Insect Egg Counting on Mass Rearing Oviposition Pads by Image Analysis

T. C. Pearson, R. H. Edwards, A. P. Mossman, D. F. Wood, P. C. Yu, E. L. Miller

Abstract. *Variability of egg quantities laid on rearing trays during large–scale production of sterile insects can cause economic losses due to overfeeding or underfeeding. In addition, quality control measures to monitor egg numbers are both tedious and laborious. Methods were developed to count pink bollworm egg numbers on oviposition pads using histogram features from 10–bit grayscale images of the pads. Egg count prediction by regression analysis produced highly significant predictive model equations when individual histogram bin values or cumulative histogram bins (multiple thresholds) were used as independent variables. Models using cumulative histogram bin values appeared more robust as the mean squared error values were slightly lower. There was little difference in prediction results when using camera resolutions of 183, 139, and 94 pixels/cm. Automation of this methodology may allow a mass rearing facility to obtain oviposition pad segments with the desired number of eggs ±10%, increasing insect yield and/or quality and minimizing diet costs. In addition, the methodology will allow entomologists to quickly and accurately count eggs from laboratory or large–scale experiments.*

Keywords. *Image analysis, Egg counting, Pink bollworm, Mass rearing.*

Accurate estimation of egg numbers on oviposition pads is important during production of sterile insects for biocontrol sterile release programs, such as with pink bollworm, or during research and development of rearing techniques where egg, larvae, or pupae counts need to be ascertained. During commercial production of sterile pink bollworm insects, annular egg laden pads are collected from the oviposition cages daily for nine days. The first collection begins after the first 24 hours. Eggs from collection days 2 through 8 are considered first quality eggs. Eggs from days 1 and 9 are second quality and are used only if there are insufficient numbers of first quality eggs. In general, the annular pads are radially cut into eight equal sections of 45° each although manual corrections are sometimes made for pads which appear to have an extra heavy or light load of eggs. In these cases, the pads are cut into fewer or more sections, as appropriate. Thus, there is limited control of egg numbers per rearing tray, despite the fact that the optimum number per pad appears to be approximately 5000 ± 500 eggs (Miller, 1999). Production costs could be minimized if the optimum number of eggs were applied to each rearing container. If the number of eggs is lower than the optimum, then diet is being wasted. If more than the optimum number is used, quantity is being swapped for quality.

Article was submitted for review in April 2001; approved for publication by the Biological Engineering Division of ASAE in October 2001.

This project was supported in part by funding from the Cotton Pest Control Board of the California Cotton Growers Association. Mention of a trade name, proprietary product, or specified equipment does not constitute a guarantee or warranty by the USDA and does not imply approval of a product to the exclusion of others that may be suitable.

The authors are **Tom C. Pearson, ASAE Member Engineer**, Lead Scientist and Research Agricultural Engineer, **Richard H. Edwards**, General Engineer, **Albert P. Mossman**, Chemist, **Delilah F. Wood** (formerly published as Delilah W. Irving), Botanist, **Pauline C. Yu**, Technician, Microscopy & Imaging Lab, Western Regional Research Center, USDA–ARS, Albany, California, (currently Graduate Student, University of Southern California); and **Ernie L. Miller**, Research Entomologist, USDA–APHIS, Phoenix Methods Development Center, Phoenix, Arizona. **Corresponding author**: Tom C. Pearson, Grain Marketing and Production Research Center, 1515 College Ave., Manhattan, KS 66502; phone: 785–776–2729; fax: 785–539–5550, e–mail: tpearson@gmprc.ksu.edu.

Review of Literature

Image processing techniques are being used more and more frequently to count items, orient pieces, or discriminate between objects with different visual characteristics. Most automated image systems perform counting by segmenting the item to be counted from the background by applying a threshold based on the pixel intensity and/or intensity slope (or rate of change in intensities). Using this methodology, automated imaging systems have been developed to count white spots on dyed fabric (Han, 1998), to count specks on semolina (Symons, 1996), and to count specks of bran in flour (Kim, 1999). In addition, there are several commercial software programs that use this methodology to count objects given a digitized image, for example, Optimas (Media Cybernetics, Silver Spring, Md). These object counting systems can work well if the objects are only one layer thick and have good contrast from the background. At present, no work has been reported to count insect eggs on oviposition pads, perhaps because eggs tend to be layered two or three thick. But, a rapid and non–destructive method for counting eggs on oviposition pads might be very useful in entomology research laboratories and for quality control purposes.

