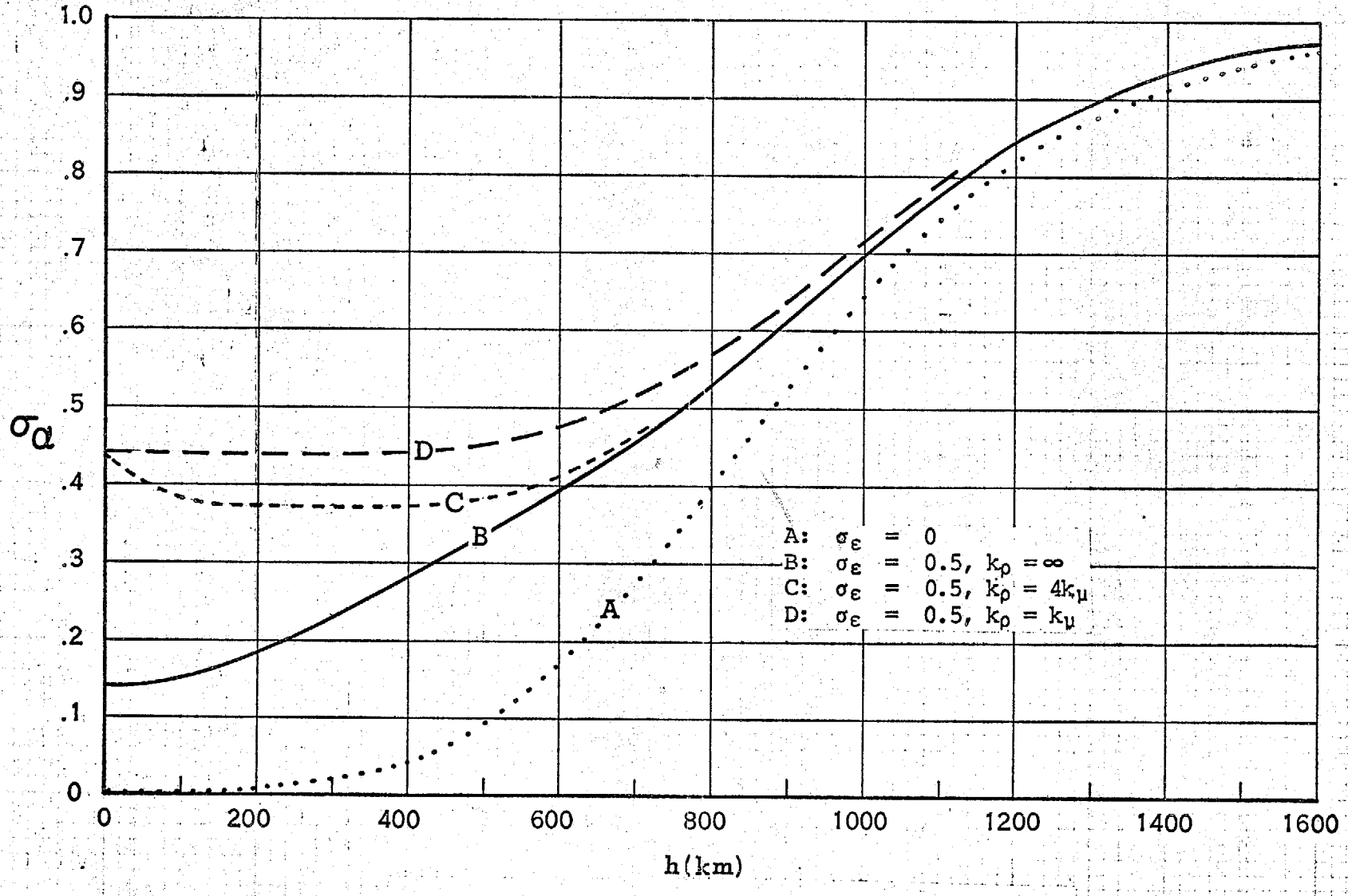

Figure 4 Same as Figure 3, but for $\sigma_\varepsilon = 0$ and .50.

