* Manuscript
Click here to view linked References

Effective Interpolation of Incomplete Satellite-Derived Leaf-Area Index Time Series for the Continental United States

Jordan S. Borak[a,*] and Michael F. Jasinski[b]

[a]Wyle Information Systems, 7515 Mission Drive, Suite A100, Lanham, MD 20706, USA
Fax: +1-301-352-0437; email: Jordan.S.Borak@nasa.gov

[b]NASA Goddard Space Flight Center, Hydrological Sciences Branch, Code 614.3, Greenbelt, MD 20771, USA
Fax: +1-301-614-5808; email: Michael.F.Jasinski@nasa.gov

Keywords: Interpolation; Leaf-area index; MODIS; Time series

*Corresponding author.

ABSTRACT

Many earth science modeling applications employ continuous input data fields derived from satellite data. Environmental factors, sensor limitations and algorithmic constraints lead to data products of inherently variable quality. This necessitates interpolation of one form or another in order to produce high quality input fields free of missing data. The present research tests several interpolation techniques as applied to satellite-derived leaf-area index, an important quantity in many global climate and ecological models. The study evaluates and applies a variety of interpolation techniques for the Moderate Resolution Imaging Spectroradiometer (MODIS) Leaf-Area Index Product over the time period 2001-2006 for a region containing the conterminous United States. Results indicate that the accuracy of an individual interpolation technique depends upon the underlying land cover. Spatial interpolation provides better results in forested areas, while temporal interpolation performs more effectively over non-forest cover types. Combination of spatial and temporal approaches offers superior interpolative capabilities to any single method, and in fact, generation of continuous data fields requires a hybrid approach such as this.

1. INTRODUCTION

The acquisition of continuous gridded fields of many land surface properties is a critical challenge within global climate and ecological modeling. As models move toward finer and finer spatial resolutions, there is increasing reliance on satellite-derived data products for model initialization or forcing (e.g., Oleson and Bonan, 2000; Niemeyer and Vogt, 2001). Leaf-area index (LAI), or the leaf area per unit ground area, is especially important in many global terrestrial processes for computing the exchanges of energy, water and other gases (Running et al., 1986).

Despite the increasing need however, quality and availability of LAI data products are variable, whether considered at the global scale or pixel scale, due to such factors as persistent cloud cover, elevated aerosol loading, or sparse observations owing to insufficient repeat frequency of the satellite sensor. Products that are of consistently poor quality or are unavailable over large regions cannot be used at all. In such instances, the parameter of interest may be estimated using an appropriate interpolation technique. Certainly, low quality inputs would be expected to negatively impact any derived products using those data. It is also important to avoid spurious assumptions in the interpolation process so that the resulting estimate does not actually reduce the usability of the data.

Various interpolation techniques of differing degree of complexity and applicability exist (e.g., Hartkamp et. al, 1999). These can be based on statistical or deterministic, and spatial or temporal approaches, and include such methods as kriging (Cressie, 1991),

inverse distance weighting (Shepard, 1968), cubic splines (e.g., Mitášsová and Hofierka (1993); Alexandrov et al. (2001)), and spectral analysis (e.g., Kondrashov and Ghil, 2006). Alternatively, if the requirements of the data user permit, it is possible to alleviate some missing data issues by spatiotemporal filtering, or by coarsening spatial or temporal resolution by aggregating the original data into products with lower frequency sampling. However, there is no clear understanding as to which approach produces the best results for a given land cover type or season. Consequently a systematic evaluation of the strengths and weaknesses of some of the more common methods is warranted.

The intent of the current research is to develop a systematic approach for interpolating missing data from a time series of MODIS LAI data product over the continental United States. The methodology relies entirely on remote sensing data products, avoiding the drawbacks of "look-up" methods or ancillary requirements. A variety of individual and hybrid techniques is tested using actual satellite observations as "truth" in a data denial experiment. Results are evaluated by comparing the interpolated data to the original data for each interpolation approach.

2. DATA AND METHODS

2.1. Study Area

The study area for this research consists principally of the conterminous United States, plus southern Canada, northern Mexico and some Caribbean land mass, as shown in Figure 1. The intent is to use a geographic domain of sufficient extent that regional

effects are relevant. Areas outside the United States are not considered germane in the sense that they are merely included by default as the minimum geographic bounding box necessitates their inclusion.

The region is characterized primarily by mid-latitude climatology and vegetation types, although the northern and southern extremes extend into the boreal and subtropical zones. Vegetation types consist mainly of forest types in the north, east and west, while grassland and cropland dominate the interior of the domain. During winter months, areas of missing data tend to be concentrated in the north due to frequent presence of clouds, snow cover, and low sun angles. In the summer, missing data are more prevalent in the east and west due to cloud cover related to cyclonic activity.

2.2. Input Data

The input data employed in this research consist of Collection 4 MODIS land products. Land cover data have been derived from the 2001 Terra MODIS Land Cover Product (Friedl et al., 2002), International Geosphere-Biosphere (Belward, 1996) data layer (Figure 1). Leaf-area index data are compiled for the period 2001 to 2006 from the archived Terra MODIS LAI product, MOD15A2 (Myneni et al., 2002). No additional inputs are necessary. The land cover data are produced on a yearly basis, while the LAI data are produced every eight days as composites of daily data. Quality statistics are included in the LAI product for each pixel (Roy et al., 2002).

The leaf-area index data are screened to remove low quality observations. When LAI data do not meet the following criteria, then for the purposes of this paper, they are considered missing, and an interpolation is required:

1. The main radiative transfer algorithm was used to compute LAI, not the empirical backup method.
2. MODIS Land mandatory QA flags are set to "OK" or "Best" quality, and
3. The cloud flag is labeled either as "clear" or "not defined, assumed clear."

2.3. Methodology

Since there is no single interpolation approach that would be appropriate for every circumstance, several approaches that might address the different categories of missing data are examined. For instance, if missing observations are scattered in a "salt-and-pepper" spatial pattern, techniques such as local window operations may be very effective. Conversely, if extensive areas of missing data are prevalent (e.g., due to widespread cloud cover), temporal approaches may be more effective. Regardless of the approach implemented, an effective sampling design must be in place in order to test the interpolation results. Thus, the analysis is not intended to be exhaustive of all existing techniques, but rather the approaches are representative of the interpolation needs for estimating missing LAI data.

2.3.1. Data Denial Sampling

A data denial approach is implemented in order to compare the estimation capabilities of various interpolation methods. For each IGBP land cover type, 100 samples have been randomly drawn over the spatiotemporal domain of the LAI time series. The samples are required to be of acceptable quality so that it would be sound to compare the original values to interpolated estimates. Then, the observations are removed from the LAI fields so that the interpolation algorithms could be applied.

