B09

Opium yield estimates in Afghanistan using remote sensing

D. M. Simms | Cranfield University | Cranfield, Bedfordshire | UK
T. W. Waine | Cranfield University | Cranfield, Bedfordshire | UK
DOI: 10.1481/icasVII.2016.b09c

ABSTRACT

Accurate estimates of opium production are essential for informing counter-narcotics policy in Afghanistan. The cultivated area of opium poppy is estimated remotely by interpretation or digital classification of very high resolution (VHR) satellite imagery at sample locations. Obtaining an accurate estimate of average yield is more challenging as poor security prevents access to a sufficient number of field locations to collect a representative sample. Previous work carried out in the UK developed a regression estimator methodology using the empirical relationship between the remotely sensed normalised difference vegetation index (NDVI) and the yield indicator mature capsule volume. The application of the remote sensing approach was investigated in the context of the existing annual opium survey conducted by the United Nations Office on Drugs and Crime and Afghanistan's Ministry of Counter Narcotics (UNODC/MCN) and indicated the potential for bias correction of yield estimates from a small targeted field sample. In this study we test the approach in Afghanistan using yield data and VHR satellite imagery collected by the UNODC/MCN surveys in 2013 and 2014.

Field averaged measurements of capsule volume were compared to field averaged NDVI extracted using visual interpretation of poppy fields. The study compares the empirical relationships from the UK field trials with the Afghanistan data and discusses the challenges of developing an operational methodology for accurate opium yield estimation from the limited sample possible in Afghanistan.

Keywords: Opium poppy, remote sensing, yield bias correction

PAPER

1. Introduction

The United Nations Office on Drugs and Crime and Afghanistan's Ministry of Counter Narcotics (UNODC/MCN) produce annual opium production estimates based on survey data collected across Afghanistan. The opium yield estimate is based on the empirical relationship of dry opium gum to mature capsule volume per unit area, measured at sample locations. The indirect yield measurement of mature capsule volume is used as it is quicker than removing and drying capsules for weighing.

Sample sites are chosen randomly from a sample frame of the villages located in the poppy producing provinces. At each selected village three fields are subjectively chosen by a field surveyor representing a poor, average and good quality poppy crop. In each field, measurements of capsule volume (capsule height and diameter) and the number of yielding capsules are made within three 1 m^2 quadrats arranged randomly along a transect. The average dry opium gum yield (Y) in kg ha^{-1} is calculated for each quadrat using the model (UNDCP, 2001)

where Vc is the mature capsule volume ($cm^3\ m^{-2}$). The average yield per unit area weighted by province area is then calculated and multiplied by the cultivated area of poppy to estimate the total production.

$$Y = \frac{(V_c + 1495) - ((V_c + 1495)^2 - 395.259 \cdot V_c)^{0.5}}{1.795} \quad (1)$$

There are difficulties with implementing the yield survey because security limits access to poppy fields within the major opium producing areas. Surveyors are unable to access many of the field locations from the sample frame which results in a low number of surveyed fields and a non-random distribution of samples (UNODC, 2012). The statistically un-representative sampling will bias the average dry opium yield and hence the yearly estimate of opium production.

A methodology for bias correction of opium yield estimates was proposed by Waine et al.(2014) using the existing very high resolution (VHR) imagery and the non-random field observations from the opium yield survey. The approach makes use of VHR imagery collected at random sample locations for visual interpretation of poppy crops as part of the cultivation component of the UN-ODC's annual opium survey. For each image, a bias-corrected estimate is made by substituting the field average Normalised Difference Vegetation Index (NDVI) of all poppy fields (population) into the linear regression of NDVI and mature capsule volume determined using a small number of high-quality field observations. As each image is randomly selected an unbiased estimate of opium yield can then be made.

Remote sensing of yield exploits the relationship between linear combinations of red and near-infrared reflectance, known as vegetation indices (VI), and crop biophysical parameters. VIs such as the NDVI are correlated to leaf area index and above-ground biomass which are in turn correlated to final yield in crops such as wheat, millet, soybean, cotton, barley, and maize (Domenikiotis et al., 2004; Liu and Kogan, 2002; Prasad et al., 2006; Rasmussen, 1997; Tucker et al., 1980; Weissteiner and Kühbauch, 2005).