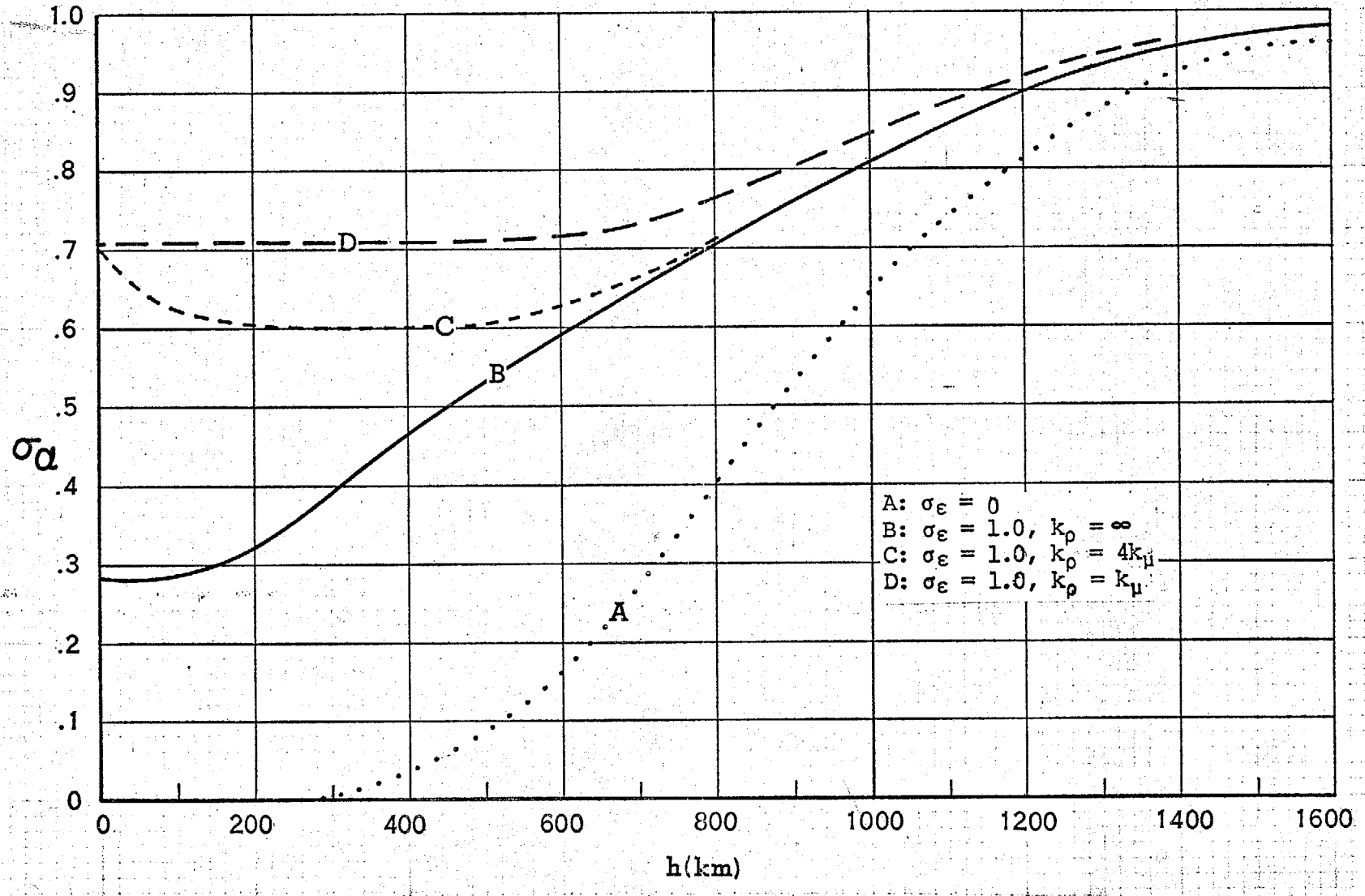

Figure 5 Same as Figure 3, but for $\sigma_\varepsilon = 0$ and 1.00.

