MEMORANDUM

Fr: Bryce Contor
To: ESPAM2 Model Files
ESHMC
Date: 6 July 2009

Re: Irrigated Lands and Reduction for Non-irrigated Inclusions

---

This memo is intended to record and report on irrigated lands data and reductions for non-irrigated inclusions, for ESPAM2 modeling. The goal is to present this information to the ESHMC more quickly than would be possible with a formal design document. This memo has not been internally reviewed within IWRRI, nor reviewed by IDWR personnel. We request response from the ESHMC by 20 July 2009.

Background

An important input to the aquifer model is a map of irrigated lands. Typically, maps omit small-scale details such as corrals, homesteads, field roads, etc., which are referred to as "non-irrigated inclusions." To impart the correct acreage for calculation of incidental recharge from surface-water irrigation and net pumpage due to groundwater irrigation, a second important input is a reduction fraction to account for non-irrigated inclusions.

In ESPAM1.1, and as currently proposed for ESPAM2, the Recharge Tool calculates irrigated acreage in each model cell as follows:

$$IrrAcr = GIS\_Acr \times (1 - RED) \qquad (1)$$

IrrAcr = actual irrigated acres in model cell
GIS_Acr = acreage determined by Geographic Information System (GIS) analysis[1]
RED = reduction for non-irrigated inclusions.

In ESPAM1.1, three irrigated-lands data sets were available, but only one data set of high-resolution data for calculation of parameter RED. The three irrigated-lands data sets were developed using different data sources and different methodology. The general trend of declining irrigated acreage over time was counter to expectations. IWRRI determined (and discussed with the ESHMC) the distinct possibility that the differences were due to differences in data and/or

---

[1] Calculation of GIS_Acr also includes consideration of mixed-source lands and source fractions, but that detail is not necessary for the discussion here.

methodology, rather than due to actual changes in irrigated acreage. Therefore, we used a single irrigated lands data set for the entire calibration period (spring 1980 through spring 2002). The RED parameter for ESPAM1.1 was 0.12, for both sprinkler and gravity irrigated lands.

ESPAM1.1 design documents discuss these data, the calculation process and this decision in more detail.

Data Available for ESPAM2

Irrigated Lands Data: Data available for ESPAM2 include the data sets considered for ESPAM1.1: The 1980 RASA irrigated lands data (obtained from LANDSAT 5 images), the 1992 irrigated lands data that were used in ESPAM1.1, (developed from aerial photos by IDWR and BOR) and year-2000 irrigated lands data from LANDSAT 7 developed by IDWR using seasonal patterns of visual-band color changes.

Additional data include a 1986 irrigated lands map and a "placeholder" 2006 irrigated lands map derived from Normalized Difference Vegetative Index (NDVI) analysis of LANDSAT data by IDWR. High-resolution 2006 irrigated lands maps derived from the NDVI analysis, USDA polygons of parcel boundaries, and painstaking hand analysis by IDWR will replace the "placeholder" data set when available.

Data for Determination of Actual Irrigated Area: In order to utilize all the irrigated-lands data sets, despite differences in data source and methodology, we decided to develop unique "true acreage" sample data for each data set. This would allow construction of unique reduction factors for each data set, overcoming differences in methods and underlying data.

Images for "true acreage" samples vary in presentation and resolution, as described in Table 1.

Table 1
Images Available for "True Acreage" Samples
for ESPAM2

Year	Data Source	Color	Resolution	Coverage
1980	U2 aerial transparencies	near infrared false color	~ 30 meter[2]	partial
1983	U2 aerial transparencies	"	"	"
1987[3]	IDWR	”	~10 meter[4]	nearly

[2] Original images are very high resolution but we obtained scans at only 70 dpi.

Year	Data Source	Color	Resolution	Coverage
	Adjudication			complete
1992[5]	DOQQ aerial photos	black and white	"	partial
2004	NAIP	true color	1 meter	nearly complete
2006	NAIP	true color	2 meter	complete

Conceptual Background

The underlying concept for calculation of reduction factors is to isolate a statistical sample of irrigated parcels, and for each parcel compare the nominal acreage from the given data set with acreage of detailed maps hand drawn from high resolution data. In Figure 1, the light polygon represents the nominal irrigated parcel from one of the irrigated-lands data sets and the heavy line represents the hand-drawn parcel from one of the high-resolution images. If the nominal parcel were 100 acres and the hand-drawn parcel were 95 acres, the RED value for this parcel would be 0.05 [(100 - 95) / 100].