2.3.2. Spatial Interpolation

Per-Class Means (PCM)

One of the simplest approaches to interpolation is to replace missing observations with their per-class mean. Wherever data are missing, the field is filled by the mean LAI for all locations sharing the same IGBP land cover type as the missing observation during the same time period. The size of the domain over which the per-class mean is computed can vary. This method is advantageous for large domains in that it is insensitive to extensive areas of missing data and it guarantees that all locations contain valid observations in the output data field. A significant drawback is that the per-class mean dampens the data signal, which is of greater concern as domain extent increases. The accuracy of the estimated value decreases with increasing domain size, although calculating the class means on a finer spatial scale comes at the expense of guaranteed complete spatial coverage of the interpolation.

2.3.3. Temporal Interpolation

2.3.3.1. Temporal Linear Averaging (TLA)

Another relatively simple approach to interpolation is to replace missing data with the mean value of the observations recorded for that location in the preceding time period and the subsequent time period. In situations where only one of the two surrounding observations is usable, the interpolated value assumes the LAI from the single available adjacent time period. Although the TLA method includes no spatial component, it provides an efficient means of estimating temporal LAI interpolates since usually the relative change in LAI values is small between 8-day compositing periods for most of the year. Still, this technique may fail in locations characterized by persistently poor data quality.

2.3.3.2. Temporal Cubic Splines (TCS)

In earth science studies, splines are generally used for spatial interpolation (e.g., Mitáš and Mitášová, 1999). Possible implementations include two-dimensional per-period splines or one-dimensional per-location splines. In the current research it is not practical to produce two-dimensional splines for each composite in the time series. Instead, individual splines for each location are generated, effectively performing a separate temporal interpolation for each pixel within the spatial domain. A primary benefit of the TCS method is that interpolation is possible over extensive areas of missing data within a single time period. Disadvantages include 1) lack of spatial context; 2) poor

characterization of processes as gaps between valid observations increase and 3) unconstrained function behavior at endpoints.

2.3.3.3. Climatology (TC)

For locations where spatially and temporally neighboring observations are sparse, it may be beneficial to interpolate to climatological values. The concept of climatology in this context denotes the average usable LAI value observed in a given location, during a specific compositing period over all years in the time series. In other words, when an observation is missing, it is replaced by the average LAI observed over that specific location during the same time period from 2001-2006. This more rigorous definition of climatology contrasts with per-class mean climatology commonly used in earth system modeling.

2.3.4. Hybrid Approaches

Kang Method (H1)

The approach of Kang *et al.* (2005) consists of both spatial and temporal components. The primary interpolation is a spatial per-class mean defined over a small window. It replaces missing observations with the mean of all observations in the surrounding 5x5 pixel window where the land cover type is identical to that of the missing observation. If the primary method is unsuccessful, the algorithm selects the previous observation (interpolated or otherwise) for the location of interest. This is an attractive option as it is relatively uncomplicated in nature, and attempts to exploit the available spatial and

temporal information. As the authors note, however, the interpolation algorithm is still susceptible to an inability to interpolate in situations where data loss is extensive in space and only slightly persistent in time.

3. RESULTS

3.1. Comparison of Data Denial Results

The above spatial, temporal and hybrid interpolation approaches, as well as combinations of those approaches, were applied to the data denial experiment. Once the interpolated estimates for a given approach were generated for each sample withheld via data denial, the estimates were plotted against the true LAI values. Next, linear regressions were fitted to the bivariate data. The associated R-square values for the regressions were used to gauge the efficacy of each interpolation method. These results are summarized in Table 1 for the principal land cover classes within the study area.

The spatial per-class means approach using the entire domain (PCM) resulted in relatively poor estimates of missing LAI data, particularly for the non-forest classes. This was unsurprising as the extent of the domain was such that vegetation phenology would be inconsistent across the domain for a given class. For example, wide variations in latitude within the domain would tend to dictate shifts in cropping types and schedules which would affect the results of the *cropland* and *cropland mosaics* classes.

A second per-class mean approach narrowed the domain used for estimation to the immediate area (5x5 kilometers/pixels) surrounding the missing observation. This method (LPCM) was similar to Kang et al (2005) without the temporal backup component. It actually improved the fits in all categories relative to the results for PCM, as shown in Table 1.

The temporal linear averaging (TLA) produced encouraging results, particularly for the non-forest classes, and for *deciduous broadleaf forest*. This was likely due in large part to the location-specific nature of the algorithm. It would be expected that land cover types with strong phenologies would be well characterized by an interpolation technique that focused on the *in situ* temporal evolution of the vegetation.

Similarly, temporal cubic splines (TCS) performed well when applied to land cover types with well-defined phenologies. As with the TLA method, *deciduous broadleaf forest* and cropped areas were well-suited for temporal interpolation. Interpolation of the *evergreen needleleaf* and *mixed forest* cover types was not as accurate. This is likely due to the fact that these forest types tend to experience more persistent cloud cover, which would result in more frequent loss of either preceding or subsequent observations, and in turn reducing the ability of splines to capture the temporally local variability in the vegetation.

The use of climatological values (TC) produced results inferior to both of the preceding methods for nearly all of the cover types. Even so, predictive capability was better for non-forest classes than for forest classes, which makes sense given the temporal nature of

the climatology. It may be that intraannual variability was so important for the non-forest classes that the interpolation methods which focused on immediate temporal variation were more effective for those vegetation types. It was somewhat surprising that the results for the forest classes were so discouraging, and this was likely related to interannual variability. Unfortunately, due to the relatively short temporal record of MODIS (2001 being the first complete year), there were insufficient data to develop a true climatology for each pixel location and therefore to test whether interannual and intraannual variability differed in a statistically meaningful way.

The method of Kang *et al.* (2005) (H1) produced encouraging results for nearly all of the land covers listed in Table 1, although not much different from the LPCM approach. However, it was able to provide greater coverage of interpolation estimates than the local per-class mean approach alone, particularly when missing data occurred over extensive regions. By exploiting both the spatial and temporal information in the LAI dataset, the technique was able to estimate missing data values over a broad range of vegetation types. The use of spatial information would be expected to provide improved estimating capability over cover types that do not exhibit pronounced seasonality.

To this point, the H1 method produced promising results, although it did not guarantee that every missing observation could be interpolated. Areas of extensively and/or persistently poor quality data (e.g., cloud cover in mountainous regions) could not be interpolated by the stock algorithm. As such, several modifications were tested in order to produce LAI fields completely free of missing data.

The first modification involved the expansion of Kang's spatial window from 5x5 to 9x9 (H2). The intent was to capture more spatial data in the vicinity of missing observations. Based on earlier results, this would yield improved interpolations, particularly over forested locations.. This did not turn out to be the case for the *evergreen needleleaf forest* class, as indicated in Table 1. The *deciduous broadleaf forest* class fared somewhat better, but still showed a decline in interpolation efficacy when compared with the original H1 method. The fits for those classes actually degraded, while for *mixed forest*, there was little change. The results for the non-forest classes were mixed – there was marked improvement for *grassland*, a slight improvement for *cropland*, and slight declines for the rest of the vegetation types.

The next hybrid method variant focused on enhancing temporal aspects of the H1 algorithm. Instead of examining the previous observation when the primary spatial technique failed, the secondary technique used the TCS splining technique. Interestingly, there was very little difference in results between this approach (labeled H3), and the original H1 method (see Table 1). Given the additional computational resources required

to fit splines for each pixel as opposed to the original H1 algorithm, this approach was rejected.

