RESOLUTION ERRORS ASSOCIATED WITH GRIDDED PRECIPITATION FIELDS

C. J. WILLMOTT* and M. L. JOHNSON

Center for Climatic Research, Department of Geography, University of Delaware, 216 Pearson Hall, Newark, DE 19716, USA

Received 4 February 2004
Revised 22 February 2005
Accepted 16 May 2005

ABSTRACT

Spatial-resolution errors are inherent in gridded precipitation (P) fields – such as those produced by climate models and from satellite observations – and they can be sizeable when P is averaged spatially onto a coarse grid. They can also vary dramatically over space and time. In this paper, we illustrate the importance of evaluating resolution errors associated with gridded P fields by investigating the relationships between grid resolution and resolution error for monthly P within the Amazon Basin.

Spatial-resolution errors within gridded-monthly and average-monthly P fields over the Amazon Basin are evaluated for grid resolutions ranging from 0.1° to 5.0°. A resolution error occurs when P is estimated for a location of interest within a grid-cell from the unbiased, grid-cell average P. Graphs of January, July and annual resolution errors versus resolution show that, at the higher resolutions (<3°), aggregation quickly increases resolution error. Resolution error then begins to level off as the grid becomes coarser. Within the Amazon Basin, the largest resolution errors occur during January (summer), but the largest percentage errors appear in July (winter). In January of 1980, e.g., resolution errors of 29, 52 and 65 mm – or 11, 19 and 24% of the grid-cell means – were estimated at resolutions of 1.0°, 3.0° and 5.0°. In July of 1980, however, the percentage errors at these three resolutions were considerably larger, that is, 15%, 27% and 33% of the grid-cell means. Copyright © 2005 Royal Meteorological Society.

KEY WORDS: resolution error; precipitation; gridded data

1. INTRODUCTION

Spatial-resolution errors associated with gridded precipitation fields are of increasing concern to climatologists (Giorgi and Marinucci, 1996; Bresnahan and Miller, 1997). Time-integrated precipitation (P) is often highly variable over space, especially when the integration period is short, and much of its (P’s) spatial variability can be lost when the grid is too coarse. Resolution errors reduce the opportunity to observe local extremes, and the high-frequency spatial variability in P. Typical climate-model grid cells (on the order of 2° or 5.0 × 10^4 km^2), for instance, can miss important components of the spatial variability in P (Houghton et al., 2001; Chapter 7.2.3, page 431). Satellite-derived, historical P fields also have tended to be rather coarse; that is, they often have been made available at resolutions of 2.5° or lower (Johnson, 2002). Higher-resolution fields (e.g. at 1°) have been and are being produced, although validation tends to indicate relatively high grid-cell errors (Huffman et al., 2001).

When estimating P from a gridded P field (e.g. from a satellite- or climate-model-derived P field), the error in estimating P at a single point within a grid cell from the estimated grid-cell average (\( \bar{P} \)) is (\( \bar{P} - P \)). This error can be partitioned usefully into ‘bias’ and ‘resolution’ components; that is, (\( \bar{P} - P \)) = [(\( \bar{P} - \bar{P} \)) + (\( \bar{P} - P \))], where \( \bar{P} \) is a reliable estimate of the ‘true’ grid-cell average of P made at the same
spatial resolution as $\hat{P}$. Since $\hat{P}$ and $P$ generally vary – but often differently – with the grid resolution, the relative contributions that bias and resolution components make to error also vary with resolution.

A measure of average, total error associated with a gridded $P$ field ($E_T$) can be obtained from a spatial integration of the magnitudes of ($\hat{P} - P$). Similarly, an average bias error ($E_B$) can be derived from a spatial integration of the magnitudes of ($\hat{P} - P$), while a mean spatial-resolution error ($E_R$) can be taken from a spatial integration of the magnitudes of ($P - \hat{P}$). Average spatial-resolution error ($E_R$) – the focus of this paper – generally increases as a $P$ field is spatially averaged onto progressively coarser grids. It is also true that $E_T = E_B + E_R$, with conservative measures of $E_B$ and $E_R$. A serious problem can arise when two or more gridded $P$ fields are compared with one another, and only $E_B$ is evaluated (cf. Costa and Foley, 1998); that is, without a concomitant evaluation of $E_R$, it is impossible to assess the accuracy ($E_T$) of a gridded data set. Interpretations of $E_B$ also can be faulty because of the inverse correlation that usually exists between grid resolution and $E_B$; that is, $E_B$ tends to get smaller as the grid resolution becomes coarser.