The objective of this article was to determine the feasibility of using image processing techniques to count the number of pink bollworm eggs on an oviposition pad.

Experimental Procedures

Imaging System

Images of 30° pad samples were taken with a 10–bit grayscale CCD progressive scan camera, 1300 × 1030 pixels (TM1300, Pulnix, Sunnyvale, Calif.) using a zoom video lens (Zoom 7000,18–108 mm, f2.5, Navitar, Rochester, N.Y.), and were captured on a frame grabber board (IC–4DIG–16D, Imaging Technology, Bedford, Mass.). The images were imported into an image analysis program (Optimas, v 6.5, Media Cybernetics, Silver Spring, Md.) running on a personal computer (PC).

A schematic of the imaging system is shown in figure 1. The camera was mounted on the arm of a photographic enlarger in the center of a dual circular fluorescent light fixture (Mdl C–2072, American Fluorescent, Waukegan, Ill.). The outer and inner bulbs were 40 and 32 watts, respectively, and the vertical distances from the samples to the camera lens and the light bulbs were 50 and 43 cm, respectively. The pad segments were placed on a 22.9–cm diameter aluminum plate sanded with 100 grit paper. The aluminum surface provided a near white reflection to the camera. All segments were positioned radially to the center. Moving the zoom lens to specific positions marked on the lens and housing changed image resolution. Images from each sample were taken at resolutions of 183, 139, and 94 pixels/cm. Images of 30° pads were used so the entire segment could be captured at all three camera resolutions. Once the image was focused, the focusing ring was fixed in place for all experiments. All images were saved for off–line analysis.

Egg Oviposition Pads

Current mass rearing pink bollworm oviposition pads consist of annular rings, 10.2 cm ID by 20.4 cm OD die cut from brown paper toweling (No. 6086K, Tagsons Papers Inc., Mechanicville, N.Y.). Two changes to the oviposition pads were made for these experiments. First, the facility pads had a nonuniform color, due to paper fibers of various shades of brown running throughout, which would complicate image processing required to count eggs. Consequently, a white paper toweling (No. 1040, Kimberly Clark, Roswell, Ga.) was substituted to reduce this problem during these concept–testing runs. Tests with the white paper prior to its experimental use showed that fecundity (no. of eggs laid per unit time) was actually higher by 16.8% than using the brown paper (Miller, 1999). The second problem with the brown oviposition papers at the facility was the frequent presence of eggs on the underside of the pad near the inner and outer edges. The brown oviposition papers deflected upward allowing insects to crawl to the underside of the paper and lay eggs even though the paper was held down by the weight of a silicone ring. It was found that use of a slightly smaller diameter silicone ring held the paper flat and prevented eggs from being laid on the underside of the pad.

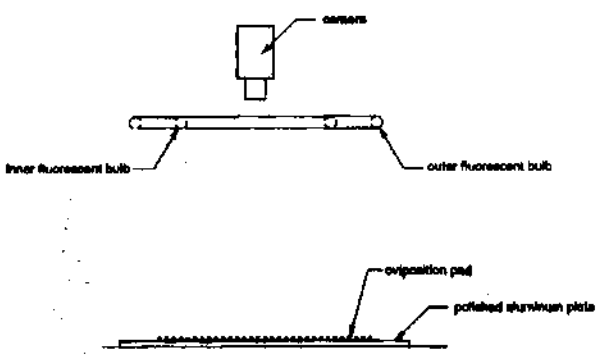

Figure 1. Schematic of imaging system.

Oviposition Pad Processing

Once a week, three white pads were placed in mass rearing oviposition cages at the USDA–APHIS Phoenix facility containing moths that had been laying eggs for 0 to 8 days. Egg–laden pads were removed daily after the normal time in the cages, as previously described. Timing was such that each batch of pads, designated day 1 to day 9, was removed on a Monday. Each pad was placed in a separate plastic bag and the lot was air shipped in an insulated box to Albany, California, arriving the following morning. One pad from each oviposition day was randomly selected, cut into four or more 30° segments, and two of these cut segments were imaged as previously discussed. When placed on the aluminum plate, a margin of about 6 mm of the plate around the outside edge was included in each image. This procedure was repeated weekly for eight weeks, commencing 26 April 1999, and resulted in a total of 144 images of 30° pad segments.