Wood et al. (2003) demonstrated how image derived NDVI could be used for mapping within-field variability in crops of wheat (Triticum aestivum) and barley (Hordeum vulgare L.) using a small number of ground samples to construct a regression based model. They found as few as eight sample observations could be used to correlate yield indicators to NDVI for groups of fields, with a single observation per field. Based on this approach, the relationship between opium yield indicators and NDVI was inves-tigated by Taylor et al. (2010) as part of wider research into improving opium production estimates using remote sensing. Field trials conducted in the UK showed a high correlation between NDVI and measurements of mature capsule volume ($R^2 = 0.70$) across multiple fields and three growing seasons (Waine et al., 2014).

Investigations using 2011 yield and VHR image data from the UNODC's annual survey in-dicated the potential for operational use but there were too few ground observations per image to validate the methodology. The following conditions were recommended by Waine et al. (2014) for field measurements in support of the bias-correction approach: (1) accurately geo-referenced ground data; (2) samples should be distributed across the range of crop variation with respect to NDVI; (3) ground measurements of poppy crop parameters must be accurately co-registered with ortho-rectified imagery; and (4) imagery used to derive NDVI should be targeted for collection around the flowering growth stage to maximise the correlation between yield indicators and NDVI. This conference paper presents investigations into a bias-correction methodology using ground observations and image data from the UNODC's 2013 and 2014 annual opium surveys.

2. Data and methodology

2.1. Yield datasets

Field observations from the annual opium survey in 2013 and 2014 were supplied by the UN-ODC for sample sites with coincident VHR satellite images collected for poppy interpretation. This data included quadrat level measurements of the number of capsules/flowers and mature capsule vol-ume together with ground photography and GPS data. The UNODC also supplied interpretations of poppy fields as digital boundary files for each image. Quadrat level data was used to calculate the mean volume of mature capsules per unit area ($cm^3\ m^{-2}$) for each sampled field. Confidence inter-vals (95%) for the mature capsule volume (field average) for each image were calculated using the formula:

$$CI = \bar{y} \pm \left(\frac{s}{\sqrt{n}} \cdot 1.96\right) \quad (2)$$

where $\bar{y}$ is the mean mature capsule volume for the image,

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}, \quad (3)$$

y_i is a sample observation (field average) and n is the number of sample observations.

2.2. Image data

Multispectral QuickBird (QB) and WordView2 (WV2) satellite image scenes were supplied by the UNDOC. The image data had a spatial resolution of 3.6 m (QB) and 2 m (WV2), spectral resolutions of 630–690 nm (QB and WV2) in the red and 760–900 nm (QB) and 772–890 nm (WV2) in the near-infrared. The timing of each VHR image in relation to crop growth-stage was assessed using high-frequency satellite imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Simms et al., 2014). Timing of growth-stage was found to be similar within individual scenes. Table 1 shows a summary of the yield datasets for 2013 and 2014. Hilmand 2013 (held8) was not considered for further analysis as the satellite imagery was collected too late in the growing cycle. The Uruzgan images in 2014 (uruz-erad1 & uruz4) were collected on the same orbital pass and treated as a single image scene. Each image was corrected to top-of-atmosphere reflectance to minimise spectral differences between the WorldView 2 and QuickBird sensors (Pacifici et al., 2014). The images were then or-thorectified to reduce geometric distortion and

NDVI calculated as (ρNIR −ρR)/(ρNIR +ρR), where ρNIR is near-infrared and ρR is the ref reflectance.

2.3. Mature capsule volume and NDVI

Within each selected image the location of the sampled fields was identified using a combination of GPS coordinates, geo-tagged ground photography and visual interpretation of pan-sharpened VHR imagery. The field-average NDVI was then extracted for each sample field as the mean of the image pixels intersecting the field polygon. For each image the empirical relationship between mature capsule volume and NDVI was de-termined using linear regression of the form y = a(NDVI) + b, where y is the mature capsule volume

Table 1: *Summary of 2013 and 2014 yield datasets provided by UNODC/MCN for this study and multispectral satellites QuickBird (QB), WorldView2 (WV2) at 3.6 and 2 m pixel size respectively. Image held8 excluded as poppy crop was cleared*