In an effort to strike a balance between the simple temporal component of H1 (previous time period) and splines, it was determined that a hybrid of localized 5x5 per-class means (LPCM) with temporal linear averaging might improve interpolation, particularly for vegetation classes with strong seasonality. As shown in Table 1, this hybrid method (labeled H4), improved interpolation for *deciduous broadleaf forest* as compared to the original H1 method, but there was little difference otherwise. In fact, results were actually degraded for the grassland class, as compared to the interpolations generated using just the LPCM algorithm with no temporal component at all.

Finally, the H4 method was augmented with a tertiary interpolation component, domain-wide PCM. This change did not alter the statistical results significantly, (method H5 shown in Table 1), however it did guarantee valid data values at every location in the domain. Figures 2-3 depict data denial results obtained for two land cover types. For *evergreen needleleaf forest* (Figure 2), the linear relationship, with an R-square value of 0.78, indicated a reasonably good fit for the interpolation over a wide range of LAI values. In the case of *grassland* (Figure 3), the range of LAI was narrower , and the R-square value of 0.71 was only slightly lower than for the forest class.

3.2. Application of H5 to MODIS imagery

A second set of analyses was conducted to analyze the sequential interpolation of the LAI fields using the H5 approach for a typical compositing period over the continental United States. The goal was to examine which component of the interpolation scheme, LPCM, TLA, or domain PCM, was used to estimate the missing data. Results are depicted graphically in Figures 4-5. Figure 4(a) displays the available MODIS LAI data product following the criteria established in Section 2.2 for the eight-day compositing period from May 17 – May 24, 2006, when green-up was either in progress or had already completed, depending on location within the study area. The figure indicates that missing data were scattered throughout the continent but especially in the eastern section, where large patches of missing data occurred, particularly over the Appalachian Mountains, and the far northeast. This was attributed to an extensive stretch of cloud cover that prevented satellite-based estimates over a large region at that time. Figure 4(b) shows a breakdown of which component of the H5 technique was used in estimating the interpolated data. For most of the continent, the isolated missing data were estimated using the LPCM component, followed by TLA. In the mountains, however, clouds persisted beyond the adjacent dates, and the algorithm resorted to using domain-wide PCM. In the far northeast, clouds were less persistent and the temporal linear averaging component was used. Figure 4(c) displays the resulting complete interpolated LAI field using the entire H5 approach.

It was also of interest to analyze the interpolation results outside the growing season. Figure 5(a) contains screened MODIS LAI data for the eight-day period from December 3 – December 10, 2006. Significant areas of missing data were apparent in the northern extent of the study area, as well as over the central and western United States. Again, it was expected that the temporal component of the interpolation scheme would be dominant, but the results indicated otherwise. As illustrated in Figure 5(b), data missing from the northern and central regions of the study area were mainly interpolated using domain-wide PCM. In wintertime, particularly at high latitudes, it can be difficult to acquire sufficient satellite observations to perform effective LAI retrievals due to cloud cover, snow cover and low sun angles. If this situation was persistent, domain-wide PCM would be the only of the presented methods that could generate interpolates. Figure 5(c) displays the complete LAI field, including interpolates.

An analysis of the temporal variability in the mechanism of interpolation uncovered a distinctive seasonal pattern for each component of the hybrid method. Figure 6 illustrates these patterns. During the early and late phases of the growing season, use of original, non-interpolated data was at its highest frequencies (usually 70 – 90%), and LPCM was responsible for interpolating nearly all of the missing observations. At midsummer, interpolation requirements increased, mostly due to declining data quality in the eastern U.S. associated with the peak of summertime cyclonic activity.

After the end of the growing season, data quality declined considerably, due mainly to the presence of cloud and snow cover in northern regions of the study area, as well as low

sun angles. Figure 5(a) illustrates this graphically. Because these conditions can be persistent, the domain-wide PCM component became crucial, often accounting for 30% or more of the study area.

Figures 7 and 8 also depict the temporal variability in interpolation component, but for individual land cover classes. Figure 7 shows the temporal variability for *evergreen needleleaf forest* over the study area. The basic patterns were similar to those for the full domain, although frequency of interpolation was greater for this class. As indicated in Figure 1, *evergreen needleleaf forest* tends to be located in areas of significant cloud cover (e.g., windward side of mountain ranges) and snow cover (northern U.S. and Canada). The LPCM method's frequency was elevated somewhat, but the domain-wide PCM interpolation component reached wintertime maxima of greater than 70% at times.

For *grassland*, the frequency of interpolation decreased as compared to the domain as a whole. As illustrated in Figure 1, most grassland in the study area is located in relatively dry regions of the continent. The decrease in precipitation frequency led to higher quality data, particularly during the summer.

4. DISCUSSION

From the results generated using the various interpolation methods, it became clear that no single interpolation technique was superior in all ways (see Table 1). However, several patterns were evident. The TLA method was particularly effective for estimating missing observations within land covers that exhibit strong seasonality. For the *cropland* and *cropland mosaic* classes in particular, spatially-based interpolations would be expected to introduce errors, as crop characteristics may differ considerably even at the 1-km scale.

Interpolation of forest classes was most accurate with the use of spatial techniques. This was borne out by the results for all of the variants of the H1 method, which used a 5x5 LPCM as its primary interpolation algorithm. Temporally-based techniques generally fared more poorly.

In the current research, kriging (Cressie, 1991) might have been used for one-dimensional temporal interpolation, two-dimensional spatial interpolation or multidimensional spatiotemporal interpolation. However, kriging would involve considerable manual effort to select the optimal model and associated parameters. A one-dimensional interpolation would require a separate model for each location, and a two-dimensional approach would necessitate separate models for each time period. Multidimensional kriging of a dataset of this size would have been impractical due to computational constraints.

The use of domain-wide PCM to assure complete coverage in the LAI fields is problematic, as shown in the results in Table 1. Over a domain of this size, these means are not necessarily representative of local conditions, particularly with respect to those classes that exhibit considerable seasonality. The modeling communities typically require complete fields, with no missing data. As one of the primary reasons for generating these fields is for modeling studies, it is essential to force the interpolating algorithm to produce an LAI value at each location during each period of the time series. Alternative methods for doing so merit further investigation.

5. CONCLUSIONS

Interpolation of satellite data for modeling purposes is essential for producing data fields of acceptable quality for modeling applications at continental spatial scales. Satellite data products are inherently of variable quality due to environmental factors and sensor limitations. Additionally, algorithmic constraints often result in data of questionable quality.

From the results presented here, it is difficult (if not impossible) to produce complete data fields using a single simple approach to interpolation. Each of the stand-alone methods tested has weaknesses. A hybrid approach offers a means whereby complete fields can be generated under any input data scenario.

