INT. J. REMOTE SENSING, 2000, VOL. 21, NO. 4, 839–843

Beware of per-pixel characterization of land cover

J. R. G. TOWNSHEND†*, C. HUANG, S. N. V. KALLURI,
R. S. DEFRIES, S. LIANG

2181 LeFrak Hall, Department of Geography, University of Maryland, College Park, MD 20742, USA; e-mail: jt59@umail.umd.edu

and K. YANG

Science Systems and Applications, Inc., Code 922, NASA-GSFC, Greenbelt MD 20771, USA

(*Received 21 May 1999; in final form 3 September 1999*)

Abstract. A simulation experiment was carried out to analyse the effects of the modulation transfer function on our ability to estimate the proportions of land cover within a pixel by linear mixture modelling. In the simulated landscape the proportion of each land cover type in every pixel was known exactly. The standard error of the estimate (SEE) between percentages derived from mixture modelling and the actual land cover percentages was 11%. Substantial improvements in estimating the percentages can be obtained simply by deriving estimates for pixels of twice the original dimensions, the SEE dropping to 4.16%, though this is with the obvious consequence of a final product with a coarser spatial resolution. Alternatively by deconvolving the input bands using a linear approximation of the point spread function the SEE can be reduced by almost as much, namely to 5.11%. If we combine the two approaches, by first doconvolving the bands, estimating the percentages and then aggregating resultant pixels to twice their original linear dimensions, the SEE drops to 2.24%.

1. Introduction

Attempts have been made recently to characterize land cover carried out by estimating the proportions of cover types present within each pixel using techniques such as linear mixture modelling (Gong *et al.* 1991, Asner *et al.* 1997, DeFries *et al.* 1999).

A significant, but usually ignored problem with per-pixel characterization of land cover is that a substantial proportion of the signal apparently coming from the land area represented by a pixel comes from the surrounding pixels (Townshend 1981). This is a consequence of many factors including the optics of the instrument, the detector and electronics, as well as atmospheric effects. These effects are described by the Modulation Transfer Function (MTF), which describes how the true contrast between high contrast bars is progressively reduced in the image as their width

†Also at the Institute for Advanced Computer Studies, 2131 A. V. Williams Building, University of Maryland Institute for Advanced Computer Studies (UMIACS), University of Maryland, College Park, MD 20742, USA.

ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals/tf/01431161.html

decreases (Markham 1985). Its inverse Fourier transform, the Point Spread Function (PSF) depicts how a point of light will be represented spatially and in terms of intensity on the images.

In this letter we use a simulated landscape to examine the effects of the MTF on our ability to extract land cover proportions from remotely sensed data using linear mixture modelling. We then suggest two approaches to reduce the impacts.

2. Mixture modelling

In mixture modelling the radiance (R) of a pixel is regarded as the areally weighted linear sum of the radiances (r) of the c pure cover types present occupying area a (Adams and Adams 1984). The spectral values of each of the component cover types is represented by an end-member, which is the value of a pixel if occupied solely by one cover type, e is an error term.

$$R = \sum_{i=1}^{c} r_i a_i + e \quad (1)$$

Knowing R and assuming we can estimate r_i, equation (1) can be inverted to obtain estimates of the areas a. Estimating the values of end-members can be difficult, but in our simulation the values are known exactly since we allocated the mean value to each of the pixels at the 30 m spatial resolution. A previous analysis indicated that the type of unmixing algorithm had little impact on the results (Kalluri *et al.* 1997).

3. Derivation of simulated data sets

There are major practical problems in measuring accurately land cover proportions in the field, complicated by the fact that there are always inherent uncertainties in knowing exactly where the nominal pixel boundaries occur. Moreover if we use real data, atmospheric effects will add 'noise'. Additionally the inherent spectral variability of cover types will add a further element of uncertainty. To avoid these problems, a simulated landscape was generated.