Year	Image	Reference	Province	Acquisition	Sensor	Poppy growth stage (image)	Fields analysed	Poppy parcel interpretation
2013	*held8*		Hilmand	22/04/2013	QB	cleared	0/2	yes
	beld	HLD	Hilmand	27/04/2013	WV2	sceneced	5/10	no
	kand-cnc p3-4	KDR	Kandahar	14/04/2013	WV2	capsule	6/6	yes
	uruz-target1	URZ	Uruzgan	27/04/2013	QB	capsule	3/9	yes
	farah-erad1	FRH	Farah	13/04/2013	WV2	capsule	12/12	yes
2014	beld	HLD	Hilmand	24/04/2014	WV2	sceneced	6/8	yes - part
	uruz-erad1 & uruz4	URZ	Uruzgan	27/04/2014	WV2	capsule	9/9	no
	fara	FRH	Farah	15/04/2014	QB	flowering	12/12	yes
	bada-erad3-p1-2	BKN2	Badakhshan	26/06/2014	WV2	capsule	15/15	yes
	bada-erad1	BKN1	Badakhshan	26/06/2014	WV2	capsule	9/9	yes
	kand	KDR	Kandahar	10/04/2014	WV2	capsule	3/9	no

per unit area, a is the slope and b is the offset. The 95% confidence intervals of the mean expected value of y for a given NDVI (x^*) were constructed using:

$$\hat{y} \pm t_{0.025,n-2} \cdot s_y \cdot \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{(n-1) \cdot s_x^2}} \quad (4)$$

where s_y is the standard deviation of the residuals, $t_{0.025,n-2}$ is the student t-value for 95% CI (two tailed), s_x^2 is the variance and $\bar{x}$ is the mean NDVI.

The field delineations of poppy crops supplied by UNODC from their cultivation survey were used to extract the average NDVI for each poppy field in the image. Geometric differences between the field polygons and the orthorectified images were corrected by manually shifting individual polygons to align them with the imagery. A 2 m internal buffer was applied to the field parcels to reduce the edge effects.

The bias-corrected mature capsule volume was calculated by substituting the mean field average NDVI for all poppy fields in an image (population mean, $\bar{x}_{pop}$) into the empirical relationship between mature capsule volume and NDVI from the sampled fields.

3. Results

The linear regression of mature capsule volume and field average NDVI for 2013 and 2014 images are shown in figures 1 and 2 respectively. Image 2014 BKN1 was excluded from the analysis as there was no relationship between the mature capsule volume and NDVI from the image. Bias correction using the linear regression models above is summarised in table 2 for six of the nine images from the 2013 and 2014 data. There were no poppy field interpretations for HLD 2013, URZ 2014 and KDR 2014 as either the image was too late, the R^2 was low or there were too few field observations to warrant image interpretation. The correlation between mature capsule volume and NDVI is highly variable between images, with R^2 ranging from 0.08 to 0.42 for images with more than three fields. Images FRH 2014 and BKN2 2014 have an R^2 value greater than 0.4 and yield figure corrections of -17% and 17% respectively, with improved or similar confidence intervals compared to the sample data alone. KDR 2013 and HLD 2014 have low R^2 values indicating high variance in the relationship between mature capsule volume and NDVI. This is shown in the increased CI of the bias corrected mature capsule volume. The bias-corrected yield for FRH 2013 is within 1% of the sample yield despite a very low R^2 of 0.14.

Figure 1 - Field average NDVI plotted against mature capsule volume per unit area for individual image sites in 2013. Dashed lines show 95% confidence interval for the regression line

