International Journal of Remote Sensing, 2017, Volume 28, Issue 13, pp3785-3799
DOI:10.1080/01431161.2017.1303219

RESEARCH ARTICLE

*International Journal of Remote Sensing*

# Improved estimates of opium cultivation in Afghanistan using imagery based stratification

Daniel M. Simms, Toby W. Waine*, John C. Taylor

*Cranfield University, Bedfordshire, MK43 0AL, UK*

**ARTICLE HISTORY**
Compiled, December 2016

**ABSTRACT**
The United Nations Office on Drugs and Crime and the US Government make extensive use of remote sensing to quantify and monitor trends in Afghanistans illicit opium production. Cultivation figures from their independent annual surveys can vary because of systematic differences in survey methodologies relating to spectral stratification and the addition of a pixel buffer to the agricultural area. We investigated the effect of stratification and buffering on area estimates of opium poppy using SPOT5 imagery covering the main opium cultivation area of Helmand province and sample data of poppy fields interpreted from very high resolution satellite imagery. The effect of resolution was investigated by resampling the original 10 m pixels to 20, 30 and 60 m, representing the range of available imagery. The number of strata (1, 4, 8, 13, 23, 40) and sample fraction (0.2 to 2%) used in the estimate were also investigated. Stratification reduced the confidence interval by improving the precision of estimates. Cultivation estimates of poppy using 40 spectral strata and a sample fraction of 1.1% had a similar precision to direct expansion estimates using a 2% sample fraction. Stratified estimates were more robust to changes in sample size and distribution. The mapping of the agricultural area had a significant effect on poppy cultivation estimates in Afghanistan, where the area of total agricultural production can vary significantly between years. The findings of this research explain differences in cultivation figures of the opium monitoring programmes in Afghanistan and recommendations can be applied to improve resource monitoring in other geographic areas.

**KEYWORDS**
Opium poppy; stratification; agricultural mask, cultivation estimates; buffering

## 1. Introduction

Afghanistan is the source of 70% of the World's opium and almost all of the heroin in the UK (UNODC 2016). The two main organisations monitoring the cultivation of illicit opium are the United Nations Office on Drugs and Crime/Afghanistan's Ministry of Counter Narcotics (UNODC), and the US Government. For both programmes the annual cultivated area is obtained by measuring poppy at random sample sites from image-interpretation of very high resolution (VHR) satellite imagery. The two surveys are independent and use different sources of VHR imagery, sample fractions,

---

*Corresponding author. Email: t.w.waine@cranfield.ac.uk

sample sizes and methodology for expanding the sample observations to provincial and national estimates.

Remote sensing can be used to define sample units and sampling frames, optimise the allocation and size of sampling units and improve sample estimates by regression or calibration (Carfagna and Gallego 2005). In Afghanistan, the UNODC and US annual opium surveys use remote sensing as there is limited capacity to collect accurate field data because of poor security after decades of war. Imagery is used to map the area of agricultural production and for visual interpretation of crops at sample locations.

Agricultural production in Afghanistan takes place in blocks of irrigated land close to the main rivers, ribbons running along upland river valleys, and extensive rain fed areas. These areas have strong contrast with large mountainous areas and desert. In both surveys the agricultural area, known as the agricultural mask, is determined by digital classification and visual interpretation of medium resolution imagery, such as SPOT (10–20 m), DMC (32 m) and AWiFs (56 m), to exclude areas of non-agriculture from the sample frame. Images must have a wide enough swath and short enough revisit time to allow for suitably timed cloud-free collections covering the poppy producing provinces. Differences in the agricultural mask between surveys directly affect cultivation estimates as the agricultural area is a multiplier in the statistical expansion of the sample proportion of poppy to the provincial scale.

A major difference between the US and UNODC methodologies is the use of medium resolution imagery to stratify the agricultural area within the mask based on spectral response. A stratified estimate is designed to minimise the within-stratum variance compared to the variance between strata to improve the accuracy of a ratio estimate (Cochran 1977). In Afghanistan, Luders, Wilson, and Gardener (2004) tested a stratified approach using spectral classes from Landsat 7 imagery and found that stratification lowered the variance of the overall estimate. They also suggest buffering the mask to ensure sparse fields and the edges of agricultural areas are not excluded from the analysis because of the resolution of the imagery. Their work forms the basis of the current US survey methodology.

A degree of homogeneity in poppy cultivation within the strata, referred to in statistical terms as weak or second order spatial stationarity, is required for any increase in precision (Koeln and Kollasch 2000). At the scale of agricultural production in Afghanistan, where individual fields are typically $<0.5$ ha, the spectral signal from medium resolution imagery ($>20$ m) is a mixture of crop types, bare soil and infrastructure such as roads and compounds. Variation in crop phenology across image scenes will also affect the spectral response. The effect of stratification on area estimates of poppy was identified as a potential source of discrepancy between UNODC and US cultivation figures.

In 2003 the UK Government commissioned research into explaining conflicting information from the cultivation figures produced by the UNODC and US Government in an attempt to harmonise the results of Afghanistan's opium surveys (Taylor et al. 2010). As part of trials conducted alongside the main surveys, we produced stratified and unstratified estimates of poppy using image data from the Disaster Monitoring Constellation (DMC). 30 strata were chosen as a compromise between maximising the number of separable classes in the classification and avoiding the creation of under-sampled strata. Table 1 shows the estimates for selected provinces between 2006 and 2009. There are differences in the most probable estimates between stratified and unstratified estimates. In Helmand and Uruzgan the stratification reduced the estimate in every year (by 12,204 ha in Helmand 2007). In the other provinces stratification increased or decreased the estimate in different years. Stratified estimates have improved

**Table 1.** Opium poppy cultivation estimates and 90% confidence intervals (%) 2006 to 2009 for selected provinces. Province boundaries do not match those from other published sources and figures should not be directly compared.