Actual Egg Count Determinations

The number of eggs on a pad segment was determined by one of two methods. In the first, a 472 pixels/cm color scan (Mdl. 6200Cse, Hewlett Packard, Greely, Colo.) of the segment was enlarged to an image approximately 60 cm square at a resolution of 49 pixels/cm using Adobe Photo Deluxe (Business Edition 1.0, Adobe Systems, San Jose, Calif.) and printed using a large format color printer (450C, Hewlett Packard, Greely, Colo.) on 91–cm wide opaque bond paper. The number of eggs in a cluster or area could be hand counted from the print after referring to the high–resolution scan of that area on the computer screen. This method was useful for pad segments with fewer than 2000 eggs without multiple layers of eggs.

The second counting method used a chlorine–containing compound to remove the eggs from the pad and an image analysis program to count the number of eggs present. A pad segment was gently stirred with a magnetic stirring bar in 200 mL of a 5–wt% solution of sodium dichloro–s–triazinetrione dihydrate (stabilized spa chlorine granules). Spa chlorine granules were used instead of bleach because the granules did not disintegrate the white paper toweling as aggressively as sodium hypochlorite. After approximately 30 min, the eggs were free from the pad, and the pad was removed. Eggs still adhering to the pad after removal from the solution were counted by hand under a binocular microscope. An additional 30 min in the solution separated all eggs from one another. The eggs were recovered by pouring the mixture through a 7.6–cm diameter, 60–mesh screen and rinsing the eggs from the screen into a 30–mL beaker with about 10 g of water from a squeeze bottle. The eggs were stained in aqueous 0.125% (wt/vol) methylene blue for 2 hours. Using this procedure, the eggs were dyed blue but the white color of any cellulose fiber carried along

with the eggs was unchanged. The eggs then were screened, washed with distilled water, and transferred to a flat glass–bottomed tray, 22 cm wide by 108 cm long with 12–mm aluminum sides, which covered the surface of the color scanner. The eggs were covered with 3 mm of water and spread evenly over the center of the tray by hand using a thin wire pick with the objective of separating eggs touching one another. The plate containing the blue dyed eggs was scanned in 36–bit color and imported into Optimas for counting. The image was converted to 10–bit grayscale and thresholded to discern the dyed eggs from the background. Single eggs were found to have minimum perimeters of 0.7–0.8 mm and minimum areas of 0.040 mm^2. Particles with areas greater than 0.140 mm^2 were considered to be doublets — two eggs touching one another. The criteria for single or doublet eggs were entered into Optimas. The final egg count was the number of single eggs in the scan plus two times the number of doublets plus the number of hand counted eggs that had adhered to the oviposition paper during removal from the chlorine solution.

Six sample segments were run by both of the above methods to confirm that the results were equivalent. Each pad contained 500 to 2000 eggs. The comparisons resulted in egg counts obtained by hand and by counting with image analysis after spa chlorine removal, respectively, of 508/506, 806/799, 1676/1652, 1742/1737, 1751/1746, and 1929/1951. The two methods produced results within 1.4% or less of each other.

Prediction of Egg Counts by Image Analysis Techniques

Figure 2 shows a representative set of day 1 through 9 pad segments. It is easily seen that egg laying does not follow a bell–shaped curve with days of oviposition. Rather, the segments from any oviposition day can contain high or low numbers of eggs, and the eggs are not evenly distributed on the pads. When eggs were first laid on the pad, they were in groups of one to four eggs, fairly evenly distributed over the surface of the pad. When less open space was available,