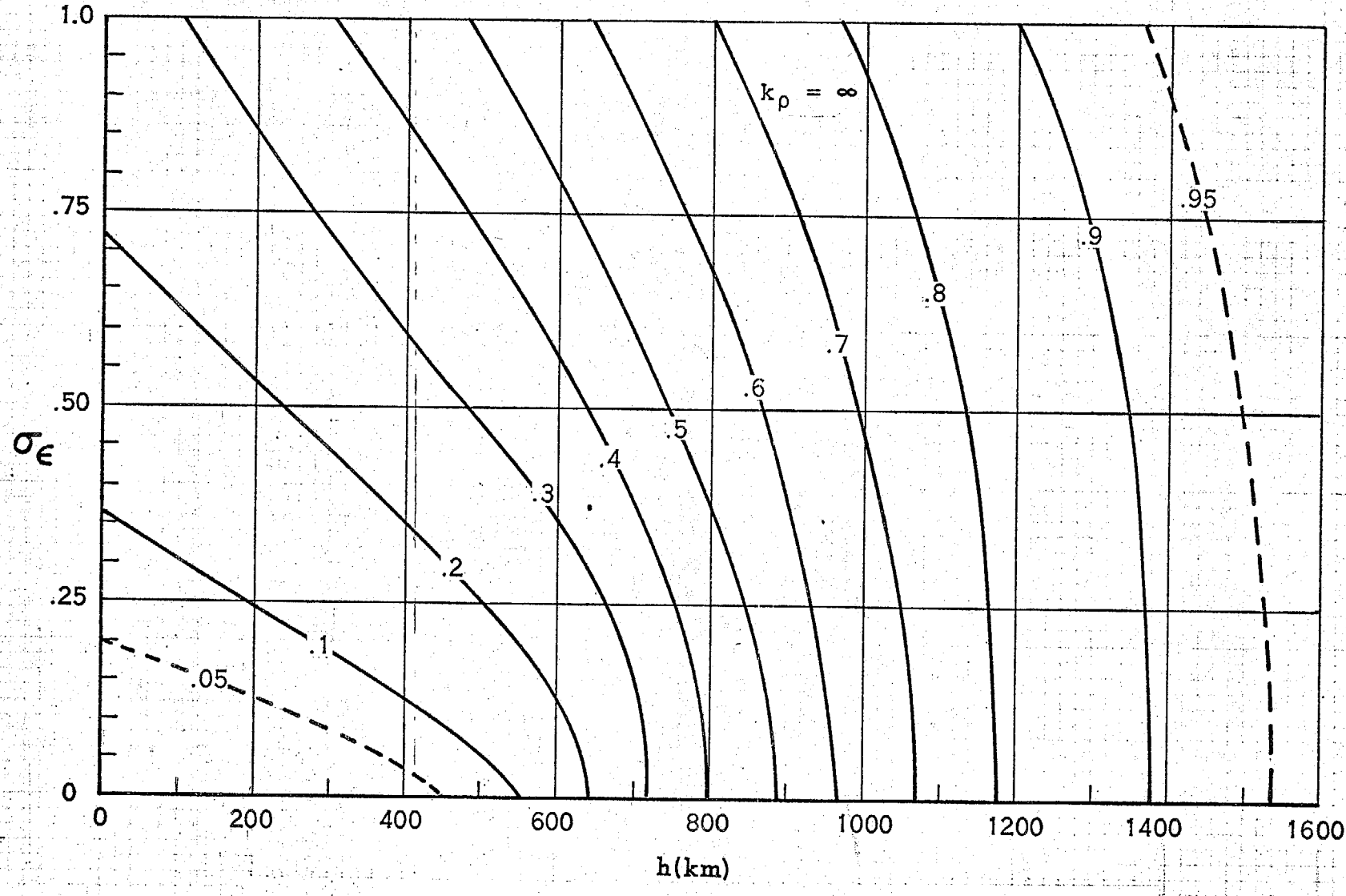

Figure 6 Contours of normalized standard deviation of analysis error σ_α as a function of normalized standard deviation of observational error σ_ϵ and observational spacing h for the case $k_\rho = \infty$.