| | | Unstratified | | | Stratified (30 strata) | | |
|---|---|---|---|---|---|---|---|
| Province | Year | Poppy, ha | -CI90 | +CI90 | Poppy, ha | -CI90 | +CI90 |
| Helmand | 2006 | 73 930 | 12 | 12 | 61 945 | 8 | 8 |
| | 2007 | 100 894 | 9 | 9 | 88 690 | 6 | 6 |
| | 2008 | 97 348 | 7 | 7 | 81 256 | 5 | 6 |
| | 2009 | 78 361 | 8 | 8 | 76 383 | 7 | 6 |
| Kandahar | 2006 | 7 777 | 33 | 46 | 9 244 | 26 | 30 |
| | 2007 | 12 140 | 31 | 38 | 13 634 | 28 | 32 |
| | 2008 | 16 851 | 25 | 26 | 16 658 | 17 | 14 |
| | 2009 | 20 372 | 23 | 24 | 19 198 | 18 | 17 |
| Uruzgan | 2006 | - | - | - | - | - | - |
| | 2007 | 10 659 | 32 | 31 | 8 487 | 23 | 21 |
| | 2008 | 10 803 | 26 | 26 | 7 094 | 16 | 16 |
| | 2008 | 8 304 | 21 | 22 | 6 841 | 17 | 19 |
| Nangarhar | 2006 | 4 067 | 62 | 76 | 5 182 | 63 | 72 |
| | 2007 | 20 349 | 25 | 25 | 19 222 | 22 | 20 |
| | 2008 | <500 | - | - | <500 | - | - |
| | 2009 | <500 | - | - | <500 | - | - |
| Badakhshan | 2006 | 3 149 | 62 | 70 | 3 337 | 67 | 69 |
| | 2007 | 2 303 | 36 | 39 | 2 322 | 37 | 37 |
| | 2008 | 659 | 42 | 46 | 561 | 41 | 44 |
| | 2009 | - | - | - | - | - | - |

confidence intervals for all provinces in all years.

These cultivation figures shown how unexplained differences in estimates can undermine confidence in opium estimates and impede policy formation on counter narcotics. This paper presents work prompted by technical discussions with the UNODC and US Government on the use of spectral stratification. It investigates the effect of buffering the agricultural area, resolution of the imagery used for stratification, the number of strata, and sample fraction on cultivation estimates of opium poppy.

## 2. Stratification, resolution and buffering

The effect of stratification, image resolution and buffering on estimates was investigated using SPOT5 10 m resolution multispectral imagery, targeted for a study area in the main production area of Helmand province in 2007 (figure 1). Image acquisition was timed to coincide with the cultivation of crops using an information system based on time series NDVI from the Moderate Resolution Spectroradiometer (Simms et al. 2014). The SPOT imagery was orthorectified using a controlled image base (CIB) and a 30 m digital elevation model (DEM) for control, to achieve sub-pixel geometric accuracy. The image was then resampled using bilinear interpolation to 20, 30 and 60 m pixels, representing a selection of suitable medium resolution imagery sources.

The stratified area frame sampling methodology can be split into five steps: (1) agricultural mask creation; (2) stratification; (3) sample selection; (4) image interpretation of samples; and (5) final estimate and confidence interval.

An agricultural mask was created for each resolution by classification using the Iterative Self-Organising Data Analysis Technique (ISODATA). The resulting classified pixels were grouped into agriculture and non-agriculture information classes by

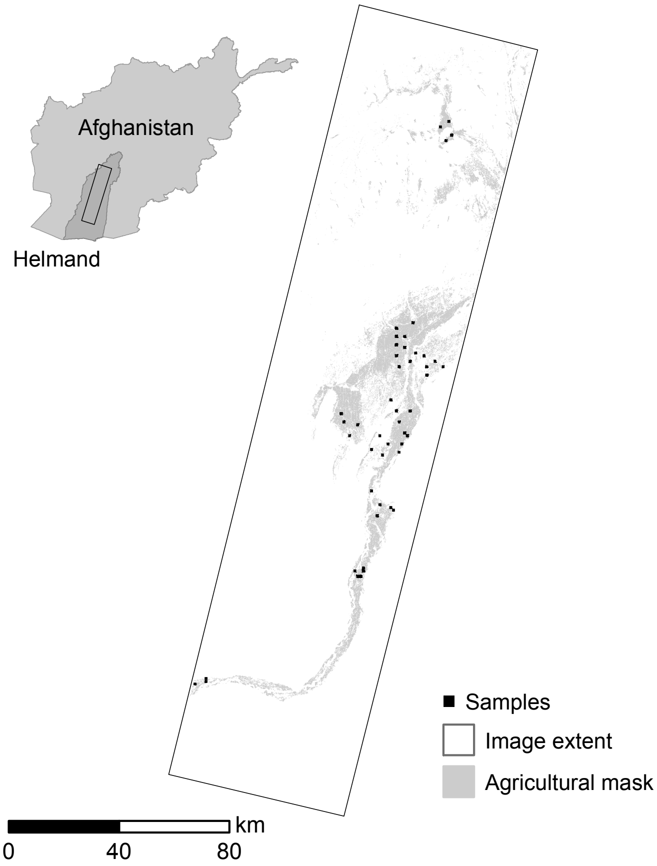


**Figure 1.** Map showing the extent of the SPOT5 image acquired on 2 April 2007 and 48 sample locations in Helmand Province, Afghanistan.

visual image-interpretation. The agricultural area at each resolution was buffered by one pixel and stratified by clipping the resampled multispectral imagery to the mask before clustering into 30 spectral vegetation classes using ISODATA, resulting in three buffered strata layers. The buffer pixels of each strata were then removed to create an additional three unbuffered strata layers.

The sample was designed within the existing UNODC sample selection to allow direct comparison and make use of their ground survey data. A sample size of 1 km × 1 km and sample fraction of 2% of the agricultural area was chosen based on experience of monitoring crop inventories in Europe assisted by remote sensing (Taylor et al. 1997), resulting in 48 samples (figure 1).

A VHR image (IKONOS or Quickbird–2) was targeted to coincide with the flowering period of opium poppy for each block using the same crop information system used to target the medium resolution acquisitions. The flowering period lasts for approximately two weeks and is the optimum time in the crop cycle for discrimination of opium poppy from other crops (Jia et al. 2011). Each acquired VHR image was pan-sharpened to 1 m resolution and ortho-resampled using the vendor supplied Rational Polynomial Camera model refined using control points from the CIB and a 30 m DEM (Grodecki and Dial 2001).