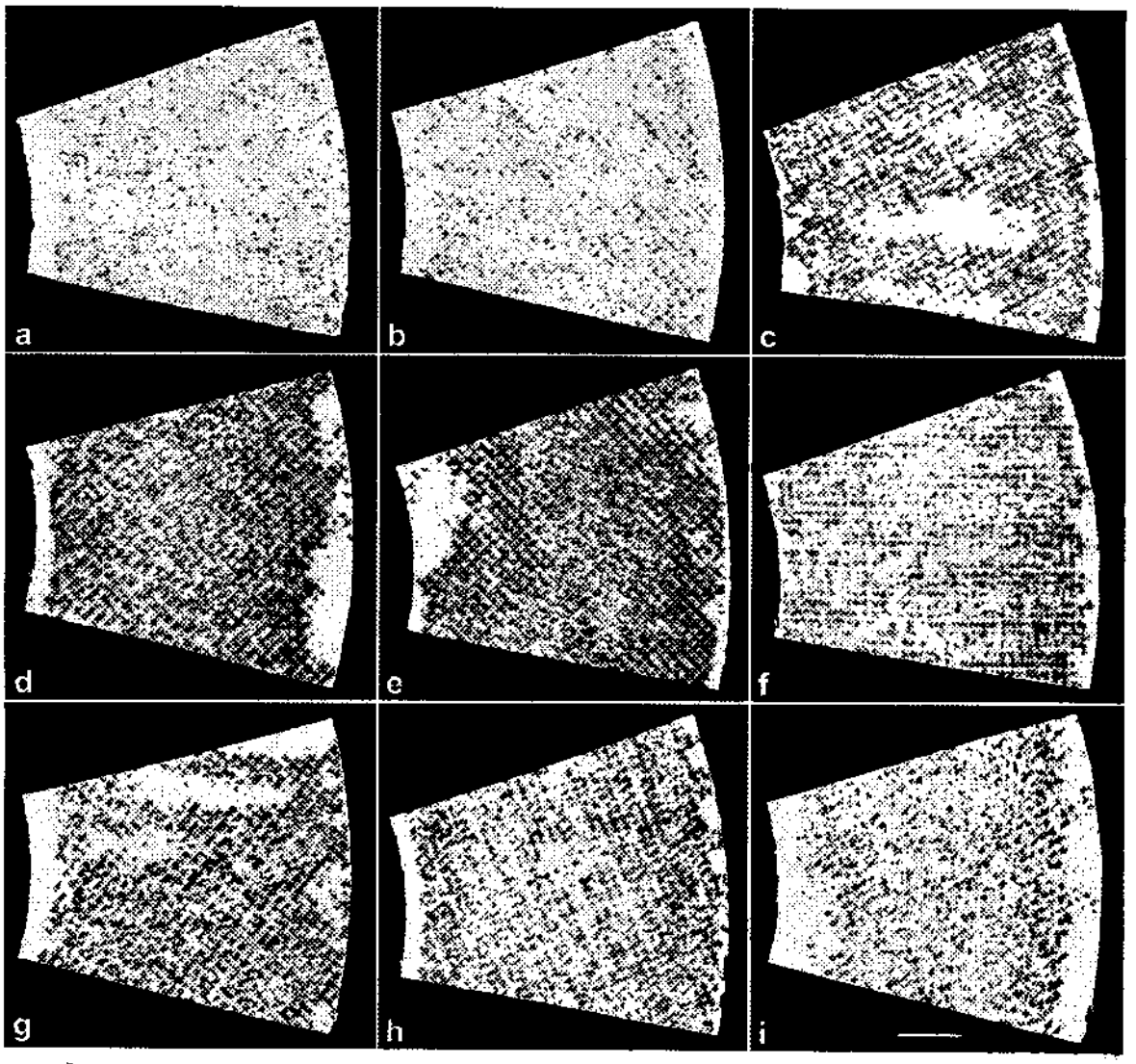

Figure 2. Typical white oviposition pad segments from days 1 through 9 (a–i, respectively). Magnification bar (in i) = 1 cm.

eggs tended to be laid in trails along ridges and valleys in the pad with many eggs touching each other (fig. 3a). As laying proceeded, trails intersected, forming a lattice–like network (fig. 3b) which could be very large. Eventually, masses were formed which were more than one egg thick (fig. 3c). Thick egg masses appeared darker in the imaged pad than single thickness egg patterns. In addition, the darkness of the eggs depended on their orientation since egg shape is typically ellipsoid.