In this paper, we illustrate the importance of estimating and interpreting $E_R$ by investigating relationships between grid resolution and $E_R$ for monthly $P$ within the Amazon Basin. We first create very-high-resolution gridded-monthly $P$ fields over the Amazon Basin – by spatially interpolating from high-resolution raingage observations of monthly $P$ – and then aggregate (regrid) those very-high-resolution gridded $P$ fields to progressively lower (more coarse) spatial resolutions. For each of the lower-resolution gridded fields, we evaluate the expected resolution error as the sub-grid-scale variability exhibited within the corresponding, very-high-resolution gridded field.

2. MONTHLY RAINGAGE RECORDS

South American station records of monthly $P$ – compiled, documented and gridded by Willmott and Webber (1998) and Webber and Willmott (1998) – serve as the primary source of data for our analysis. Willmott and Webber’s and Webber and Willmott’s precipitation-data archives for South America (together referred to as WW’s data) were obtained from several existing data sets. These include the Global Historical Climatology Network Version 1 (GHCNv1); an archive compiled by Frederick Wernstedt, with World Meteorological Organization (WMO) assistance; and an archive assembled from the Brazilian National Water and Electric Power Department (Divisão Nacional de Águas e Energia Elétrica – DNAEE) and other data as part of the Earth Observing System (EOS) Amazon Project, and provided to us by Tom Dunne. Monthly $P$ observations within the WW time-series archive fall within the period from 1846 to 1993, with the greatest number of station records available in the early 1970s.

Only those WW station records from within or near (to reduce edge effects) the Amazon Basin are used in our analysis. Over the time period of interest (1960–1990), the number of stations available varied from 1010 to 1758. Within the Amazon Basin, station density associated with the 31-year climatology was high – approximately 95 per million km² (Figure 1) – even with some stations dropping out and others coming on line throughout the 31-year period. One of the years with good station coverage was 1980. Within the Basin, it had roughly 82 stations per million km²; therefore, 1980 was chosen to represent the spatial variability associated with shorter-term integrals of $P$. These two average station densities are uncommonly high, when compared to station densities over other time periods and parts of the world (Willmott et al., 1994).

3. GRIDDED RAINGAGE FIELDS

Very-high-resolution, gridded-monthly raingage ($P$) fields for the Amazon Basin were spatially interpolated (according to Willmott et al., 1985) from the station data described earlier to a 0.05° spherical grid. When applied to monthly $P$ observations, the Willmott et al. (1985) interpolation algorithm estimates monthly $P$ totals at each grid node (grid-cell ‘center’), and generally performs well (Chen et al., 2002). A spatial average
of our grid-node estimates of $P$ can then be obtained by multiplying each grid node $P$ by its 0.05° grid-cell area, summing the products, and then dividing by total area. It should be noted that the 0.05° resolution is higher than the average spatial resolution of our raingage data ($\approx 0.92°$), which makes our resolution error estimates for grids finer than 0.92° extrapolations. This also suggests that our interpolated 0.05° $P$ fields are likely to be smoother than the real $P$ fields resolved onto a 0.05° grid; as a consequence, our estimates of resolution error may be somewhat low. These finely gridded $P$ data then were spatially averaged (aggregated) to a range of more coarse, grid resolutions; that is, to resolutions of 0.10°, 0.25°, 0.50°, 0.75°, 1.00°, 1.50°, 2.00°, 2.50°, 3.00°, 3.50°, 4.00°, 4.50° and 5.00°.