Consequently, it is recommended, at least for the use of MODIS-derived LAI data products, that a hybrid technique be used for interpolation of all LAI data because of the varying characteristics of the land cover classes and the causes of missing data. For forest classes, the approach would be first to use a 5x5 LPCM method, followed by TLA, followed by a domain-wide PCM to ensure there are no missing data. This is essentially similar to Kang et al. (2005) but with an adjustment to the secondary method from simply using the previous observation to employing temporal linear averaging and adding a teritiary fallback of domain-wide per-class means. For non-forest classes, the data denial analysis indicates one should also use a hybrid approach but switch the order of the first two components. This results in substantial improvements over H1.

The use of domain-wide PCM as a backup method is not optimal due to its overgeneralization, but it is necessary. In an ideal situation, the fallback component would be less prone to problems related to generalization. This is an area for future research.

Acknowledgements

This research was funded by the NASA Earth Science Division Applied Sciences Homeland Security Program and the Terrestrial Hydrology Program.

REFERENCES

Alexandrov, M.D., Cairns, B., Lacis, A.A., Carlson, B.E., 2001. Integrated study of MFRSR-derived parameters of atmospheric aerosols and trace gases over the ARM Cloud and Radiation Testbed (CART) site extended facilities. Eleventh ARM Science Team Meeting Proceedings.

Belward, A.S. (Ed.), 1996. The IGBP-DIS global 1 km land cover data set (DISCover): proposal and implementation plans. IGBP-DIS Working Paper No. 13, IGBP-DIS Office, Meteo-France, Toulouse, France, 61 pp.

Cressie, N.A.C., 1991. Statistics for Spatial Data. John Wiley and Sons, Inc., New York, pp. 105-275.

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A.,Cooper, A., Baccini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing Environ. 83, 287–302.

Hartkamp, A.D., De Beurs, K., Stein, A., White, J.W., 1999. Interpolation Techniques for Climate Variables. NRG-GIS Series 99-01, Mexico, D.F.: CIMMYT.

Kang, S., Running, S.W., Zhao, M., Kimball, J.S., Glassy, J., 2005. Improving continuity of MODIS terrestrial photosynthesis products using an interpolation scheme for cloudy pixels. Int. J. Remote Sens. 26, 1659-1676.

Kondrashov, D., Ghil, M., 2006. Spatio-temporal filling of missing points in geophysical data sets. Nonlin. Proc. Geophys. 13, 151-159.

Mitáš, L., Mitášová, H., 1999, Spatial Interpolation. In: Longley, P., Goodchild, M.F., Maguire, D.J., Rhind, D.W. (Eds.), Geographical Information Systems: Principles, Techniques, Management and Applications. John Wiley and Sons, New York, pp. 481-492

Mitášová, H., Hofierka, J., 1993. Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis. Math. Geol. 25, 657-669.

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R., Running, S.W., 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sens. Environ. 83, 214–231.

Niemeyer, S., Vogt, J., 2001. Integration of operationally available remote sensing and synoptic data for surface energy balance modelling and environmental applications on the regional scale. In: Beniston, M., Verstraete, M.M. (Eds.), Remote Sensing and Climate

Modeling: Synergies and Limitations. Advances in Global Change Research 7, Kluwer Academic Publishers, Dordrecht, pp. 329-343.

Oleson, K., Bonan, G.B., 2000. The effects of remotely-sensed plant functional type and leaf area index on simulations of boreal forest surface fluxes by the NCAR land surface model. J. Hydrometeor. 1, 431–446.

Roy, D.P., Borak, J.S., Devadiga, S., Wolfe, R.E., Zheng, M., Descloitres, J., 2002. The MODIS Land product quality assessment approach. Remote Sensing Environ. 83, 62-76.

Running, S., Peterson, D.L., Spanner, M.A,, Teuber, K.B., 1986. Remote sensing of coniferous forest leaf area. Ecology 67, 273-276.

Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 23rd ACM National Conference, pp. 517-524.

Land Cover Type	Interpolation Technique									
	Domain Per-Class Means (PCM)	**Localized Per-Class Means (LPCM)**	**Temporal Linear Averaging (TLA)**	**Temporal Cubic Splines (TCS)**	**Climatology (TC)**	**Kang Method (H1)**	**Kang 9x9 (H2)**	**H1/TCS Hybrid (H3)**	**LPCM/ TLA Hybrid (H4)**	**LPCM/TLA/ PCM Hybrid (H5)**
Evergreen Needleleaf Forest	0.34	0.76	0.71	0.46	0.30	0.75	0.66	0.77	0.77	0.78*
Deciduous Broadleaf Forest	0.65	0.96*	0.87	0.78	0.56	0.91	0.85	0.90	0.96*	0.95
Mixed Forest	0.58	0.84	0.75	0.60	0.41	0.83	0.84	0.85*	0.84	0.84
Open Shrublands	0.00	0.86	0.84	0.47	0.40	0.88*	0.83	0.88*	0.88*	0.87
Woody Savannas	0.13	0.89	0.87*	0.67	0.67	0.80	0.74	0.78	0.79	0.74
Grassland	0.04	0.71	0.97*	0.84	0.57	0.71	0.86	0.71	0.71	0.71
Cropland	0.14	0.69	0.82*	0.75	0.74	0.69	0.73	0.69	0.69	0.69
Cropland Mosaics	0.24	0.81	0.90*	0.83	0.62	0.81	0.78	0.81	0.81	0.81

Table 1: R-squared values of linear regression between actual observations and interpolated estimates for the tested methods. The results presented here show outcomes for dominant land covers. Highest R-squared values for each class are indicated by asterisks.

Figure Captions

Figure 1: MODIS Land Cover Product data mapped to the conterminous U.S. domain of this study. Cover types are labeled with the IGBP legend.

Figure 2: Data denial results for *evergreen needleleaf forest* using the LPCM/TLA/PCM hybrid method, H5. The R-squared value for the linear fit is 0.78.

Figure 3: Data denial results for *grassland* using the LPCM/TLA/PCM hybrid method, H5. The R-squared value for the linear fit is 0.71.

Figure 4: Interpolation of MODIS LAI Product over the study site during the growing season, May 17 – May 24, 2006. The original data field is screened to remove low quality observations, with missing data over land mapped in black (a), interpolated via the various methods of the H5 hybrid approach (b) and output as a complete data field (c). Locations that are never interpolated, such as water and barren land are mapped in white.

Figure 5: Interpolation of MODIS LAI Product over the study site after the growing season, December 3 – December 10, 2006. The original data field is screened to remove low quality observations, with missing land data mapped in black (a), interpolated via the various methods of the H5 hybrid approach (b) and output as a complete data field (c). Locations that are never interpolated, such as water and barren land are mapped in white.

Figure 6: Interpolation method frequency, full domain.