To overcome the problem of egg orientation and layering, gray level processing rather than binary blob analysis was performed to estimate the number of eggs on a pad. The method for counting eggs on the imaged pads was a statistical approach using stepwise linear regression (SAS, 1987; 1991). The independent variables consisted of 1024 gray level histogram bins into which the image was divided. Pixel counts from all 1024 bins from each image were used in the process. The stepwise procedure was used to select only a limited number of variables to be used in the prediction model using a significance of 0.001 to enter or exit from the model.

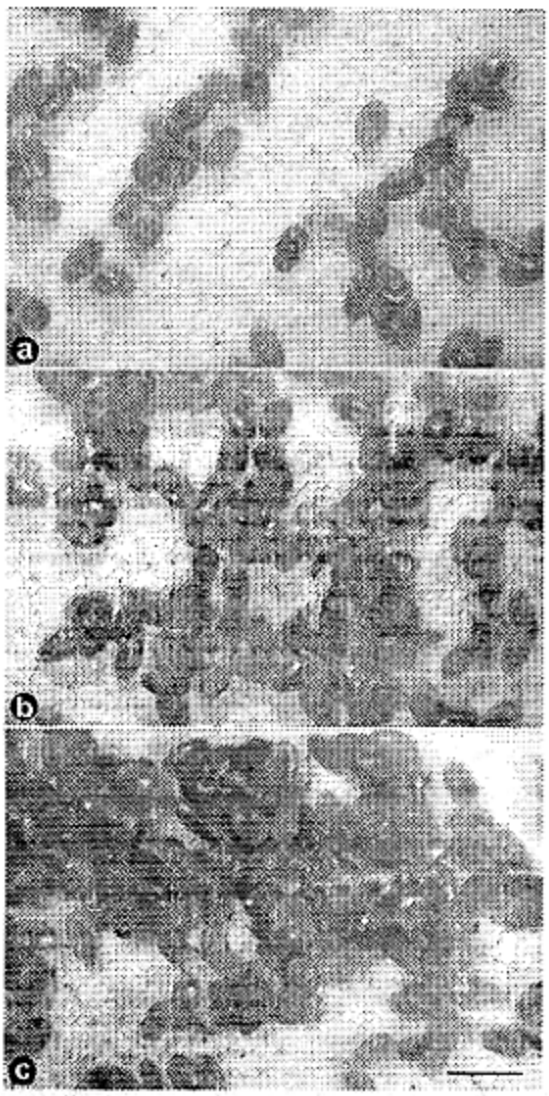

Figure 3. Typical range of egg patterns on oviposition pads as egg density increases. Trails of eggs and single eggs (a); trails of larger areas producing a lattice–like arrangement (b); and larger masses of eggs containing multiple layers (c). Magnification bar (in c) = 1 mm.

A file was created for the data from each of the three camera resolutions giving collection date, sample number, actual number of counted eggs, and pixel counts in each of the 1024 bins. Two types of histograms, the bin and cumulative, were used. With the bin histogram, a single bin value was the number of pixels having a gray level intensity corresponding to the histogram bin value. So, each bin consisted of pixels comprising an individual grayscale value. In contrast, cumulative histogram bin values were the sum of all pixels having gray level intensities from 0 to the gray level intensity corresponding to the histogram bin value. For example, the cumulative histogram value for bin 654 would be the sum of all pixels having gray level intensities from 0 to 654. Using cumulative histogram bins is equivalent to setting an image intensity threshold, or using multiple thresholds if there were more significant independent variables in the model.

Variable selection for both the bin and cumulative histogram data files was performed using half the images, often called the training set, and the prediction accuracy was validated with the other half (the validation set). Thus, data from 72 images comprised the training set, and data from 72 separate images comprised the validation set. Images were randomly assigned to the training and validation sets. Segments with egg counts below 500 were not included in the regression analyses.