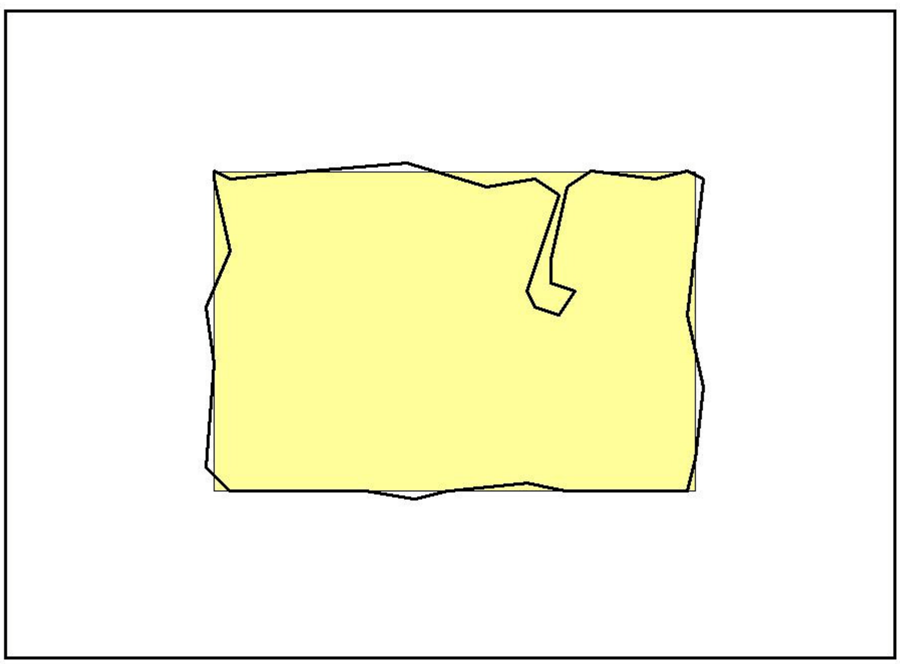

Figure 1. Hypothetical nominal irrigated parcel (colored) and hand-drawn "actual acreage" parcel (dark heavy line).

[3] Nominally 1987 but image dates appear to range from 1986 through 1992. These are administrative-basin mosaic images so the image date of any particular parcel is uncertain.
[4] Pixel size is actually smaller but underlying image resolution appears to be closer to 10 meter.
[5] Some images are 1993 and 1998, but no 1998 images were used in this effort.

By repeating this process for a statistical sample of irrigated parcels, we can calculate an overall RED value (with appropriate confidence limits) for each data set.

Challenges

Challenges for application of this concept to ESPAM2 data include the following:

1. Mismatched parcels. In order to avoid biasing the hand-drawn polygons, we did not look at the irrigated-lands polygons while drawing detailed boundaries from images. Determination of irrigated/non-irrigated status is difficult for the true-color and black and white images. It can also be difficult if a near-infrared image had been taken early or late, when crops were not actively growing. Our goal was to let the images define geometry, but accept the various data sets' determination of irrigation status. One consequence of this decision is that many sample locations include the condition illustrated in Figure 2: Some hand drawn parcels are not included in the irrigated-lands data set, and vice versa.

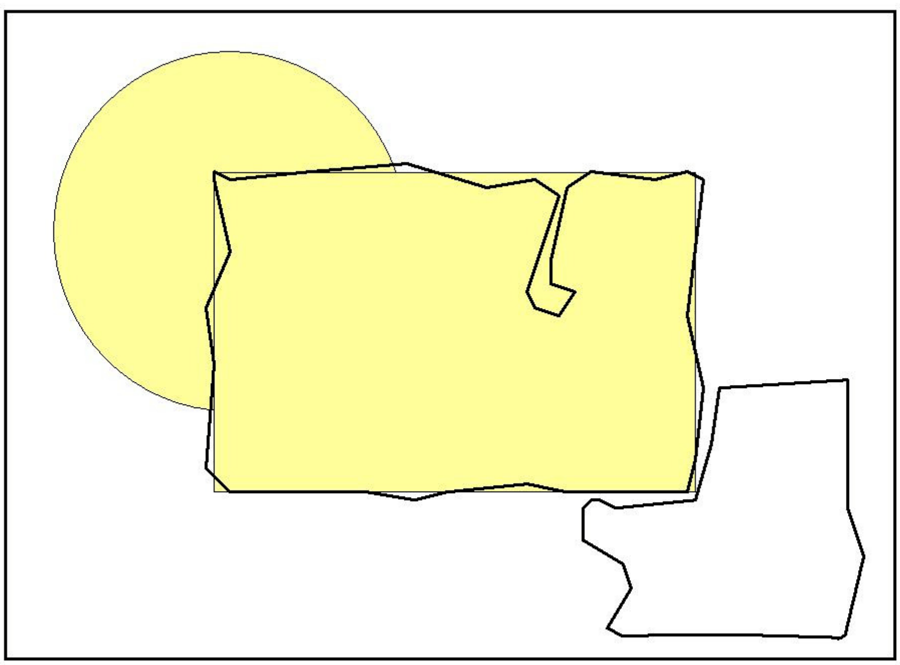

Figure 2. Hypothetical sample area where there is a mismatch between nominal irrigated parcels and hand-drawn parcels.