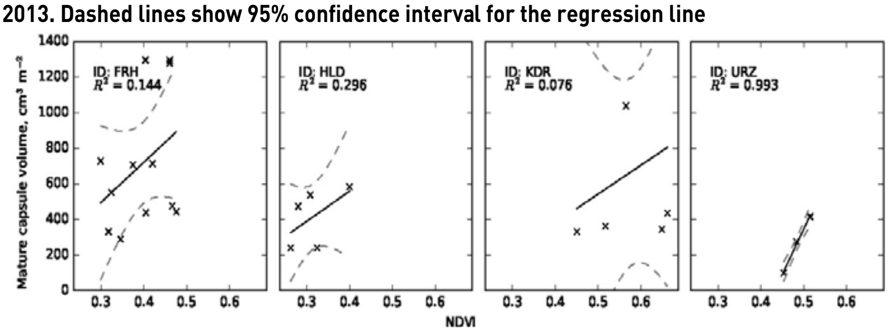
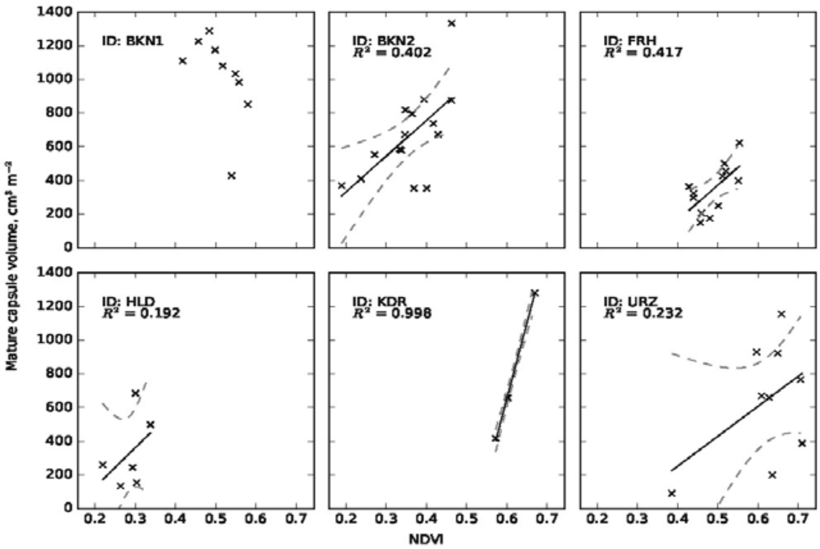


Figure 2 - Field average NDVI plotted against mature capsule volume per unit area for individual image sites in 2014. Dashed lines show 95% confidence interval for the regression line. BKN1 shows no relationship.



4. Discussion and recommendations

The empirical relationships between field average NDVI and mature capsule volume for the Afghanistan data (figures 1 and 2) are poor compared with the data from the UK (R^2 >0.7 at flower-ing). The potential reasons for the discrepancy are the differences in agricultural practices and poppy varieties, the differences in the methodology between the UK field trials and the Afghanistan data collection, and the quality of the yield observations.

The manual cultivation methods used in Afghanistan are likely to result in greater variation in the quality of crops than the standardised mechanised practices used in the UK. In the UK field trials, sample locations were co-located with pixel level accuracy (<1 m) across the range of NDVI variability. In Afghanistan accurate co-location of the quadrat locations (<1 m) is not possible using the current GPS units so the field-averaged yield and NDVI are used instead. Detailed examination of the Afghanistan field data and ground photography suggests that the average of the quadrat level data may not be representative of individual fields. Many of the quardats are located along field boundaries, which may introduce sample bias if these areas are consistently higher or lower yielding than the rest of the field.

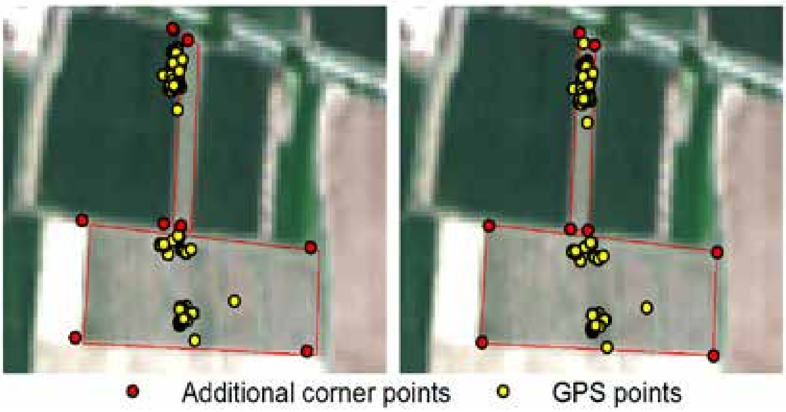
Table 2 - Results of bias correction methodology applied to 2013 and 2014 image data. NDVId is the difference between the mean NDVI of the sample fields and the mean NDVI of all poppy fields in the image. CIsamp is the 95% confidence interval of the sample yield estimate and the CIcorr is the 95%confidence interval of the regression model yield estimate.