The cultivated area of opium poppy and wheat was delineated by visual image-interpretation at each 1 × 1 km sample location. Interpretation keys were developed for opium poppy and the main first cycle crops using ground observations and photography from the UNODC's segment surveys. Consistency in interpretation was maintained by cross-checking a 5% overlap between individual interpreters.

The area of the poppy within each stratum is calculated from $n$ number of samples by

$$m_s = \frac{\sum\limits_{i=1}^{n} m_i}{\sum\limits_{i=1}^{n} a_i} A_s, \tag{1}$$

where $m_i$ is the area of poppy within stratum $s$ in sample $i$, $a_i$ is the total area of sample $i$ in stratum $s$ and $A_s$ is the total area of the stratum in the study area. The total area estimate for poppy is the combined estimates for $x$ number of strata,

$$M = \sum_{s=1}^{x} m_s. \tag{2}$$

The stratified estimate can be directly compared with an unstratified direct expansion by combining the individual strata into a single stratum representing the total agricultural area. Equation 2 becomes

$$M = \frac{\sum\limits_{i=1}^{n} m_i}{\sum\limits_{i=1}^{n} a_i} A, \tag{3}$$

where $A$ is the total agricultural area.

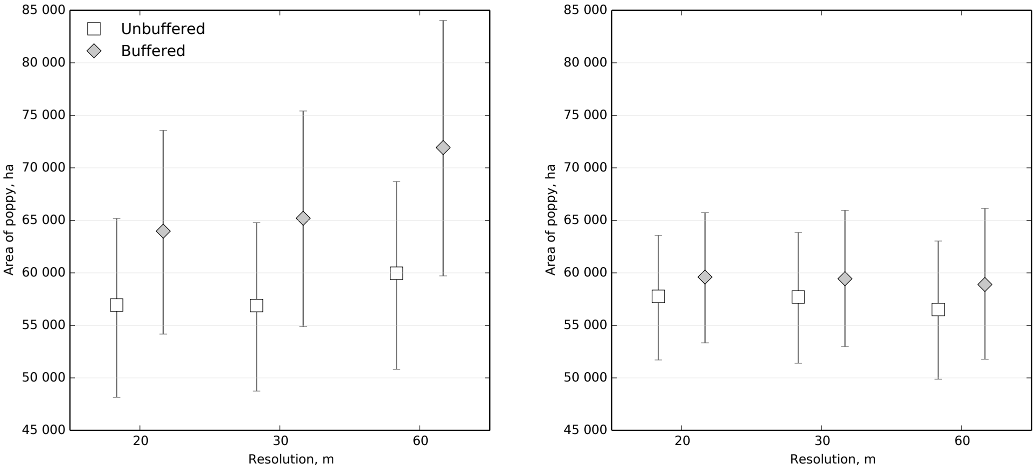
The confidence interval of the estimate is calculated by bootstrapping to approximate the distribution of $M$ from $N$ repetitions of equation 1 using a random draw of the sample,

$$M^* = M(X_1^*, \ldots, X_n^*) \tag{4}$$

where $X_1^*, \ldots, X_n^*$ is an independent random selection of the original samples with replacement. This results in $N$ calculations of $M^*$. The upper and lower confidence intervals are found by ordering the values of $M^*$ and taking the value corresponding to the percentile required. For example, $M_{500}^*$ and $M_{9500}^*$ for the 90% confidence level for $N = 10,000$.

The total area of poppy was calculated for each mask resolution using buffered and unbuffered strata (equation 2) and a single stratum (equation 3) using the same 48 image interpreted samples.

Buffering the agricultural mask increased the stratified (30 strata) and single stratum estimates of poppy at all three spatial resolutions tested (figure 2). Table 2 shows the difference in the single stratum and stratified area estimates, the total masked area and the proportion of measured poppy in the sample after buffering. The agri-



(a) Single stratum

(b) Stratified

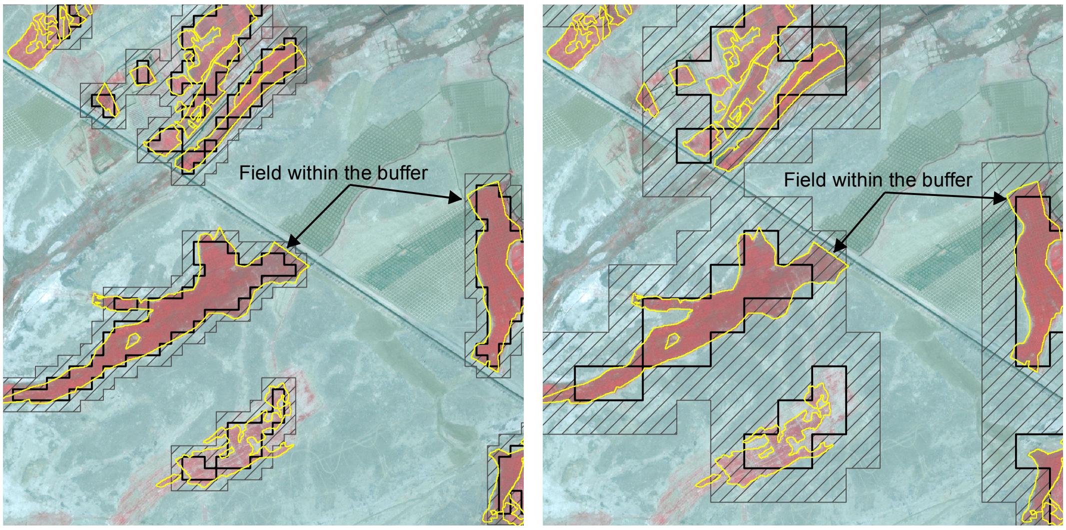
**Figure 2.** (a) Single stratum and (b) stratified area estimates for buffered and un-buffered agricultural masks created from 20, 30 and 60 m resolution imagery using 2% sample. Error bars represent 90% CI range.

cultural area increases by 36, 39 and 48% after buffering for resolutions of 20, 30 and 60 m respectively, with a corresponding decrease in the proportion of poppy within the sample of about 7%. Buffering the agricultural area includes more of the measured area of poppy within the sample for all resolutions (1.4 to 2.2%). Stratification greatly reduced the positive bias caused by buffering (figure 2), with the differences in buffered estimates in table 2 within about 4% of unbuffered estimates for the three resolutions tested, compared to differences of 12, 15 and 20% for the single stratum estimates.