We derived the simulated data as follows. The boundaries of the land cover parcels were taken from an actual landscape in Montgomery County, Maryland, USA on the left bank of the Potomac River. Field observations were used to label each of these parcels into simple cover types. The classes used consisted of water, forest, and herbaceous cover, the latter including both grassland and crops. Near contemporaneous Thematic Mapper (TM) data were obtained and the mean values for the cover types for each of the six reflective bands were calculated. We then allocated each pixel with the mean value of its cover type and thus created an artificial landscape of pure pixels. The TM bands of the subscene were then degraded to 250 m spatial resolution therefore creating mixed pixels whose composition was known exactly. The values of the pixels were solely a result of the mixing of different proportions of 30 m pure pixels. 250 m corresponds to the finest spatial resolution of the Moderate Resolution Imaging Spectrometre (MODIS) and the Japanese Global Land Imager (GLI). To ensure a realistic spatial degradation we applied a kernel derived from the MTF of MODIS. We also prepared a set of images created solely by averaging the values. In the present analysis we used estimated values based on the instrument characteristics (Barker *et al.* 1992). Subsequently we obtained actual values from the instrument but these were little different. The resultant images do not simulate MODIS spectrally since the latter has only two bands at this spatial resolution.

4. Results

To test that our mixture modelling procedures were operating properly we unmixed the set of images produced by simple averaging. As expected the estimates corresponded exactly with the known mixtures (row 1 of table 1).

When we used the data degraded with the MODIS MTF, the SEE increased to 11.00%, representing a considerable degree of uncertainty in the estimates of the percentage of cover types. Note that in a real situation the accuracy would be lower because of atmospheric effects, spectral variability of cover types and difficulties in identifying spectral end-members.

One simple way to improve the estimates of the proportions is by aggregating the 250 m pixels to produce larger ones. This should produce more accurate estimates since the impacts of neighbouring pixels are reduced. We could approach this by conducting the mixture modelling on averaged pixel values, or as done here, where we simply derived the proportions for the larger pixels by averaging the proportions from the smaller pixels. If we aggregate 250 m pixels to form 500 m pixels, the SEE is reduced to just over 4% (row 3 of table 1). This approach has the consequence of substantially reducing the spatial resolution of the final output.

As an alternative approach we attempted to deconvolve the recorded spectral values (Forster and Best 1994) using our knowledge of the PSF for MODIS, which is the same in both scan and track directions (Barker *et al.* 1992):

$$PSF(d) = \exp\left(-\frac{d^2}{2\sigma^2}\right) \tag{2}$$

where d is the distance from pixel centre, and σ defines a detector's IFOV. In this model only the impacts of immediately neighbouring pixels are significant. Thus in one dimension the above equation can be rewritten in discrete format:

$$PSF = [\alpha \quad 1-2\alpha \quad \alpha] \tag{3}$$

where α is the proportional contribution of an immediately neighbouring pixel. In

Table 1. Standard errors of the estimate (SEE)* between the original 30 m pixels and the derived 250 m pixels (at the 90% confidence limit) for the reflective bands, namely bands 1 to 5 and 7 (B1–B5 and B7) of the Thematic Mapper. The final column shows the SEE for the estimated vs. the actual percentages of land cover.

Pixel size (m) and procedure for generating images	B1 (DN)	B2 (DN)	B3 (DN)	B4 (DN)	B5 (DN)	B7 (DN)	Land cover % error
1. 250 m, derived by simple averaging	0	0	0	0	0	0	0
2. 250 m, MTF simulation	3.19	2.99	5.08	2.97	4.58	2.13	11.00
3. 500 m, aggregated from 2	1.19	1.14	1.93	1.03	1.68	0.80	4.16
4. 250 m, MTF deconvolved	1.44	1.40	2.38	1.20	2.01	0.97	5.11
5. 500 m, aggregated from 4	0.62	0.63	1.06	0.44	0.85	0.42	2.24

*The SEE measures the dispersion of estimated values relative to the actual value.

the two dimensional space, the PSF can be written as

$$\begin{bmatrix} \alpha \\ 1-2\alpha \\ \alpha \end{bmatrix} (\alpha \quad 1-2\alpha \quad \alpha) = \begin{bmatrix} \alpha^2 & \alpha(1-2\alpha) & \alpha^2 \\ \alpha(1-2\alpha) & (1-2\alpha)^2 & \alpha(1-2\alpha) \\ \alpha^2 & \alpha(1-2\alpha) & \alpha^2 \end{bmatrix} = \{PSF_{i,j}(\alpha)\} \tag{4}$$