2. Photo resolution. Different photo resolution can invite differences in the number of inclusions that are represented in the hand-drawn data. If a given parcel had not changed over time, ideally we would desire the hand-drawn parcels to be identical for all the available image data. However, a small road

may only be visible in the 2004 data set, so the hand-drawn parcel for 2004 could indicate fewer irrigated acres than the others, although nothing on the ground had changed.

3. Georeferencing. Georeferencing of irrigated-lands data sets and images may differ. In the center of a large irrigated area, the differences may self-cancel, but for parcels on the boundary between irrigated and non-irrigated areas, errors could be introduced. This is illustrated conceptually in Figure 3. The two polygons are identical to the polygons in Figure 1, but spatially offset. If the sample area were represented by the dotted red line, it would appear that the hand-drawn parcel was perhaps 80% of the size of the data-set polygon, instead of the correct value of 95%.

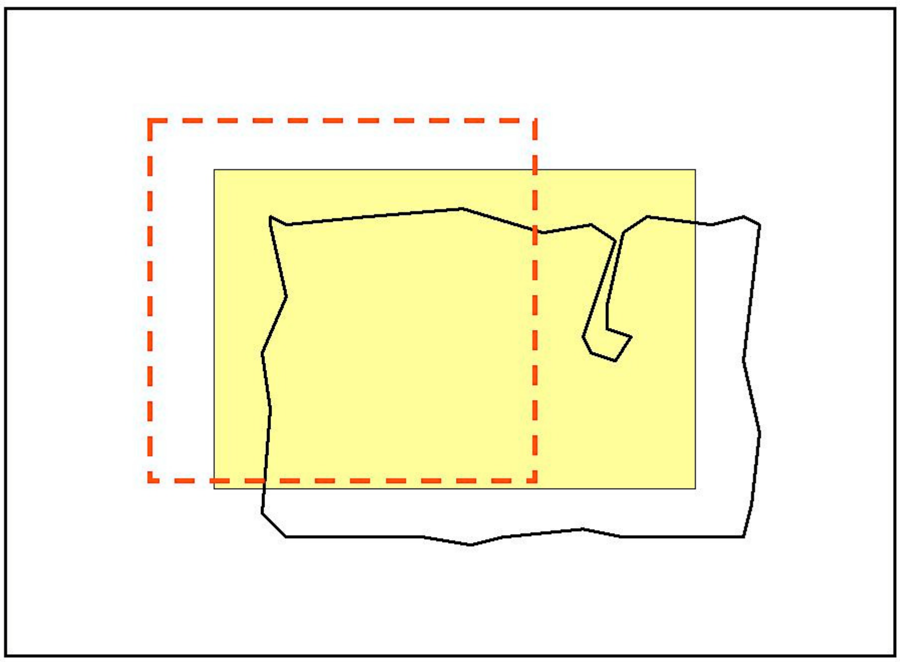

Figure 3. Hypothetical georeferencing offset between data-source parcel (light colored) and hand-drawn parcel (dark solid line). The dashed line is the hypothetical sample box.

4. Parcel boundary identification. In some areas, irregular field boundaries and low-resolution photos made determination of parcel boundaries difficult. Figure 4a shows a 1980 U2 aerial photograph of a particular sample location. Figure 4b shows the 1980 RASA data set indication of irrigated status, and Figure 4c shows the hand-drawn parcel boundaries.