**Table 2.** Effect of buffering at 20, 30 and 60 m resolution on total area of the agricultural mask, proportion of poppy, measured area of poppy within the sample, single stratum and stratified poppy area estimates.

| Resolution, m | Difference in total mask area, % | Difference in proportion of poppy, % | Difference in sampled poppy, % | Difference est. single, % | Difference est. stratified, % |
|---|---|---|---|---|---|
| 20 | 36.0 | -7.1 | 2.2 | 12.4 | 3.2 |
| 30 | 38.7 | -6.7 | 1.4 | 14.6 | 3.0 |
| 60 | 47.6 | -6.8 | 1.9 | 20.0 | 4.2 |

In general the agricultural mask was consistent with the extent of poppy cultivation in the samples with no noticeable alignment error between the orthorectified imagery used for stratification and image-interpretation. Figure 3 shows the effect of buffering on the measured area of poppy at the sample scale. At a resolution of 20 m, some of the measured field area falls outside of the unbuffered mask because the field edge digitised from VHR imagery is more detailed than the mask boundary. At 60 m (figure 3(b)), the mask boundary is even more generalised, with a greater mismatch between the mask boundary and the field edges caused by the increased pixel size of the medium resolution imagery. Buffering increases the measured area of poppy included within the sample for all mask resolutions. The buffered mask at 20 m resolution includes more of the measured area of poppy than at resolution of 30 m and 60 m across all the samples despite the increasing total area of the mask (table 2).



(a) (b)

**Figure 3.** Agricultural field boundaries in yellow, digitised from VHR imagery, overlaid with (thick line) unbuffered and (hatched area) buffered agricultural masks at resolutions of (a) 20 m and (b) 60 m. Background image false colour pan-sharpened IKONOS, acquired 17 April 2007.

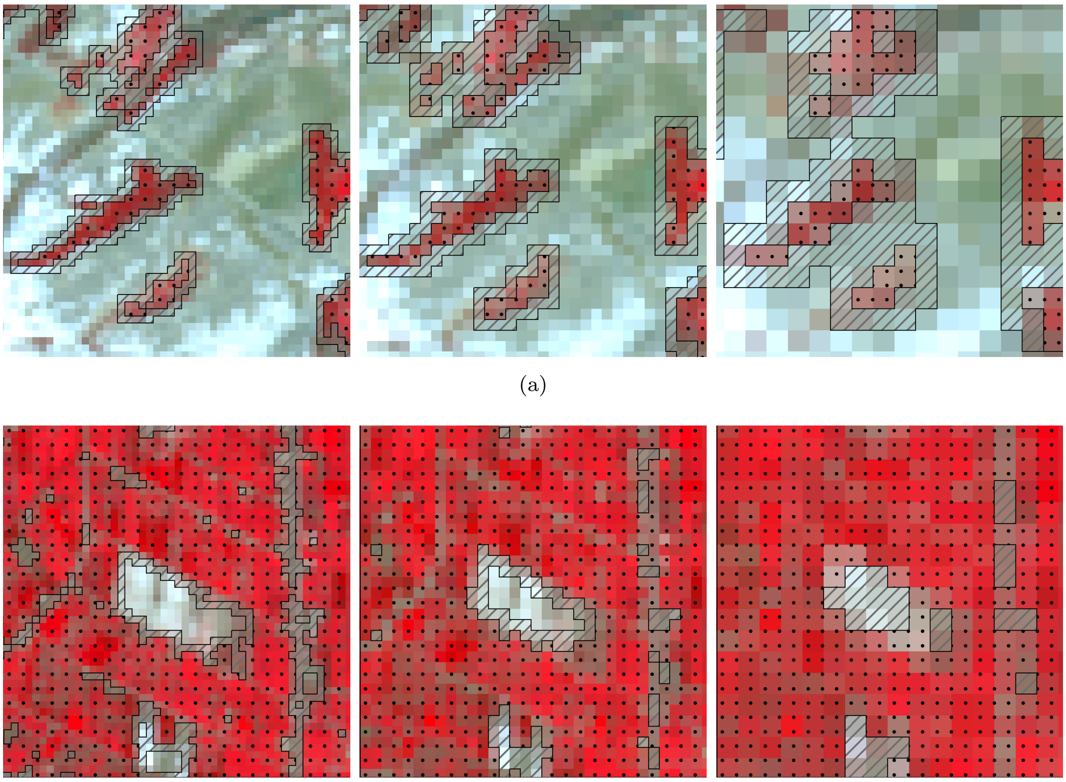
The buffering of the masked area lowers the proportion of poppy in samples as the masked area increases: reducing the average proportion of poppy. However, this does not compensate for the large increase in the masked area as the effect is limited to samples at the interface between agriculture and non-agriculture. For example, the sample in figure 4(a) is located at the mask boundary in marginal agriculture, where the masked area makes up a greater proportion of the total area of the sample compared to the location in figure 4(b), which is within the main irrigated valley. This leads to the positive bias in the area estimate seen in figure 2.

The measured area of poppy missing from the mask was caused by small areas of poor quality crops with a weak spectral response and the detail of the agricultural boundary caused by the resolution of the mask. These effects were more evident at 60 m resolution, where the proportion of active vegetation within the pixel decreases in relation to the non-vegetated surface and there is greater mismatch between the detail of the boundary mapped from VHR and medium resolution imagery.

The next section investigates the effect of strata number and sample fraction on estimates and the relationship between the spectral classes and poppy cultivation.