Figure 7: Interpolation method frequency, *evergreen needleleaf forest.*

Figure 8: Interpolation method frequency, *grassland.*

Figure 01

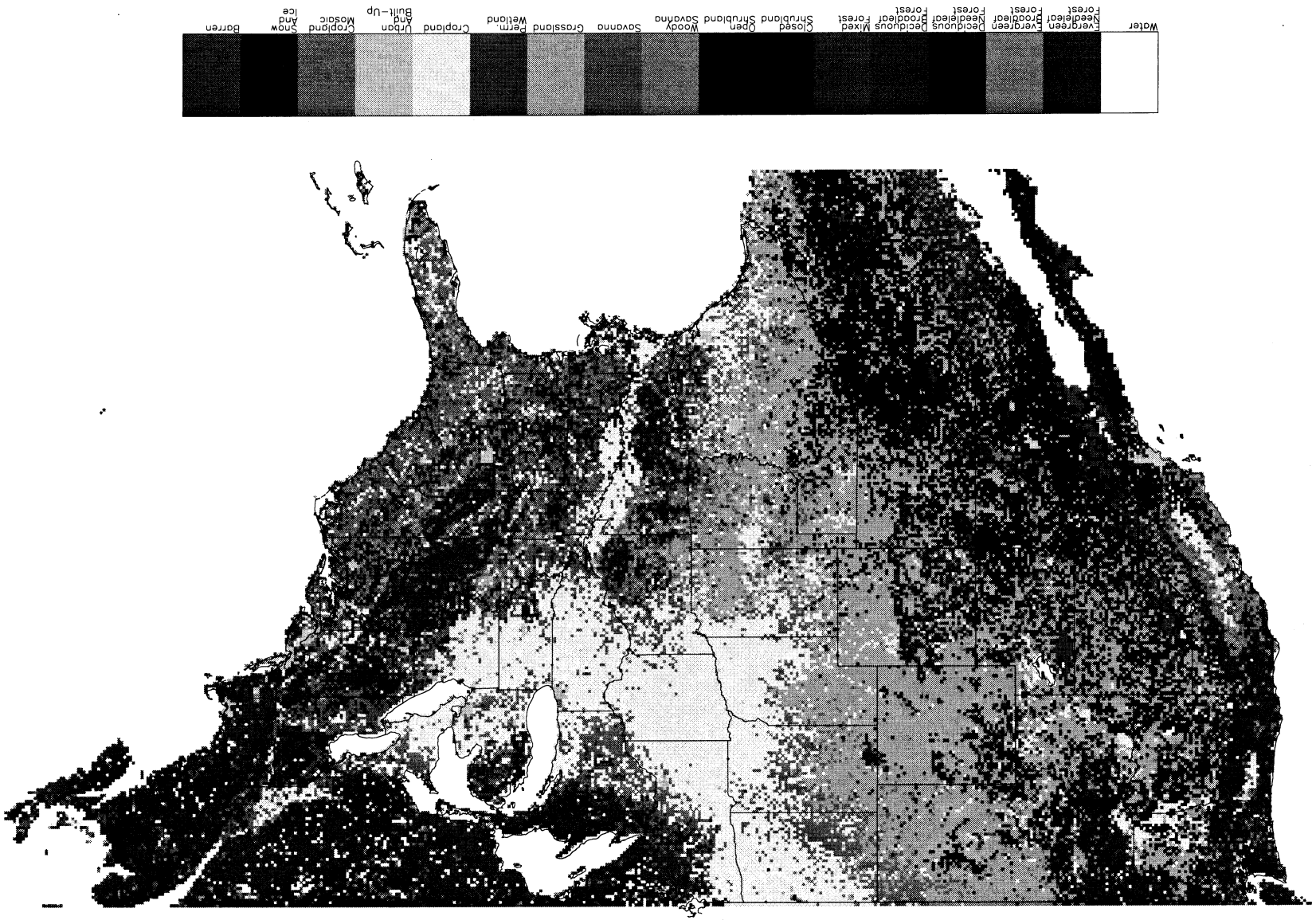

Water
Evergreen Needleleaf Forest
Evergreen Broadleaf Forest
Deciduous Needleleaf Forest
Deciduous Broadleaf Forest
Mixed Forest
Closed Shrubland
Open Shrubland
Woody Savanna
Savanna
Grassland
Perm. Wetland
Cropland
Urban And Built-Up
Cropland Mosaic
Snow And Ice
Barren