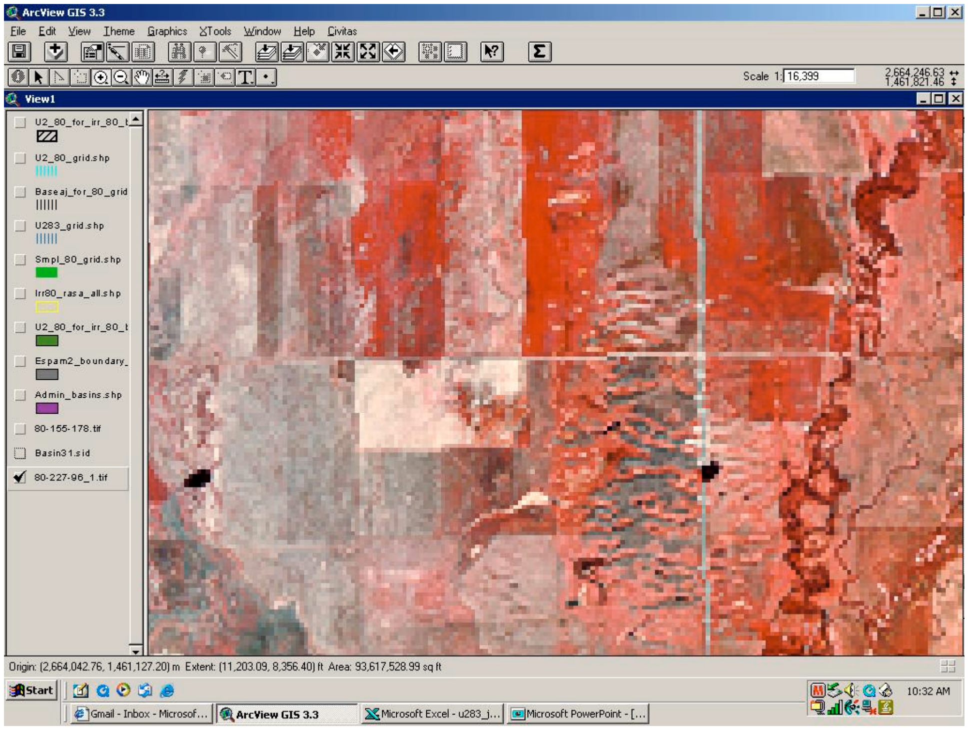

Figure 4a. 1980 U2 aerial photo of one sample location.

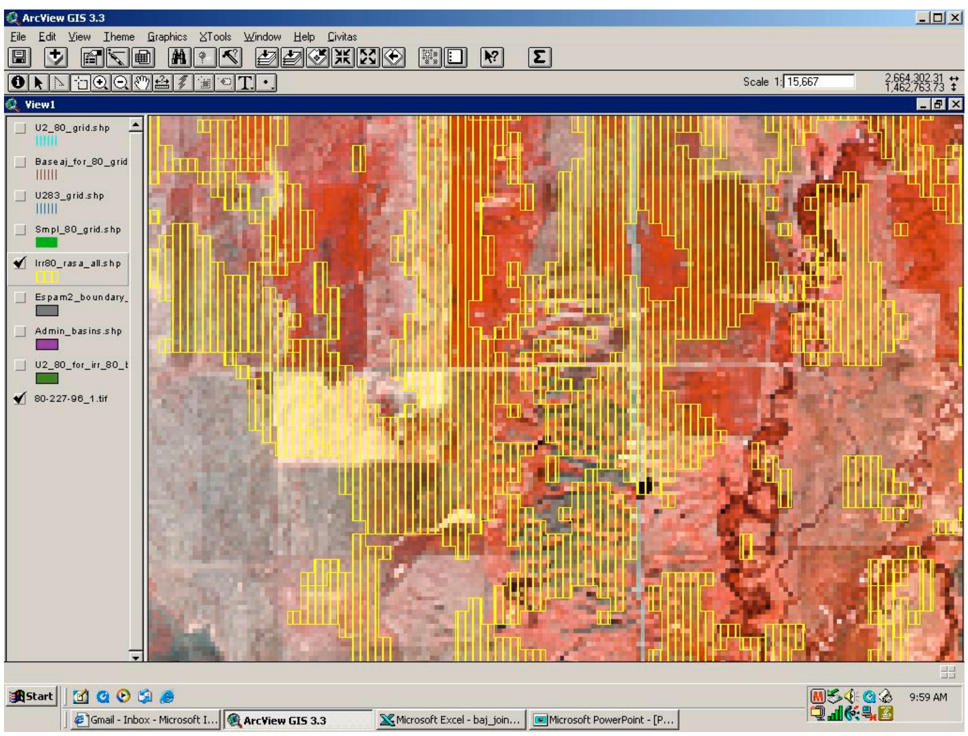

Figure 4b. 1980 RASA irrigated lands data set for location shown in Figure 4a.