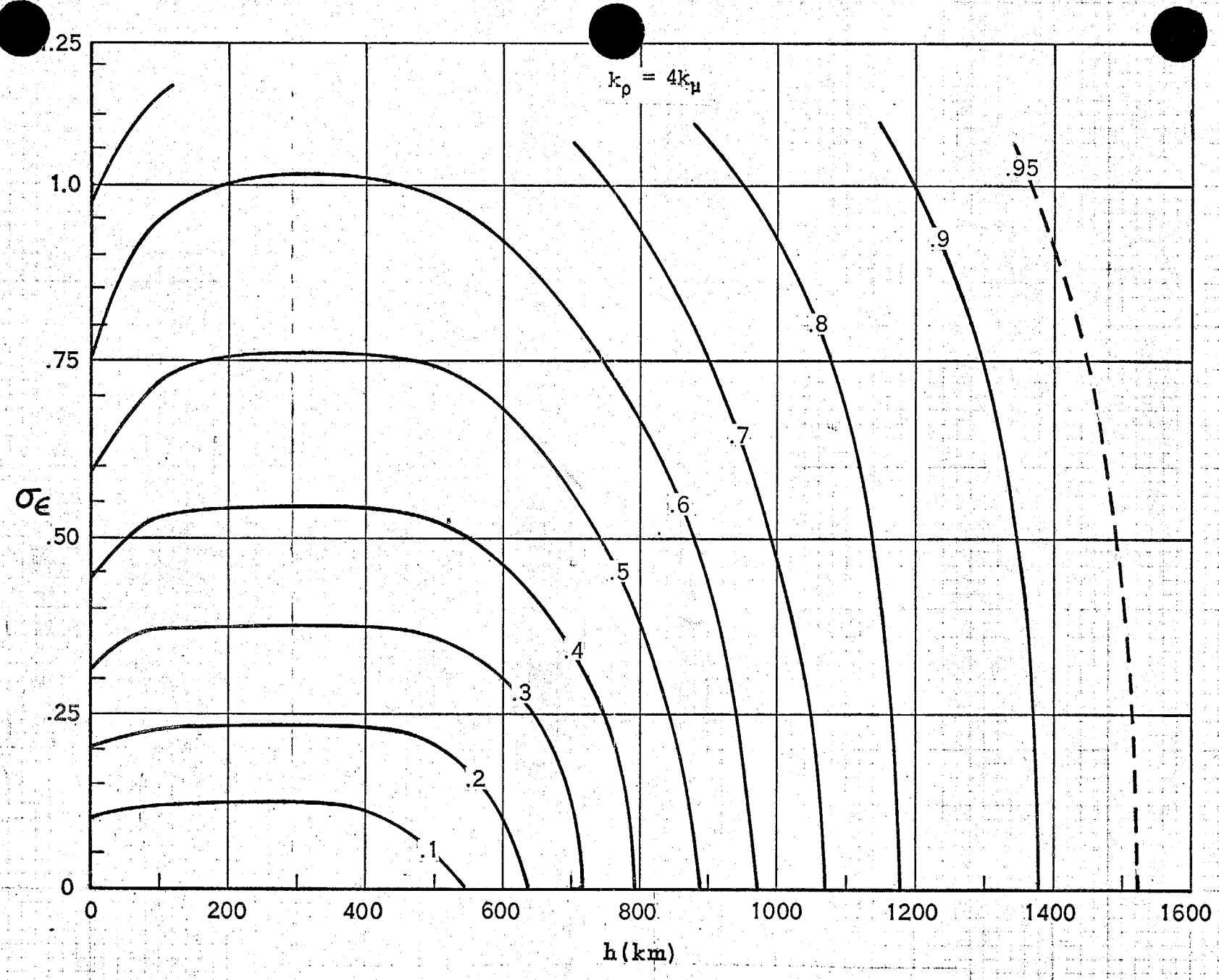

Figure 7 Same as Figure 6, but for the case $k_\rho = 4\ k_\mu$.