## 3. Number of strata and sample fraction

The effect of the number of spectral classes used in the stratified estimates was investigated using the resampled 20 m multi-spectral SPOT5 image, subset to the agricultural mask. An increasing number of strata were derived hierarchically from unsupervised clustering of the image pixels to allow direct comparison between the resulting poppy estimates. Firstly, a layer with 4 strata was created from the subset image using ISODATA. Each of these strata were then split into two by re-clustering the original image pixels within each stratum to produce an 8 strata layer. At this point the 8 strata were assessed for homogeneity using the relative standard error of the bootstrapped estimate,



(a)

(b)

**Figure 4.** Examples of two 1 x 1 km samples in (a) marginal agriculture and (b) in the main irrigated area of Helmand Province comparing the buffered (hatched) and unbuffered (dotted) agricultural mask for image resolutions of 20, 30 and 60 m (left to right) resampled from SPOT5 image, acquired 2 April 2007.

$$RSE_s = \frac{\sigma(m_s^*)}{\bar{m}_s^*} \tag{5}$$

where $m_s^* = m_s(X_1^*, \ldots, X_n^*)$ for $N = 10,000$ draws of the sample. Any stratum with a $RSE_s > 0.1$ was subdivided into two new strata, which resulted in a 13 strata layer. This splitting process was repeated twice more, resulting in layers with 23 and 40 strata. Splitting was stopped at 40 to avoid under-sampling those strata with small areas.

Estimates for the stratified layers (1–40 strata) using proportions of the original sample data at intervals of 0.2% were then calculated. Sample proportions were simulated using a random draw of the original sample squares for each iteration of the bootstrap $(X_1^*, \ldots, X_k^*)$, where $k$ is the number of samples for the required sample proportion. The sample cultivation estimate was calculated from the mean of the bootstrapped estimates $\bar{M}_p^*$ and the precision of each poppy estimate as 1 - relative standard error,

$$\text{precision}_p = 1 - \frac{\sigma(M_p^*)}{\bar{M}_p^*}. \tag{6}$$

Figure 5 shows the improvement in the precision of the area estimate of poppy as the sample fraction and the number of strata increase. The area estimate using 4 strata and a sample fraction of 1.6% has a similar precision to the original single stratum estimate with a 2% sample fraction. Increasing the number of strata to 8 lowers the required sample fraction to 1.4% for a single stratum estimate of similar precision. For estimates with more than 8 strata the gains in precision are less. A 40 strata estimate is close to half the sample fraction (1.1%) required for a direct expansion estimate of the same precision (2% sample).

The effect of the individual strata on the overall estimate can be seen in figure 6. Splitting the single stratum (agricultural mask) into 4, separated three strata with higher proportions of poppy ($> 0.25$) and lower $RSE_s$ ($< 0.11$) from a stratum with a lower proportion of poppy (0.14) and a higher $RSE_s$ (0.23). The higher $RSE_s$ showing greater variation in the stratum estimate of poppy from the sample. Separating this more mixed stratum decreases the overall $RSE$ from 0.080 to 0.067, improving the precision of the estimate.

Further splitting of the strata results in an increase in precision for some of the strata. The strata in the top half of the dendrogram (s01–s31 in figure 6) are generally more mixed with a lower proportion of poppy, where splitting does not result in a reduction in $RSE_s$, except in s21. In the lower half, the strata with a higher proportion of poppy have low $RSE_s$ and, being identified as homogeneous, are not split further (s31–s34). These results show the stratification improving the estimate by separating high production areas with lower variance in poppy cultivation from lower production areas with higher variance, reducing the overall $RSE$. This lowers the overall variance compared to the sample as a whole and reduces the size of the confidence interval in line with the findings by Luders, Wilson, and Gardener (2004).

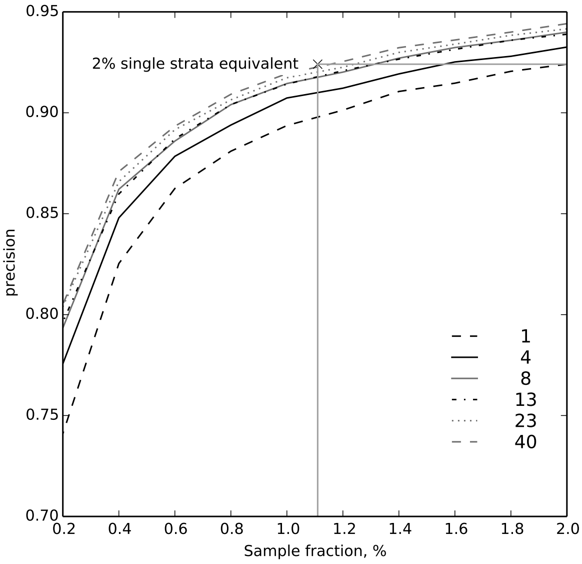


**Figure 5.** Variation in overall precision (1 - relative standard error) of bootstrapped area estimates of poppy with variation in sample fraction and the number of strata (20 m SPOT5). The 40 strata sample fraction with precision equivalent to 2% single strata is marked.

**Table 3.** Area of agricultural mask (hectares) between 2006 and 2008 in poppy producing provinces of Afghanistan. Percentage difference with previous year in brackets.

| | 2006 | 2007 | 2008 |
|---|---|---|---|
| Helmand | 277 416 | 280 228 (1) | 287 859 (3) |
| Uruzgan | 61 311 | 58 764 (-4) | 62 307 (6) |
| Nimroz | 32 681 | 35 007 (7) | 26 813 (-23) |
| Kandahar | 119 728 | 186 870 (56) | 182 120 (-3) |
| Nangarhar | 85 396 | 102 565 (20) | 101 520 (-1) |
| Badakhshan | 168 842 | 123 626 (-27) | - |
| Farah | 94 673 | 95 840 (1) | 104 520 (9) |
| Balkh | - | 310 425 (-) | 150 234 (-52) |

## 4. Discussion and conclusions

There are significant annual changes in the agricultural land under cultivation in Afghanistan caused by variation in water availability from snow melt and rainfall (Shahriar Pervez, Budde, and Rowland 2014), crop rotations, and agricultural expansion. The changes in the active agricultural mask for selected provinces from the wider project work from 2006 to 2008 are shown in table 3. The total change can be misleading as some areas are in rotation and there are shifts in cultivation of poppy crops from the main irrigated valley into newly exploited desert areas (irrigated from wells). Figure 7 shows a example from northern Helmand where the active areas is larger in the higher valleys in 2007 (purple) compared with 2008, where there is more active area in the valley (yellow). The green area is active agriculture that was consistent over the two years. The study area is located in Helmand Province, which has the lowest annual total change. Kandahar and Balkh have differences of up to 50%, showing the importance of understanding the effect of the agricultural mask on opium estimates.