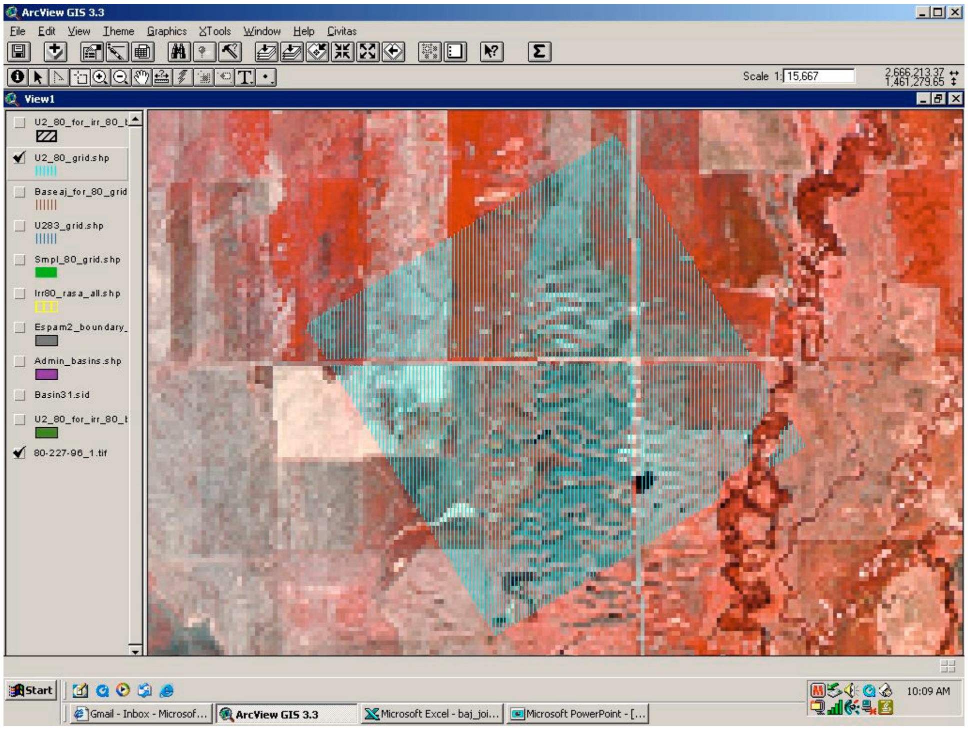

Figure 4c. Hand-drawn irrigated parcels (blue stripe pattern) for sample location shown in Figure 4a.

5. Mismatch in data years. In many cases, the available image year is not the same year as the tested data set. Results can be confounded by changes in actual parcel geometry that took place in the intervening year(s).

6. Difficulty in identifying application method. The Recharge Tool includes the capability to assign separate RED values for sprinkler irrigated and gravity irrigated parcels. While we could easily identify center-pivot sprinkler systems and border or wild-flood gravity systems in aerial photos, field testing indicated that in photo identification, we could not reliably distinguish between furrow irrigation and non-pivot sprinkler systems.

Responses to Challenges

1. Mismatched parcels. Our first attempt was to manually go through all the data sets and remove mismatched parcels. This proved to be very labor intensive, and because of image-to-image differences, we were concerned about repeatability and objectivity. Instead, we intersected all the sample parcels and hand-drawn polygons with a regularly-spaced grid and selected for comparison only those grid cells that were near both the hand-drawn "image" polygons and the "sample" polygons from the data sets. This is described in more detail below.

2. Photo resolution. We responded to this challenge in two ways. We first drew polygons from the Adjudication data set, because it covered nearly all the sample parcels and was a medium-resolution image set. For coarser-resolution images, we relied on the Adjudication data set for basic geometry and only used the coarser data to indicate changes to be made. For finer resolution images, we limited the field of view displayed on the GIS screen, so that even on the high-resolution image, the technician drawing polygons would be limited to the degree of information available in the Adjudication images.

3. Georeferencing. We first trimmed the sample data sets to a buffered data set that included the sample areas plus 500 meters. Then we manually adjusted each sample to match the underlying image, then trimmed to the exact sample borders.

4. Parcel boundary identification. The response to this challenge is discussed below in the description of the grid selection process.

5. Mismatch in data years. The response to this challenge is discussed below in the description of the grid selection process.

6. Difficulty in identifying application method. In ESPAM1.1, we were surprised that the difference between RED factors for sprinkler and gravity lands was not statistically significant. Based on this result and our inability to positively identify application method in images, we decided to calculate a single RED factor for each data set and ignore differences between sprinkler and gravity application.