Figure 02

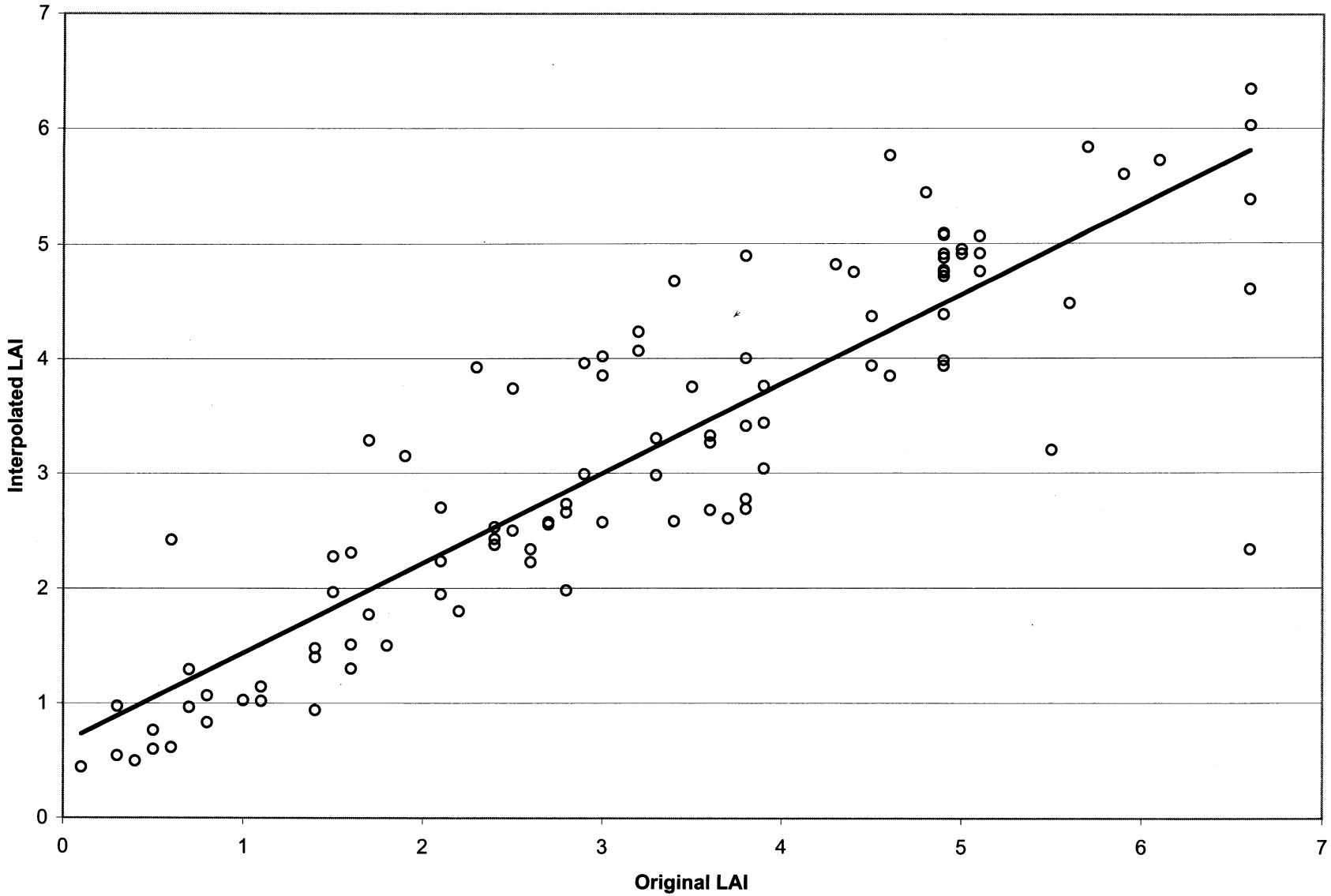


Figure 03

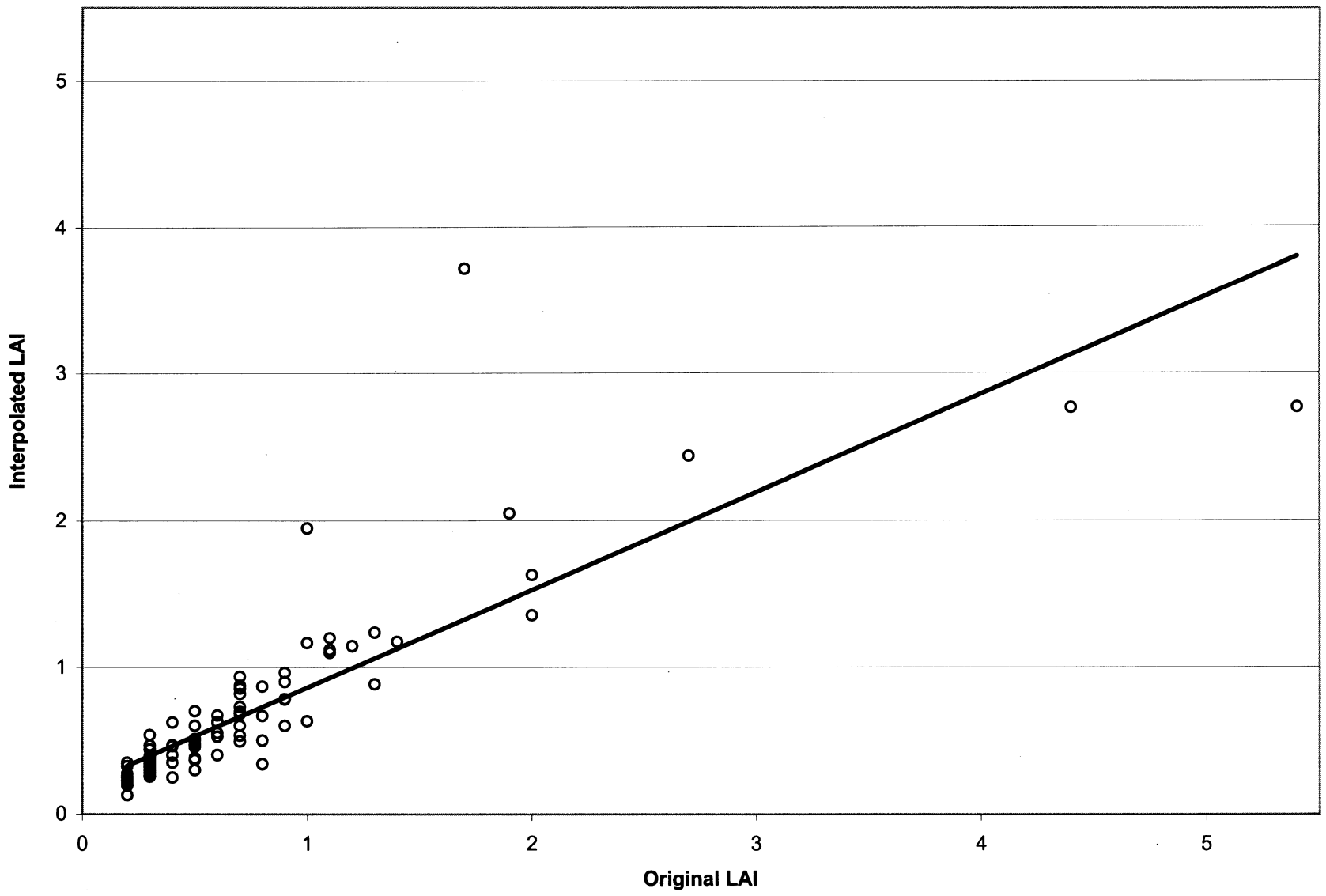


Figure 04a

Figure 04b

Figure 04c

Figure 05a

Figure 05b

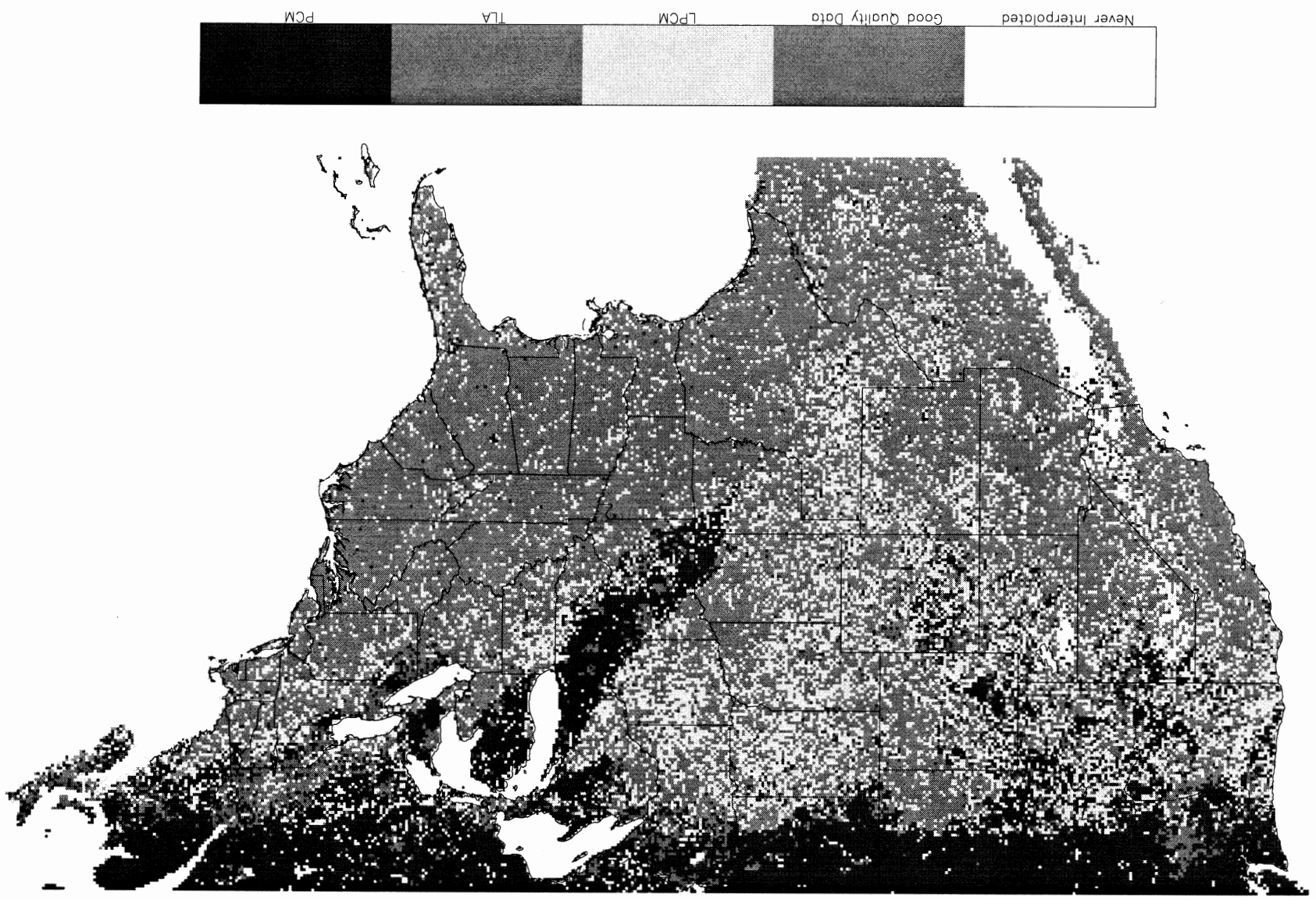


Figure 05c

Figure 06

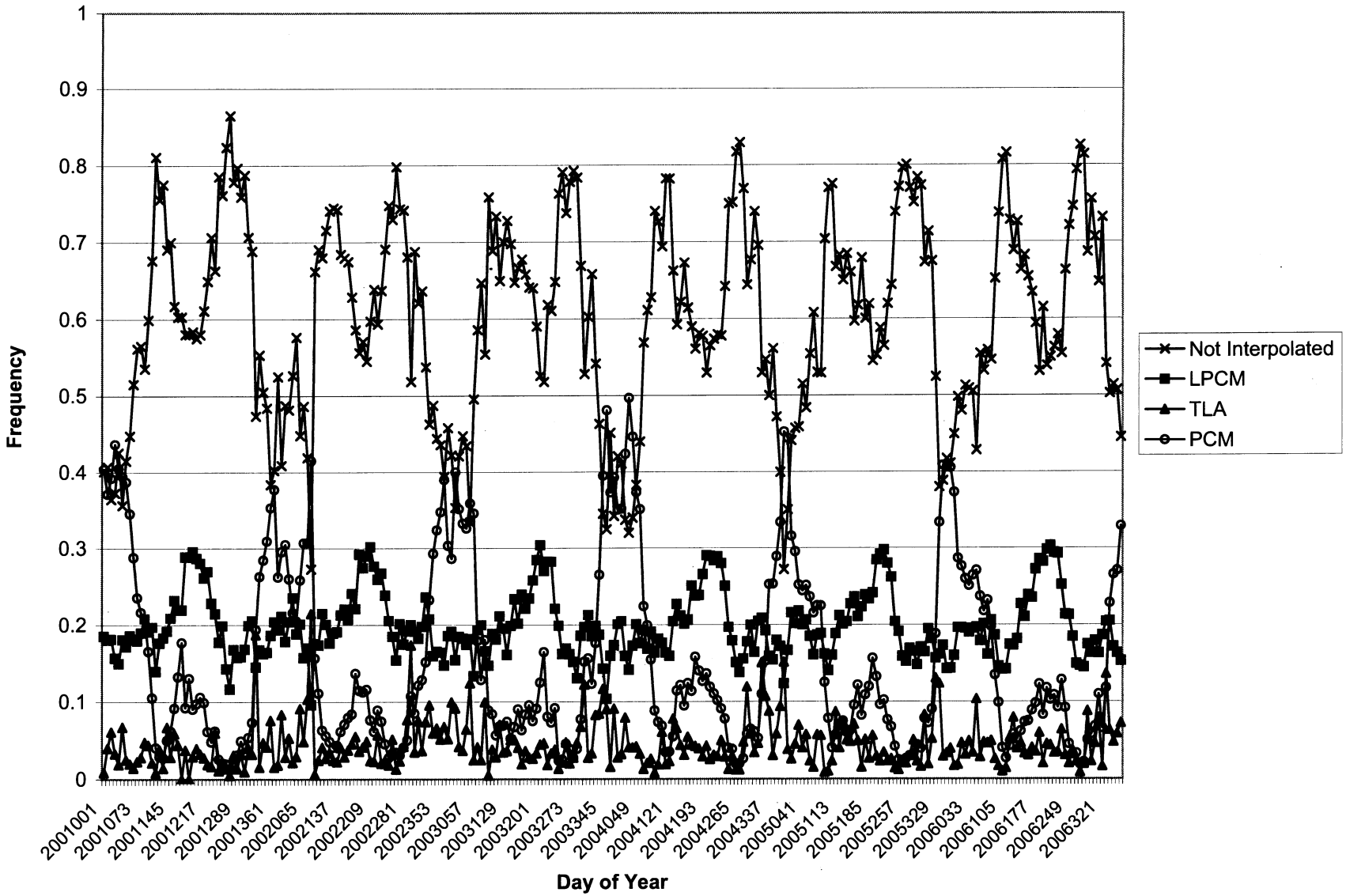

Frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2001001
2001073
2001145
2001217
2001289
2001361
2002065
2002137
2002209
2002281
2002353
2003057
2003129
2003201
2003273
2003345
2004049
2004121
2004193
2004265
2004337
2005041
2005113
2005185
2005257
2005329
2006033
2006105
2006177
2006249
2006321
Day of Year
Not Interpolated
LPCM
TLA
PCM