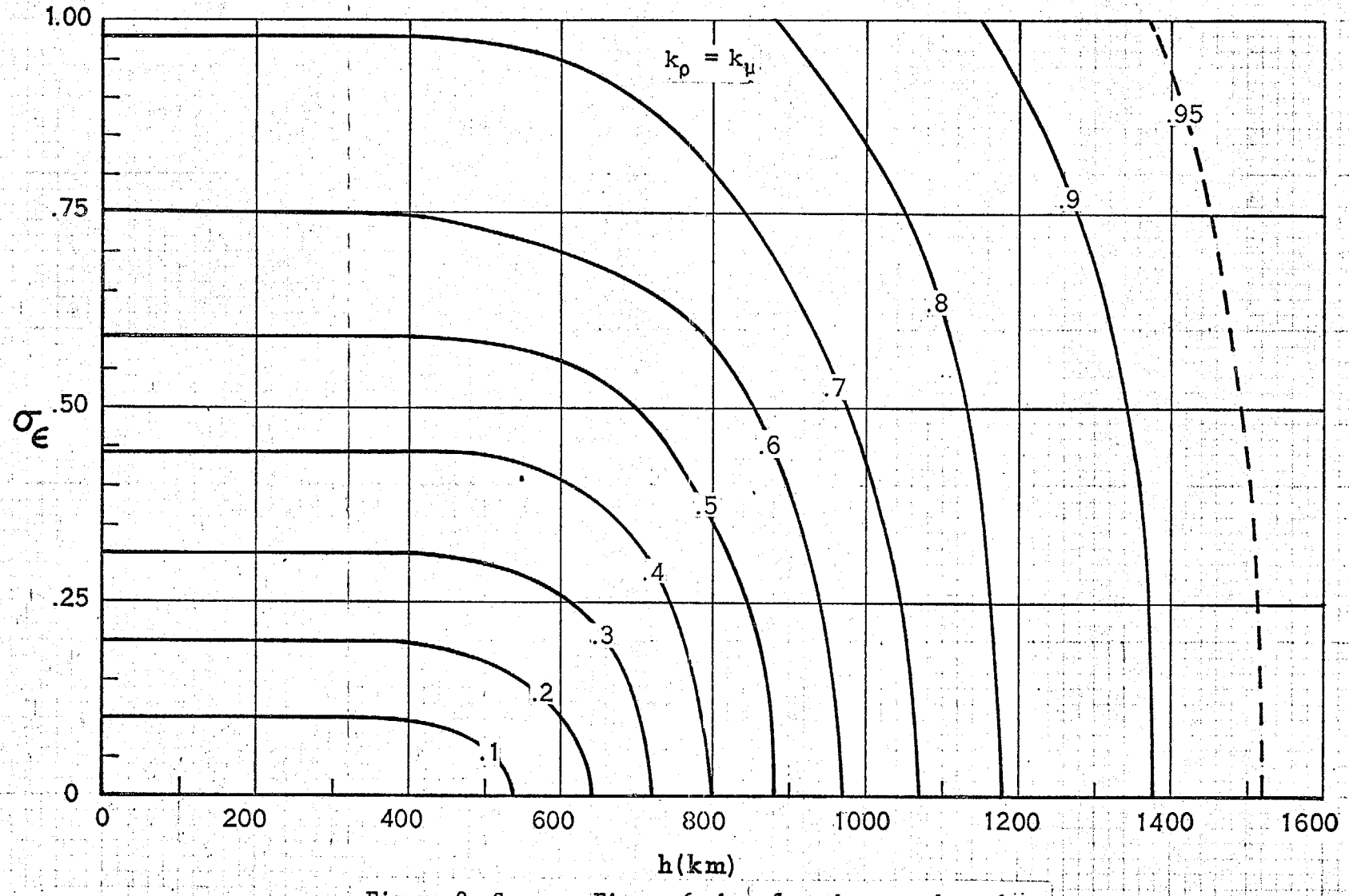

Figure 8 Same as Figure 6, but for the case $k_\rho = k_\mu$.