Hand Trimmed Data Set. We constructed a hand-trimmed data set for 1980, for evaluation of the grid selection method. From the hand-trimmed data we obtained the statistical distribution of the ratio of (image acres) to (data sample acres) for all sample polygons.[6] The hand-trimmed data set addressed the following concerns:

1. Mismatched parcels. We deleted all sample parcels that did not appear in the hand drawn image data set, as well as all parcels in the image data set that did not appear in the sample from the irrigated lands data set. In essence, we edited the Figure 2 condition to match Figure 1.

4. Parcel boundary identification. All sample locations where parcel boundaries were ambiguous were deleted from both the image and sample data sets.

5. Mismatch in data years. Whenever it appeared that the underlying geometry was different in image and sample data sets (for instance, pivots in one but not the other), we deleted the sample from both data sets.

[6] Note that RED would be equal to (1 minus ratio).

The result of hand editing was that we deleted about 15% of the sample parcels, and edited many of the rest.

Grid Selection Method. We then experimented with different search radii for the 1980 grid-intersected data sets. One extreme was to select only grid cells that intersected both image and sample data (zero radius); the other was a "no-trim" option that included all grid cells (infinite radius). Two middle-ground options were to include all grid cells within 10 or 35 meters of irrigated polygons of both data sets.

1. Mismatched parcels. The grid selection process directly addresses the concern about mismatched parcels, by eliminating any grid cells that are not nearby to both the sample and image polygons.

2. Ambiguous parcel boundaries. In the process of inspecting individual outliers, it appeared that ambiguous parcel boundaries tended to produce biased outliers; that is, the indicated fraction from the grid selection was usually higher than the fraction from hand-trimmed data. Since using the median (rather than the mean) as a measure of central tendency is an accepted method to reduce the impact of extreme values, we used median values to address the concern about ambiguous parcel boundaries.

3. Mismatch in data years. We did not find a way to explicitly address mismatch in data years, but it appears from Figure 5 that the grid selection process produces adequate results despite the mismatch in data years. Note that the mismatch problem is likely to be worst for the 1980 data set; most of the images used were from the Adjudication data, which are images taken six to 12 years after the data set. All other data sets use images within four years or closer.

Figure 5 shows the medians and confidence intervals from the various search radii, for the 1980 data set.

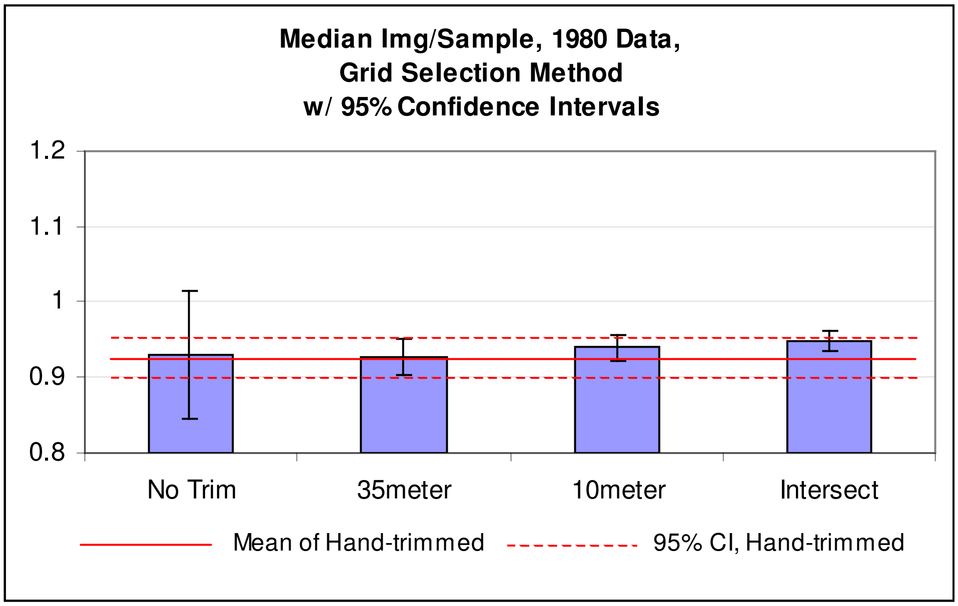

Figure 5. Medians and 95% confidence intervals for the ratio of (image acres) to (sample acres) for 1980 data, compared to mean ratio and confidence interval for hand-trimmed polygons.

While the median values of all the search radii were within the confidence interval of the mean value from hand-trimmed data, the 35-meter search radius was selected because its median was nearest the hand-trimmed mean.

Results

Figure 6 shows the estimated (image)/(sample) ratios from applying this method to the five data sets. In each case, the RED factor would be (1 minus ratio). The dotted red line shows that a single value of approximately 0.93 would be within the confidence intervals of all but one of the data sets. That indicates that we could not argue that any of those data sets had a RED value statistically different from 0.07 (1 - 0.93).[7] Similarly, we cannot argue that the RED value for the 1986 data set is statistically different from zero.