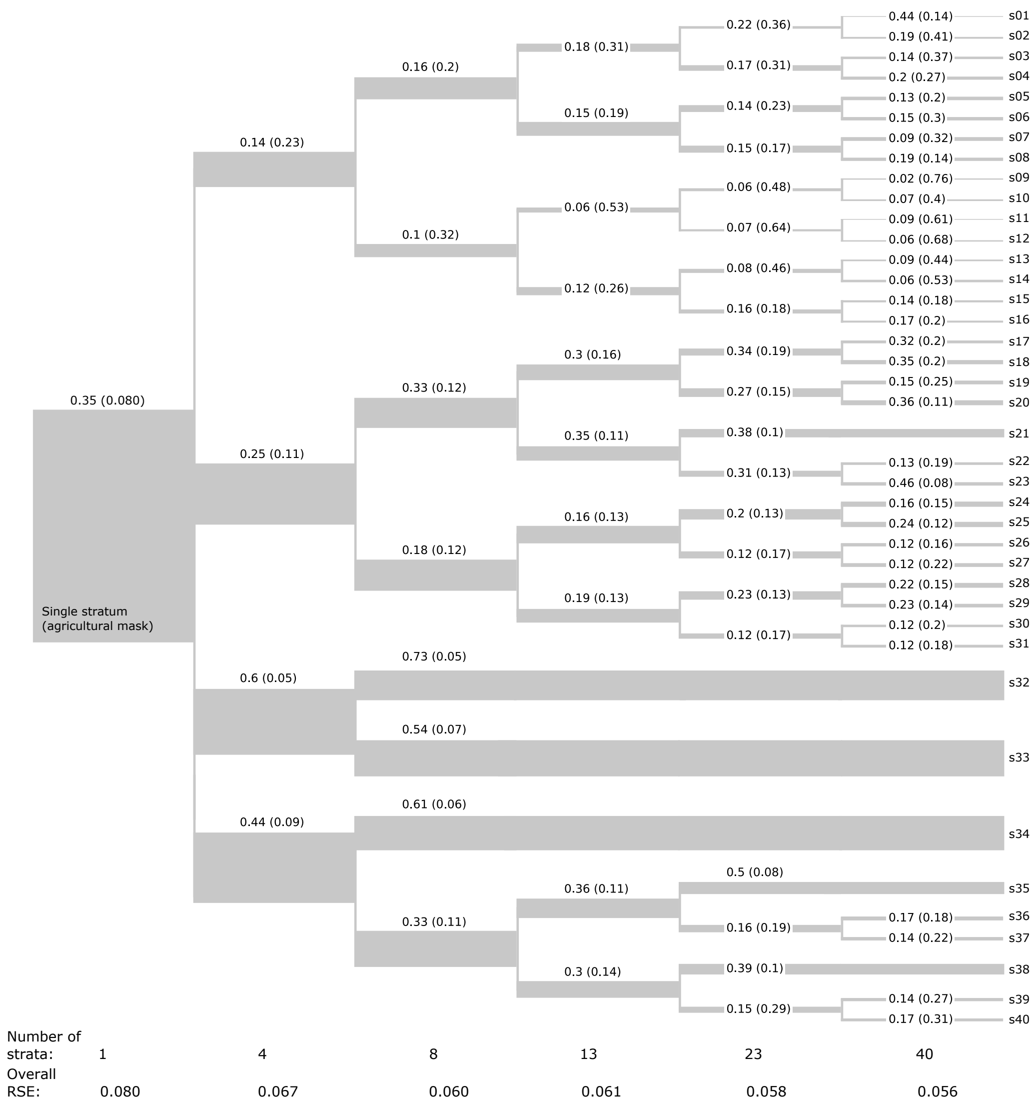


**Figure 6.** Dendrogram showing the proportion of poppy and relative standard error (in brackets) after successive splitting of heterogeneous SPOT5 20 m strata within the agricultural mask using ISODATA. The thickness of the line is proportional to the area of the stratum. Stratum labels prefixed with ‘s’.