Results and Discussion

Actual Egg Counts on Oviposition Pad Segments

The number of eggs on a pad segment varied greatly. The number of eggs found on each segment, along with the egg collection day, are shown in table 1. The eggs on pads from individual days ranged from a low of 7 for day 1 to a high of 6252 for day 3. Standard deviations by day were high in all cases. The 5000 ± 10% eggs that were considered desirable from a 45° pad segment used in the mass rearing facility would translate to a mean of 3333 and a range of 3000–3667 eggs for the 30° pad segments in these studies. A second set, based on a criteria of 3333 ± 15%, is also shown in table 1. If one assumes the egg counts were higher by 16.8%, because

Table 1. Results of egg counts from the 30° egg pad segments.[a]

	Egg Counts from Pad Segments			No. of Pad Segments With Egg Count 3333 ± 10%		With Egg Count 3333 ± 15%	
Day No.	Mean	S.E.	Range	As Imaged	Fecundity Corrected[b]	As Imaged	Fecundity Corrected[b]
1	752	684	7–1664	0	0	0	0
2	2090	1355	630–5053	1	1	2	2
3	4101	1271	1983–6252	3	4	4	0
4[c]	4512	877	3285–5829	2	0	5	0
5[c]	3885	1186	1440–5892	3	0	3	1
6[c]	3425	689	2487–4557	5	2	7	7
7[c]	2876	710	1767–4087	4	6	4	7
8	1894	863	767–3462	2	2	2	3
9	1539	674	28–2460	0	0	0	1

[a] 16 samples.
[b] Egg count divided by 1.168.
[c] 14 samples.

of the fecundity increase on 1040 white paper, dividing the white paper counts by 1.168 could be used to approximate what the number of eggs would have been on brown facility paper. These results indicate that automated egg counting on oviposition pads might increase insect quality and/or reduce diet costs.

Prediction of Egg Counts by Stepwise Regression Analysis

Results from the regression analyses are shown in table 2. All models, using either bin or cumulative histograms, at all three camera resolutions produced similar results. All models had R^2 values greater than 0.993 and all equations were significant at $p > F$ at the 0.0001 level. The cumulative histograms might be more robust because they rely on the sum of many pixel intensities rather than the pixel counts at just a few individual intensity levels, which could be affected by changes in lighting intensity and other external factors. The RMSE's for the cumulative histogram validation sets were lower and more constant than for the bin histograms. The results were essentially the same at all camera resolutions indicating that equipment for the lowest resolution could be used to lower camera system costs and decrease program running time. A graph of the predicted and actual egg counts using cumulative histogram bins to predict egg counts at the lowest resolution (94 pixels/cm) is shown in figure 4. The most significant range for a 30° segment requiring 3333 eggs per segment is approximately 2800 to 3800 eggs. Within this range, all points fell within the 90% confidence interval for the prediction equation and thus met the criteria of controlling egg counts at 3333 ± 10%. Egg dimensions were approximately 0.6 × 0.3 mm, so each egg was represented by approximately 15 pixels at 94 pixel/cm resolution. It is unlikely that much lower resolutions than 94 pixels/cm would be useful since it is important to quantify grayscale values of eggs that might be partially covered by another egg.

Actual versus predicted results were slightly more scattered at segment counts of 5000 or more which probably contained larger numbers of multilayer eggs. As can be seen in figure 4, one point with 5027 actual eggs had a predicted count of 6397. This point appears to be an outlier as the predicted egg count is more than the actual egg count by over three times the RMSE (603 eggs). This particular pad had an unusual amount of eggs layered on top of each other. However, with a lower target area of 3333 eggs per cut segment, there should be less error.

Figure 5 displays bin histograms from representative oviposition pad images containing 802, 2650, and 4265 eggs, respectively. As expected, the histogram corresponding to a pad with a low number of eggs shows less low grayscale values and more high values that represent the white pad.

Table 2. Predicted egg counts by stepwise regression analysis.