[7] We can argue that these four are significantly different from zero.

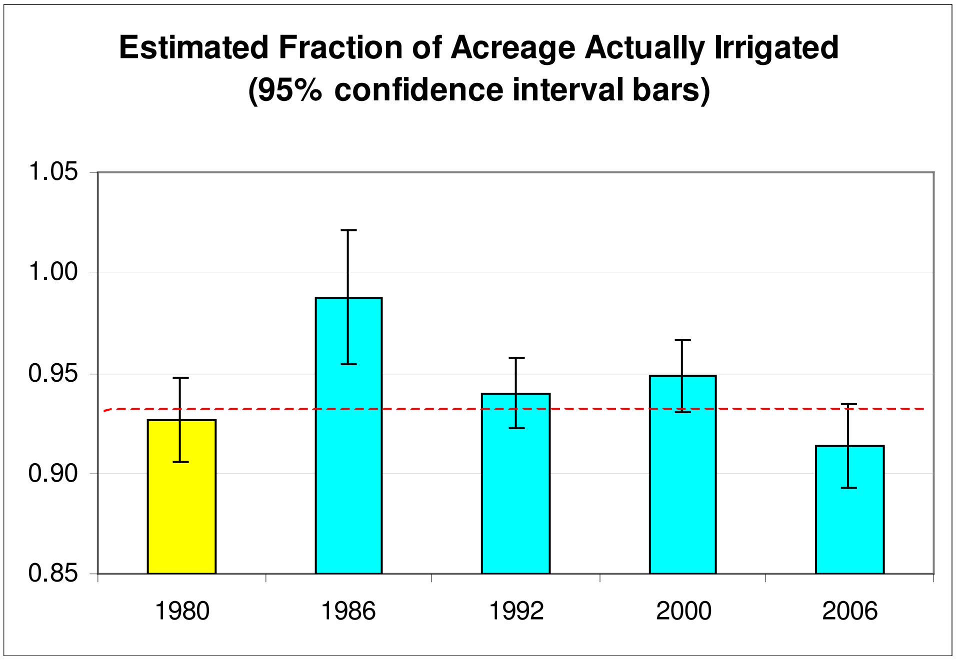

Figure 6. Estimated fraction of acreage actually irrigated [(image acres)/(sample acres)] for five irrigated-lands data sets. Value for 1980 is the mean from the hand-trimmed sample; all others are medians from grid selection method with 35 meter search radius.

Figure 7a shows the raw acreage within the ESPAM2 model boundary[8] along with the acreage after multiplying by (1 - RED). The dashed red line shows a hypothetical common acreage from which none of the acreages is significantly different. Note the wide overlap between the 95% confidence intervals of the data sets. This indicates that apparent temporal patterns could have arisen from random effects. Figure 7b shows the same data with a full-scale vertical axis, to put the temporal differences in practical context.

[8] February 2009 iteration of model boundary. Later versions differ slightly, mostly in non-irrigated areas.

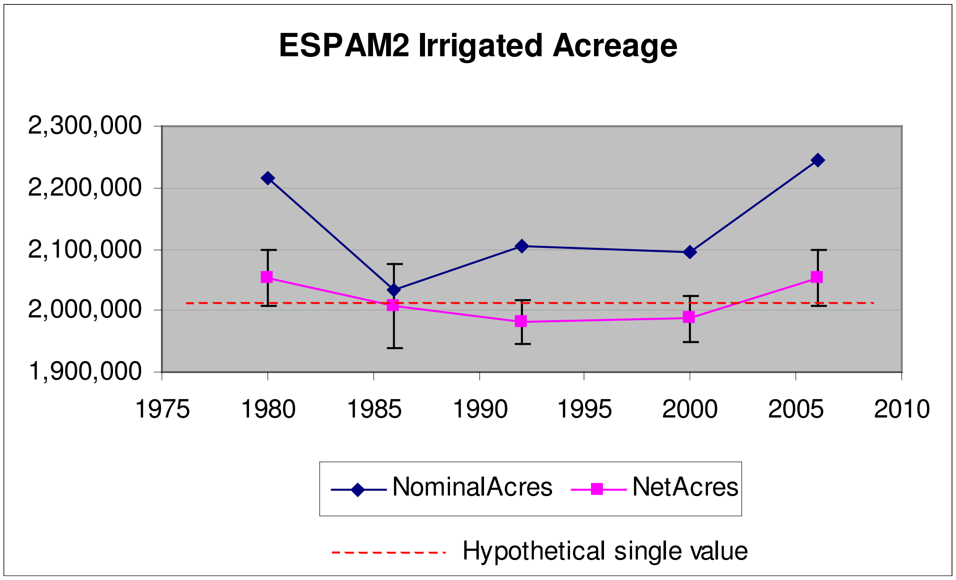

Figure 7a. Raw and adjusted acreage for five data sets, with 95% confidence intervals.