Let $x_{p,l}$ be the radiance from the footprint of pixel (p, l), $R_{p,l}$ the recorded radiance of that pixel, then the following relationship exists for all inner pixels of an image of M pixels and N lines:

$$R_{p,l} = \sum_{i=-1}^{1} \sum_{j=-1}^{1} PSF_{i+2,j+2}(\alpha) x_{p+i,l+j}, \text{ for } p=2,\ldots, M-1 \text{ and } l=2, \ldots, N-1 \tag{5}$$

With appropriate treatments of boundary pixels, we have $M \times N$ equations and $M \times N$ unknown variables. Thus the impact of the PSF can be deconvolved by solving this equation group. According to the PSF model, α should be 0.1464. However, the optimal performance of the deconvolution method was achieved when α equals 0.11. Explanation for this is still being sought: however this lower value gives the best deconvolution results for a wide range of landscapes in Bolivia, eastern USA, Egypt and Canada.

Using the resultant images leads to a substantial reduction in error, the SEE falling to just over 5% (row 4 of table 1) and is comparable with that obtained when aggregating the data. As a final experiment, we aggregated the estimates derived from the partly deconvolved images into 500 m pixels. This resulted in an SEE of just over 2%.

5. Conclusions

The results of our simulation experiment indicate that extracting useful information from individual pixels can be substantively inhibited by the contribution of signals from surrounding pixels, the MTF effect. Much more work is needed to establish the size of these effects for different landscapes and sensor types. The results indicate the difficulties of attempting to characterize land cover based solely on the spectral response of individual pixels. This implies that land cover properties should be reported at spatial resolutions coarser than the individual pixel or that the signal from individual pixels should be deconvolved. An alternative is to use contextual procedures in which observations from surrounding pixels are used to assist the characterization. These conclusions are probably relevant not only to the estimate of proportions through mixture modelling but also to other procedures such as conventional classification because the spectral signatures of the pixels are affected strongly by the surrounding cover types. Only in those situations where pixel size is small relative to the area of land cover units will these effects be unimportant.

Acknowledgments

The authors acknowledge support from NSF grant BIR9318183 and a contract from the National Aeronautics and Space Administration (NAS596060). We are grateful to the MODIS Science Data Support Team (SDST) for providing the PSF values.

References

Adams, J. B., and Adams, J. D., 1984, Geologic mapping using Landsat MSS and TM images: Removing vegetation by modeling spectral mixtures. *Proceedings of the International Symposium on Remote Sensing of Environment, Third Thematic Conference on Remote Sensing for Exploration Geology*, Colorado Springs, Colorado (Ann Arbor: Environmental Research Institute of Michigan), pp. 615–622.

Asner, G. P., Wessman, C. A., and Privette, J. L., 1997, Unmixing the directional reflectances of AVHRR sub-pixel landcovers. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, 868–878.

Barker, J. L., Markham, B. L., and Burelbach, J. W., 1992, MODIS image simulation from Landsat TM imagery. *Global Change and Education, ASPRS/ACSM/RT 92*, Washington, DC (Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing), pp. 156–165.

DeFries, R. S., Townshend, J. R. G., and Hansen, M., 1999, Continuous fields of vegetation characteristics at the global scale. *Journal of Geophysical Research*, **104**, 16911–16923.

Forster, B. C., and Best, P. P., 1994, Estimation of SPOT P-mode point spread function and derivation of a deconvolution filter. *ISPRS Journal of Photogrammetry and Remote Sensing*, **49**, 32–42.

Gong, P., Miller, J. R., Freemantle, J., and Chen, B., 1991, Spectral decomposition of Landsat Thematic Mapper data for urban land cover mapping. *Proceedings of the 14th Canadian Symposium on Remote Sensing*, Calgary, Alberta (Ottawa: Canadian Remote Sensing Society), pp. 458–461.

Kalluri, S. N. V., Huang, C., Mathieu-Marni, S., Townshend, J. R. G., Yang, K., and Chellapa, R., 1997, A comparison of mixture modeling algorithms and their applicability to the MODIS data. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Singapore (Piscataway, New Jersey: IEEE), pp. 171–173.

Markham, B. L., 1985, The Landsat sensors' spatial responses. *IEEE Geoscience and Remote Sensing*, **23**, 864–875.

Townshend, J. R. G., 1981, Spatial resolution of satellite images. *Progress in Physical Geography*, **5**, 33–55.