Year	ID	No. fields	Yield sample mean ($cm^3\ m^{-2}$)	$\pm CI_{samp}$ (%)	Yield bias corrected ($cm^3\ m^{-2}$)	$\pm CI_{corr}$ (%)	R^2	NDVI range all fields	NDVI range sam-pled	Yield differ-ence (%)	$NDVI_d$
2013	URZ	3	262	69	395	13	0.99	0.439	0.062	51	0.026
2013	FRH	12	712	30	721	32	0.14	0.401	0.177	1	0.004
2013	KDR	6	669	59	577	114	0.08	0.659	0.211	-14	-0.057
2014	FRH	12	347	23	288	30	0.42	0.215	0.125	-17	-0.028
2014	HLD	6	329	53	428	77	0.19	0.192	0.119	30	0.042
2014	BKN2	15	666	20	778	18	0.40	0.496	0.273	17	0.053

There is also the potential for increased NDVI values relating to the presence of weeds where the density of opium cropping is low. Visual inspection of the quadrat photographs for KDR 2013 showed inconsistencies between capsule counts and the recoded data. There was also ambiguity in the labelling of sampled fields and data points. There is potential for error in matching the NDVI values with the correct field data as the plot photo-locations are used for linking the fields in the database with the image interpreted field parcels.

A better correlation of NDVI with mature capsule volume could be obtained by accurately co-locating individual quadrat measurements with the image pixels. The GPS coordinates and image ortho-rectification are not sufficiently accurate to ensure pixel level geo-location. However, by col-lecting the corner positions of the sampled fields the sample data can be shifted to better align with the image data (see figure 3).

Figure 3 - Suggested additional GPS point locations at field corners allow improved placement of yield points with imagery. GPS points shifted to image (right) using field corners.



The range of NDVI values from all poppy fields in the image is much greater than the range of NDVI for the sampled fields (table 2). This suggests the current approach of sampling the range of yield by subjective assessment of crop quality is not capturing the true variability.

Some of the errors in field data collection and the placement of sample quadrats are related to the security conditions faced by the field survey teams. Taking field measurements and photographs exposes surveyors to personal risk, which is reduced by taking measurement at the corner of fields or behind compound walls. An advantage of using an imagery based approach is that NDVI data can be used to preselect fields to sample. A range of field sites representing the range of NDVI could be pre-sented to the survey teams provided that images are collected around the crop flowering period. Once in the area of the image, a sub-selection could then be made based on an in-field risk assessment. This would also improve the field selection process compared to the current subjective assessment of a low, medium or high yielding crop. Implementing this approach would require improved communication with the field survey teams. Images would need to be processed and visually interpreted for opium poppy fields in the 2-3 weeks between image collection and capsule maturity when yield observations are made in the field.

Imagery that coincides with the 3–4 week period around poppy flowering has the the highest R^2 and greatest stability of the relationship between NDVI and capsule volume in UK data (Waine et al., 2014). The timing of the UNODC imagery from both years is towards the end of the growth cycle and close to crop senescence, where the relationship between potential yield and NDVI breaks down.

The remote sensing approach aims to relax the requirement for a random distribution of sample points and to reduce the number of samples required to make a yield estimate. It does not correct for error in the yield observations, which was found to vary between the images locations. Data recording errors and mislabelling that result in outliers in the regression of mature capsule volume and NDVI will have a significant effect on the results. The suitability of the imagery also varies according to its timing in relation to the crop growth stage. Imagery collected late in the growing season will have a shorter range in NDVI and less stable relationship with potential yield resulting from a reduction in photosynthesis as the crop senesces. As such the approach could not be used for all of the 2013 and 2014 images. However, the potential of the approach can be seen in the improved yield estimates for FRH 2014 and BKN2 2014. These image sites had more than 6 field locations, imagery at flowering or capsule stage and $R^2 > 0.4$.

A small targeted yield sample remains the best approach to improve the accuracy of yield ob-servations of mature capsule volume. However, field data quality is likely to remain highly variable given the unpredictable security conditions in which surveyors work. It is recommended to apply the following criteria for each image before applying a bias-correction: (1) suitable image timing (ide-ally 2–4 week period around poppy flowering); (2) ground locations of measured fields accurately determined in imagery; (3) minimum of six ground data points, positive correlation and $R^2 > 0.4$; and (4) image mean NDVI of poppy fields within the bounds of the model. The value of R^2 suggested is supported by findings in Taylor et al. (1997), where an $R^2 > 0.3$ was recommended for application of a regression estimator in cultivation surveys.