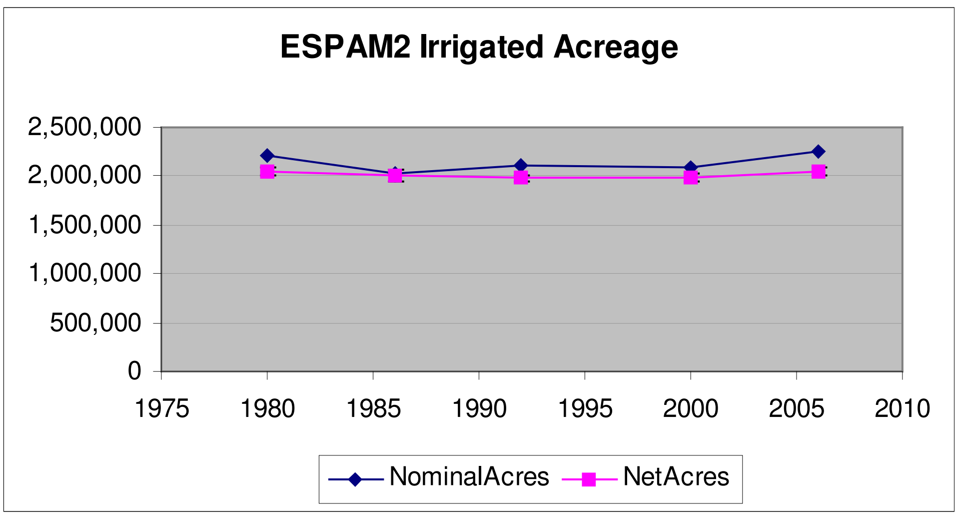

Figure 7b. Raw and adjusted acreage for five data sets with full-scale vertical axis, to show practical significance of temporal differences.

Other Considerations

ET and Inclusions. During ESPAM1.1 calibration, the ESHMC and Dr. Rick Allen discussed the possibility that small non-irrigated inclusions collect solar radiation and impart energy to adjacent crops by advection, reducing some of the impact of non-irrigated inclusions. If this is true, one could argue that no adjustment should be made for non-irrigated inclusions. However, since the non-irrigated inclusions for the 1986 data set are significantly different from some of the other data sets, the net effects must still differ. We believe that unique RED factors for each data set will accommodate differences between the data sets, and a single set of evapotranspiration adjustment factors for the full calibration period will account for any differences between nominal and actual effects of non-irrigated inclusions.

ESPAM1.1 RED vs. Proposed ESPAM2 RED. The ESPAM1.1 RED factor was 0.12, which is outside the confidence interval for all of the estimates here except for the 2006 placeholder data set. It is not known why this occurred, but there are at least two possibilities:

1. The data used in ESPAM1.1 were hand-drawn by IDWR personnel within organized canal companies and irrigation districts. The sample for the current effort was spread across the entire plain. The differences in values may reflect differences between lands within companies and districts, which tend to be older parcels, and more-recently developed lands outside companies and districts.

2. The IDWR representation was concerned with application of water for beneficial use, and the current effort was concerned with any locations where vegetation growth was enhanced by application of irrigation water. The former interpretation may have been more likely to exclude rock piles and rough areas that sprinklers pass over, which we explicitly included in the current effort.

Proposed Design Decision

Even when statistical uncertainty is large, it is still true that the estimated value is the most probable value. Therefore, it is proposed that ESPAM2 use the derived estimated values shown in Table 2, despite the lack of statistically-significant differences.

Table 2
Proposed RED Values for ESPAM2
Model Calibration

Data Set	Proposed RED Value
1980 RASA	0.07
1986 IDWR LANDSAT NDVI analysis	0.01
1992 IDWR/BOR photo analysis	0.06
2000 IDWR LANDSAT color-change analysis	0.05
2006 IDWR LANDSAT NDVI analysis (placeholder data set)	0.09
2006 IDWR final data set (when available)[9]	zero

[9] We assume that the detailed hand refinement performed by IDWR will result in a data set that incorporates all relevant non-irrigated features within irrigated parcels, so that no reduction will be required.