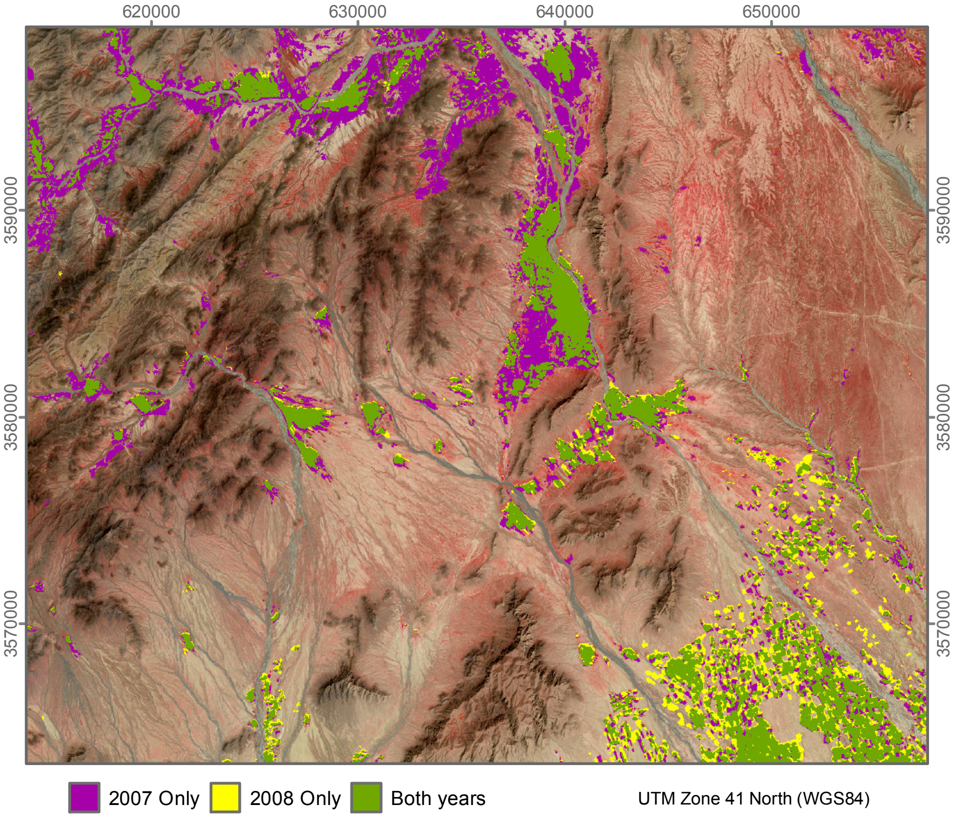


**Figure 7.** Change in active agricultural mask between 2007 and 2008, northern Helmand Province, Afghanistan. Background image SPOT5 false-colour composite, aquired 2 April 2007.

The purpose of the agricultural mask in opium surveys is to improve efficiency by reducing the total area so resources are not wasted sampling where poppy cultivation does not take place. There are two main approaches to creating the mask: the area of active agriculture; and the area of potential agriculture. The active mask is created each year and only contains those areas that are currently in production. The separation of agricultural areas from the desert, natural vegetation and high mountains is achieved by unsupervised classification and visual interpretation of medium resolution imagery such as Landsat and DMC. While time consuming, this process is straightforward as there is high spectral contrast between the active vegetation and the desert when imagery is correctly timed to coincide with the presence of crops.

In contrast, the potential mask is updated yearly with areas of new agriculture, retaining the existing mapping from previous years. It is much faster to produce once a baseline has been established as only newly exploited areas require imaging and further analysis, opposed to total coverage over the target provinces for each year.

Amending the mask to include all potential agriculture would include areas out of rotation and large areas of rain fed agriculture, which are unable to support cultivation in years with insufficient rainfall. These changes take place in both high production and marginal areas and contribute to the yearly change in the distribution of agricultural land, as shown in Simms et al. (2014). The buffering results clearly show that using a potential agricultural mask for direct expansion of the sample proportion, without a significant increase in the amount of sampling, will artificially inflate the area estimate of poppy.

The decision to produce an active or potential mask is influenced by the availability of suitably timed imagery and the resources within the monitoring programme. In general terms, imagery with larger coverage improves the chances of cloud free collections during the crop growth cycle at the expense of spatial resolution. This work shows that moving to coarser image resolutions will lead to an overestimation of the agricultural area where the boundary is complex. The source of medium resolution imagery in relation to the scale of the agriculture being sampled is important to avoid over estimation when using direct expansion.

Adding a pixel buffer to the agricultural area is a strategy to capture poppy cultivation that might be missed because of the timing or resolution of medium resolution imagery, or misalignment with the sample data. Buffering the mask increased the area of sampled poppy included in the analysis. However, this increased the area of the mask disproportionally across the samples relative to the detail in the mapped boundary between agriculture and non-agriculture. The effect was an artificial inflation of the single stratum estimate (12 to 20%). A rigorous approach to image orthorectification to ensure sample data is coincident with the agricultural area is preferable to buffering the mask for direct expansion estimates. Precise timing and appropriate resolution of medium resolution imagery will also ensure that sampled fields fall within the current active agriculture.

The stratified estimate was more robust to changes in the agricultural mask from buffering and image resolution (between 3 and 4%) compared with a single stratum direct expansion of the sample proportion (between 12 and 20%). This was despite the heterogeneity of individual strata caused by the mixed spectral response of small fields at the resolution of the imagery. The greatest effect of stratification was lowering the sample proportion for large areas of low poppy cultivation within the agricultural mask. These results are consistent with findings by McRoberts et al. (2002) that mixed strata will improve estimates where they are limited to small areas, such as boundary cases, if large areas have lower variance, with either high or low proportions of the land cover of interest.

The findings are significant for Afghanistan's opium monitoring programmes because of the large inter-annual fluctuations in the agricultural area, where the use of potential agriculture masks could lead to differences in survey estimates between the UNODC and US Government. Stratified estimates will minimise discrepancy between independent survey teams related to differences in their agricultural mask and sample design. They also reduce variance, meaning estimates can be made using a smaller sample (1.1% compared to 2% direct expansion), lowering the survey costs relating to the purchase of VHR imagery and image interpretation. Stratification does not add significantly to the survey cost as medium resolution imagery is already collected for defining the mask.

Stratification also reduced the effect of the decreased resolution on direct expansion estimates when using the 60 m mask. This allows the use of coarser resolution sensors that increase the availability of suitably timed imagery available to the surveys at low cost.

The distribution of poppy can be mapped by plotting the probability of poppy associated with each stratum for each pixel. This is useful for examining trends in cultivation, for example, figure 8 shows higher cultivation density in newly exploited areas to the north of the main irrigated area in Nad Ali, Helmand Province in 2009. However, since the probabilities are from a global estimate, the accuracy of any estimate from a single or group of pixels is unreliable (Tomppo et al. 2008). In general the homogeneity of poppy cultivation in the strata is low and the probabilities should be used with caution when used to map spatial distribution. Areas of mixed spectral response and high variance in poppy cultivation are open to misinterpretation at the local scale. A suggested improvement to the stratification is to include a spatial component to the clustering to reduce the effect of mixed strata between areas with different cultivation practices. Use of higher frequency image collections could also provide information to stratify areas where temporal cropping patterns are associated with poppy cultivation.

Stratification was performed using ISODATA with no prior information of the relationship between the clusters and the probability of poppy cultivation. The clustering results may not be optimal as they are dependant on the pre-selection of the number of clusters and the initial cluster means. The splitting of heterogeneous clusters reduced the variance in some strata but the selection of an appropriate precision for a homogeneous cluster will vary with the underlying variance in the sample data. The total number of strata splits is also limited by the sample fraction as the estimates from under-sampled strata will become unreliable. Optimising the stratification to maximise the precision of the estimate and a reliable measure of the spatial homogeneity of poppy within the strata requires further work.