5. Conclusions

A bias-correction methodology using VHR satellite imagery was applied to the yield data of the UNODCs annual opium survey from 2013 and 2014. The aim of the approach is to reduce bias caused by the non-random distribution of the sample and reduce the number of samples required to make an estimate. Of the 12 VHR images, 6 were bias corrected with 2 resulting in an improved yield estimate. The remaining images had a low R^2 for the empirical relationship between mature capsule volume and NDVI, which was attributed to the quality of the yield observations and the accuracy of field-averaged yield data calculated from three quadrat locations in heterogeneous fields. Improvements to the survey methodology were recommended to improve the empirical rela-tionships between NDVI and capsule volumes. These were to optimise the timing of VHR imagery collection to coincide with poppy flowering, improve the relative alignment of quadrat GPS locations using field corners, and using the stratified NDVI image to target field locations to sample. Criteria for applying the bias correction methodology for operational use were also suggested. The authors acknowledge that the recommendations and quality criteria cannot currently be met in all cases given the complex logistical and security situation in Afghanistan. The UNODC seeks continuing improvement in the quality of ground data collection through improved surveyor training and concentrating its effort on a small and targeted yield sample. In this context the approach should be tested further for operational use in bias correction of yield estimates going forward.

6. Acknowledgements

The authors wish to thank the UNODC and Afghanistan's Ministry for Counter Narcotics for funding the research and permission to publish their data.

References

Domenikiotis, C., M. Spiliotopoulos, E. Tsiros, and N. R. Dalezios (2004). Early cotton production assessment in Greece based on a combination of the drought Vegetation Condition Index (VCI) and the Bhalme and Mooley Drought Index (BMDI). International Journal of Remote Sensing 25.23, pp. 5373–5388.

Liu, W. T. and F. Kogan (2002). Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices. International Journal of Remote Sensing 23.6, pp. 1161–1179.

Pacifici, F., N. Longbotham, and W. J. Emery (2014). The importance of physical quantities for the analysis of multitemporal and multiangular optical very high spatial resolution images. IEEE Transactions on Geoscience and Remote Sensing 52.10, pp. 6241–6256.

Prasad, A. K., L. Chai, R. P. Singh, and M. Kafatos (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. International Journal of Applied Earth Observation and Geoinformation 8.1, pp. 26–33.

Rasmussen, M. S. (1997). Operational yield forecast using AVHRR NDVI data: Reduction of environmental and inter-annual variability. International Journal of Remote Sensing 18.5, pp. 1059–1077.

Simms, D. M., T. W. Waine, J. C. Taylor, and G. 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, pp. 6234–6254.

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: 2003-2009. Remote Sensing Letters 1.3, pp. 179–185.

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

Tucker, C. J., B. N. Holben, J. H. Elgin Jr, and J. E. McMurtrey III (1980). Relationship of spectral data to grain yield variation. Photogrammetric Engineering and Remote Sensing 46.5, pp. 657–666.

UNDCP (2001). Guidelines for Yield Assessment of Opium Gum and Coca Leaf from Brief Field Visits. UNDCP Scientific Section and Illicit Crop Monitoring Programme. New York.

United Nations Office on Drugs and Crime (2012). Afghanistan Opium Survey 2012. Vienna.

Waine, T. W., D. M. Simms, J. C. Taylor, and G. R. Juniper (2014). Towards improving the accuracy of opium yield estimates using remote sensing. International Journal of Remote Sensing 35.16, pp. 6292–6309.

Weissteiner, C. J. and W. Kühbauch (2005). Regional Yield Forecasts of Malting Barley (Hordeum vulgare L.) by NOAA-AVHRR Remote Sensing Data and Ancillary Data. Journal of Agronomy and Crop Science 191.4, pp. 308–320.

Wood, G. A., J. C. Taylor, and R. J. Godwin (2003). Calibration methodology for mapping withinfield crop variability using remote sensing. Precision Agriculture - Managing Soil and Crop Variability for Cereals 84.4, pp. 409–423.