4. PRECIPITATION CLIMATOLOGY

A very brief description of the rainfall climatology of the Amazon Basin is offered here to assist readers in interpreting our findings. Paraphrasing Johnson (2002), the Basin is exceptionally wet, with an average rainfall of 2124 mm year$^{-1}$ (177 mm month$^{-1}$). There is a distinct pattern of decreasing $P$ from north to south (Figure 2). Maximum $P$ falls on the northwestern and northeastern regions of the Basin, while minima appear along the south and southwestern edges. A single wet and single dry season occur over the course of the year, and their geographies are typified by the average January (Figure 3) and average July (Figure 4) $P$ fields. Wet season conditions persist from November through May, while the dry season lasts from June through October. Average $P$ for January (wet season) is approximately 257 mm month$^{-1}$, which strongly contrasts with the 90 mm month$^{-1}$ for July (dry season).

5. RESOLUTION ERROR

Spatial-resolution errors (sub-grid-scale variability) associated with longer-term integrals of $P$ are evaluated with monthly (January and July) and annual averages of $P$ taken over the 31-year period from 1960 through 1990. Resolution errors within fields of shorter-term integrals of $P$ are illustrated through an examination of monthly (January and July) and average $P$ for a single year, 1980.
For each $P$ field of interest, our measure of resolution error ($E_{|R|}$), in mm month$^{-1}$, is taken as

$$E_{|R|} = \frac{1}{n \times m} \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} | \bar{F}_i - 0.05P_{ij} |$$

(1)

where $\bar{F}_i$ represents a ‘reliable’ spatial mean of the original 0.05° WW $P$ values averaged up to resolution $R \times 0.05P_{ij}$ is a WW $P$ value at its interpolated (0.05°) resolution, and $w_{ij}$ is an area weight used to adjust for changes in latitude. The summation is over both the $n$ lower-resolution grid cells that fall within the Amazon Basin.
Figure 4. Thirty-one-year average July rainfall (mm month$^{-1}$) on the Amazon Basin at 0.05° resolution. This figure is available in colour online at www.interscience.wiley.com/ijoc.

Figure 5. Resolution errors ($E_{|R|}$s) over the Amazon Basin for the 1960–1990 climatology. This figure is available in colour online at www.interscience.wiley.com/ijoc.

Basin, and the $m$ very-high-resolution grid cells that fall within each lower-resolution cell. Spatial cross-validation suggests that $0.05P_{ij}$ is an unbiased estimate of $P$ at a grid node. And, when the grid-cell means of $P$ (the $\bar{P}_{ij}$s) are unbiased, the expected value of our resolution error is a measure of the sub-grid-scale variability in $P$ about the corresponding grid-cell means. For each resolution of interest, $E_{|R|}$ was calculated for the Amazon Basin, and then plotted on a graph of $E_{|R|}$ versus resolution (Figures 5 and 6).

Our use of an average-error statistic ($E_{|R|}$) that is based on the absolute values of the differences rather than on the squares of the differences was prompted by the fact that squared-difference-based statistics (like the root mean squared error (RMSE)) are functions of both the average error and the shape of the distribution of errors. Squared-difference-based statistics, therefore, are inappropriate measures of average error, even though they are commonly used for this purpose.
6. INTERPRETATIONS

A strong relationship between resolution error ($E_{|R|}$) and the seasonal rainfall cycle is apparent. Resolution error is greater during the summer (January) than during the winter (July), by as much as a factor of 2 or 3 (Figures 5 and 6). This is due to the fact that most of the $P$ falls during the summer, and its convective origin gives it a relatively high-frequency spatial variability. Since temporal averaging tends to damp the spatial variability, this summer–winter difference is considerably more pronounced in 1980 (Figure 6) than it is in the 31-year climatology (Figure 5). Other individual years, from within the 31 years of record, were also evaluated (not shown), and the associated $E_{|R|}$s were similar in magnitude to the $E_{|R|}$s for 1980, although the other January values tended to be somewhat lower than those for 1980 while the other July values were slightly higher.