Figure 07

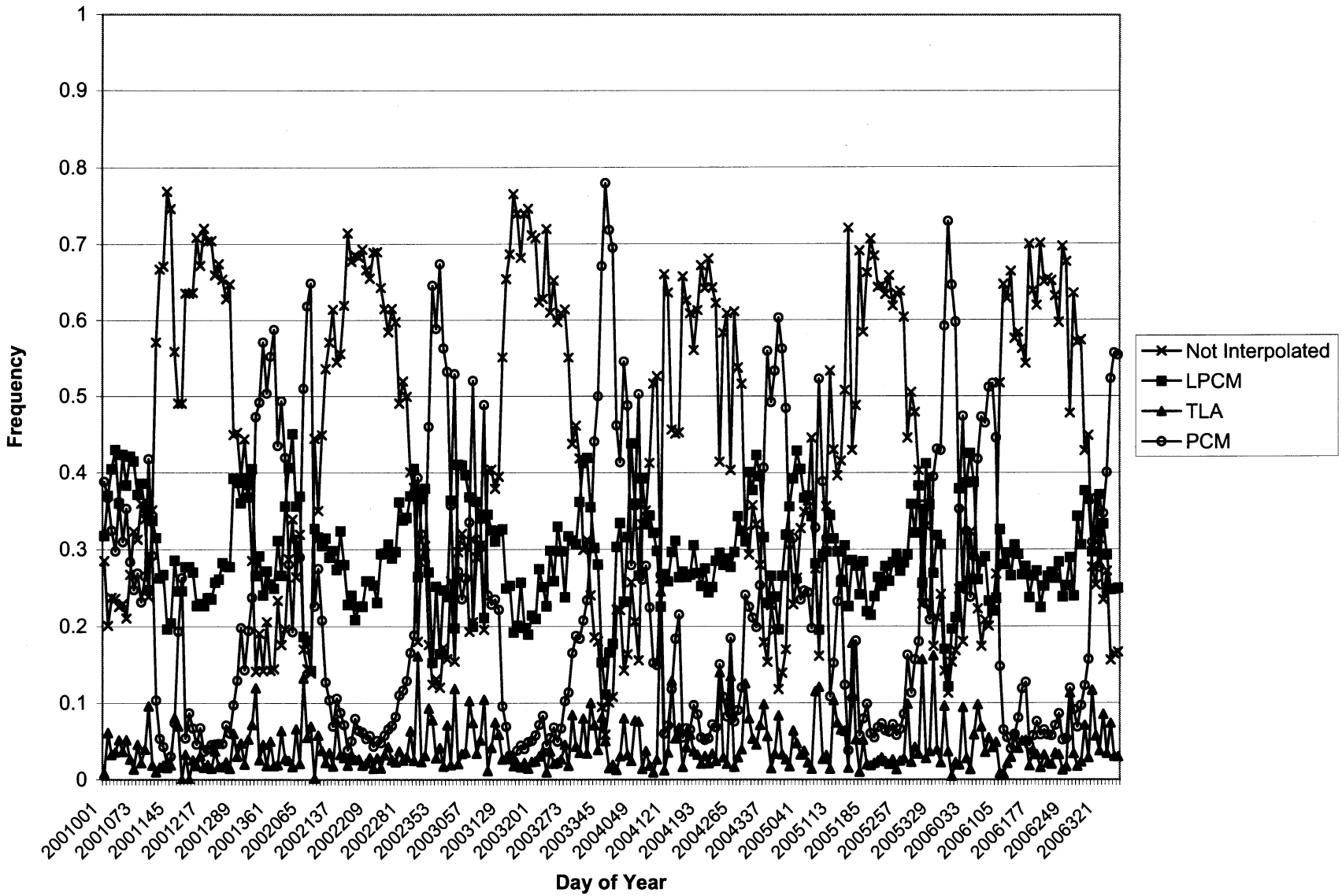
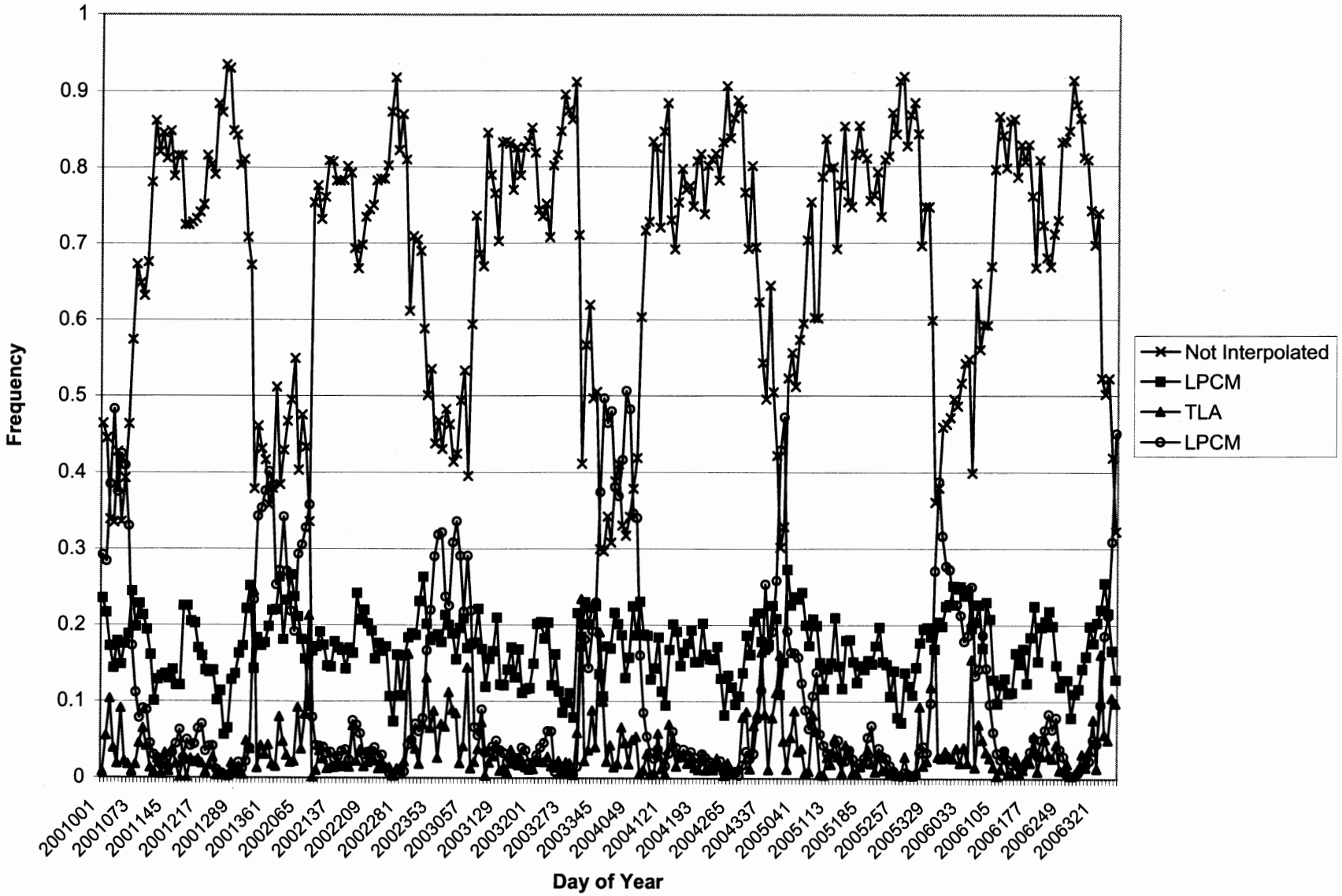

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Frequency
2001001
2001073
2001145
2001217
2001289
2001361
2002065
2002137
2002209
2002281
2002353
2003057
2003129
2003201
2003273
2003345
2004049
2004121
2004193
2004265
2004337
2005041
2005113
2005185
2005257
2005329
2006033
2006105
2006177
2006249
2006321
Day of Year
Not Interpolated
LPCM
TLA
PCM

Figure 08


Frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2001001
2001073
2001145
2001217
2001289
2001361
2002065
2002137
2002209
2002281
2002353
2003057
2003129
2003201
2003273
2003345
2004049
2004121
2004193
2004265
2004337
2005041
2005113
2005185
2005257
2005329
2006033
2006105
2006177
2006249
2006321
Day of Year
Not Interpolated
LPCM
TLA
LPCM