These findings highlight the importance of image timing and resolution when using remote sensing methodologies for crop monitoring. They show the benefits of imagery based stratification in sample surveys and support its use in other geographic regions for improving agricultural statistics.

## Acknowledgements

The authors would like to thank the UK government for sponsoring the research.

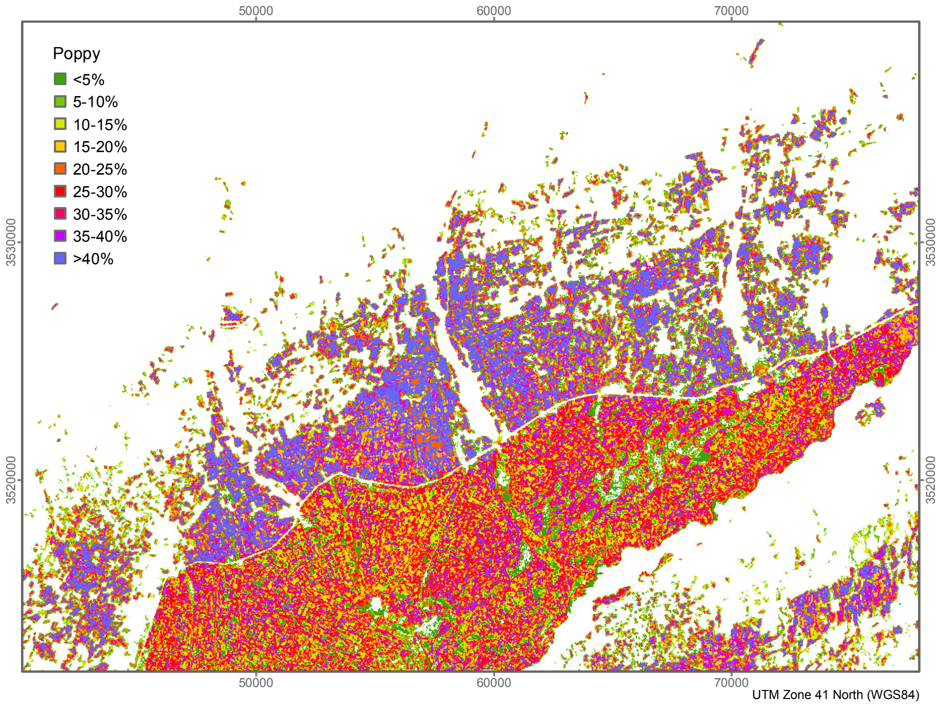


**Figure 8.** Poppy distribution map for part of Helmand Province 2009 showing a higher concentration of poppy cultivation north of the main irrigation canal.

## References

Carfagna, Elisabetta, and F Javier Gallego. 2005. "Using Remote Sensing for Agricultural Statistics." *International Statistical Review* 73 (3): 389–404. http://dx.doi.org/10.1111/j.1751-5823.2005.tb00155.x.

Cochran, W G. 1977. *Sampling Techniques*. 3rd ed. New York: John Wiley & Sons.

Grodecki, J, and G Dial. 2001. "IKONOS Geometric Accuracy." In *Proceedings of Joint Workshop of ISPRS Working Groups I/2, I/5 and IV/7 on High Resolution Mapping from Space 2001*, University of Hannover.

Jia, Kun, Bingfang Wu, Yichen Tian, Qiangzi Li, and Xin Du. 2011. "Spectral Discrimination of Opium Poppy Using Field Spectrometry." *IEEE Transactions on Geoscience and Remote Sensing* 49 (9SI): 3414–3422.

Koeln, Gregory T, and R. Peter Kollasch. 2000. "Crop area assessments using low, moderate, and high resolution imagery: a Geotools approach." Bethesda, MD, USA. American Society of Photogrammerty and Remote Sensing.

Luders, J K, W G Wilson, and R G Gardener. 2004. "Considerations when using area frame sampling tools: Afghanistan's illicit poppy cultivation estimate." Annual Conference 25–28 March, Denver, CO, USA.

McRoberts, Ronald E, Daniel G Wendt, Mark D Nelson, and Mark H Hansen. 2002. "Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates." *Remote Sensing of Environment* 81 (1): 36–44. http://www.sciencedirect.com/science/article/pii/S0034425701003303.

Shahriar Pervez, Md., Michael Budde, and James Rowland. 2014. "Mapping irrigated areas in Afghanistan over the past decade using MODIS NDVI." *Remote Sensing of Environment* 149: 155–165.

http://www.sciencedirect.com/science/article/pii/S0034425714001461.

Simms, Daniel M, Toby W Waine, John C Taylor, and Graham R Juniper. 2014. “The application of time-series MODIS NDVI profiles for the acquisition of crop information across Afghanistan.” *International Journal of Remote Sensing* 35 (16): 6234–6254.

Taylor, J C, C Sannier, J Delincé, and F J Gallego. 1997. *Regional Crop Inventories in Europe Assisted by Remote Sensing: 1988-1993.* Technical report. Luxembourg.

Taylor, J.C., T.W. Waine, G.R. Juniper, D.M. Simms, and T.R. Brewer. 2010. “Survey and monitoring of opium poppy and wheat in Afghanistan: 20032009.” *Remote Sensing Letters* 1 (3): 179–185.

Tomppo, Erkki, Håkan Olsson, Göran Ståhl, Mats Nilsson, Olle Hagner, and Matti Katila. 2008. “Combining national forest inventory field plots and remote sensing data for forest databases.” *Remote Sensing of Environment* 112 (5): 1982–1999. http://linkinghub.elsevier.com/retrieve/pii/S0034425708000242.

UNODC. 2016. *World Drug Report 2016.* Technical report.

# Improved estimates of opium cultivation in Afghanistan using imagery-based stratification

Simms, Daniel M.

2017-03-30

Daniel M. Simms, Toby W. Waine and John C. Taylor. Improved estimates of opium cultivation in Afghanistan using imagery-based stratification. International Journal of Remote Sensing, 2017, Volume 28, Issue 13, pp3785-3799

http://dx.doi.org/10.1080/01431161.2017.