As the grid becomes more coarse (Figures 5 and 6), $E_{|R|}$ increases asymptotically. The increases in $E_{|R|}$, from resolutions of 0.05° to $\approx 1.0^\circ$, are dramatic, while, by $\approx 3.0^\circ$, the increases in $E_{|R|}$ begin to level off. The shape of our relationships between $E_{|R|}$ and resolution are similar to the one observed by Bresnahan and Miller (1997), although our leveling-off point is not as well defined. The largest errors appear in January (summer), but the largest percentage errors (relative to the mean $P$) occur in July (winter).

Within the January climatology, errors – at 1.0°, 3.0° and 5.0° resolution – are 18 mm, 33 mm and 40 mm, or 7%, 13% and 16%. In January of 1980, however, the corresponding resolution errors are considerably higher at 29 mm, 52 mm and 65 mm, even though the percentage errors increased only slightly to 11%, 19% and 24%. Within the July climatology, errors – once again at 1.0°, 3.0° and 5.0° resolution – are 10 mm, 20 mm and 27 mm, or 11%, 22% and 30%. In July of 1980, the corresponding resolution errors and percentage errors are similar at 11 mm, 20 mm and 24 mm, and at 15%, 30% and 33% respectively. The annual $P$ fields exhibit intermediate $E_{|R|}$s, although the annual percentage errors are the lowest observed. Temporal averaging simply creates a smoother $P$ field. Within the annual climatology, the errors are 10 mm, 18 mm and 22 mm, or 6%, 10% and 12%; while, for 1980, they are 14 mm, 24 mm and 28 mm, or 9%, 15% and 17%.

Resolving the same precipitation field to different resolutions, of course, yields similar broad-scale $P$ patterns, but important detail can be lost with aggregation. Larger resolution errors tend to occur in areas of higher $P$, while, within areas of lower $P$, resolution errors tend to be smaller. This correlation between the spatial mean and spatial variability is not uncommon in $P$, since it is bounded on the lower end by zero and tends to be positively skewed. For all fields examined, resolution error also tends to increase with topographic ruggedness. This supports Giorgi and Marinucci’s (1996) finding that fine detail from higher resolutions are more important in areas with coastlines and nontrivial topographic variability (e.g. in the Andes Mountains and at the mouth of the Amazon River).
7. CONCLUSIONS

Resolution errors are inherent in gridded precipitation fields, such as those produced by climate models and from satellite observations, and can be a significant problem when \( P \) is averaged spatially onto a coarse grid. Within the Amazon Basin – at resolutions finer than \( \approx 3.0^\circ \) – resolution error grew rapidly with aggregation. At resolutions larger than \( \approx 3.0^\circ \), increased aggregation produced diminishing increases in resolution error. While the resolution error is substantially greater during the summer (January) than during the winter (July), higher percentage (of the \( P \) mean) errors appear in the winter. It is clear that spatial-resolution errors in monthly \( P \) can be large, and can vary dramatically in space and time. Nonetheless, the shapes of our resolution error functions are similar to one another and probably are asymptotic in nature; therefore, future work should verify whether these are really different functions or just a single function with time- and place-varying coefficients. When a gridded \( P \) field is compared with another gridded \( P \) field, without an assessment of resolution error, it is impossible to assess the accuracy of either gridded \( P \) field.

ACKNOWLEDGEMENTS

This research was supported, in part, by the Delaware Space Grant College Consortium and Fellowship Program, as well as by NASA’s Tropical Rainfall Measuring Mission (TRMM) and Large-Scale Biosphere-Atmosphere (LBA) program (grant numbers NAG5-4829 and NAG5-8443). Special thanks go to Scott Webber, who helped develop and evaluate the South American precipitation archive, and to Cynthia Brewer for the color sequence used on our maps. The constructive comments of two anonymous reviewers are also most appreciated.

REFERENCES


Johnson ML 2002. Estimating precipitation over the Amazon basin from satellite and in-situ measurements, Masters thesis, University of Delaware, Newark, DE, 68.