Histogram Type (bin/cum[b])	Camera Resolution (pixels/cm)	Predictive Regression Equation (Coefficient)	(Bin No.)	(Partial R^2)	Complete Model[a] (R^2)	($p > F$)	Validation Set (RMSE)[c]	(R^2)
bin	183	1.297×10^2	intercept	———	0.997	0.0001	226	0.995
		1.650×10^0	533	0.959				
		4.278×10^0	381	0.024				
		1.928×10^0	622	0.009				
		9.193×10^{-1}	472	0.003				
		-1.250×10^{-1}	760	0.001				
bin	139	1.021×10^2	intercept	———	0.997	0.0001	390	0.995
		3.057×10^0	546	0.968				
		2.866×10^0	655	0.011				
		1.428×10^1	394	0.012				
		-1.932×10^0	717	0.004				
		1.740×10^0	622	0.002				
bin	94	3.599×10^2	intercept	———	0.994	0.0001	239	0.992
		3.016×10^0	579	0.965				
		3.425×10^0	673	0.013				
		5.727×10^0	485	0.008				
		-7.575×10^{-1}	829	0.004				
		4.826×10^0	622	0.004				
cum	183	1.347×10^2	intercept	———				
		2.698×10^{-2}	677	0.992	0.996	0.0001	206	0.994
		-2.083×10^{-3}	808	0.003				
		-7.258×10^{-3}	578	0.001				
cum	139	1.648×10^2	intercept	———	0.995	0.0001	207	0.993
		1.361×10^{-1}	693	0.984				
		-2.901×10^{-2}	562	0.008				
		-7.982×10^{-2}	733	0.003				
cum	94	3.952×10^2	intercept	———	0.993	0.0001	201	0.992
		1.073×10^{-1}	718	0.978				
		-1.558×10^{-2}	847	0.009				
		-4.241×10^{-2}	596	0.006				

[a] Of training set.
[b] Cumulative.
[c] Root mean squared error of validation set.

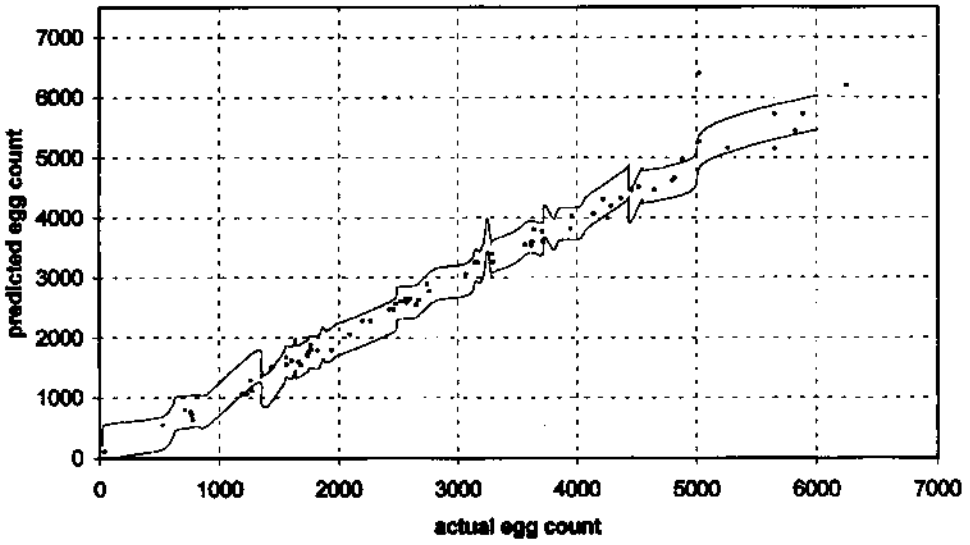

Figure 4. Predicted egg counts from the model (cumulative histogram) vs. actual egg counts from the validation set of samples at a camera resolution of 94 pixels/in. The lines indicate the upper and lower 95% confidence intervals.

As the number of eggs increases, the number of low grayscale values increases. The selected values from the regression procedure reflect this as well.

For this study, images were acquired over an 8–week period; the lengthy duration of time may cause some concern that the light output power may change during over time. However, examination of the grayscale pixel values in a 1–cm square of the aluminum base plate, present in all images acquired for this study, does not indicate that the light source output power changed over the course of the study. All pixel values in all images ranged from 1008 to 1013, and all of the 1 cm squares had a mean grayscale value of 1010. Nevertheless, better results over a long time period might be obtained by using a more stable light source since the egg count predictions do rely on grayscale values of image pixels.

A total of 144 images were used to develop and report the calibration equations used for this study. Another system using a different light source and/or different camera would need it's own calibration. When using stepwise regression selection of variables from highly correlated data, it is recommended that the training set comprise at least 10 samples for each factor chosen (Hruschka, 1987). The regression models using cumulative histogram bins contain four factors so about 40 images for the training set would be necessary for a new calibration. The validation set could be less than 40 to check that the training was not over–fitted, but more study would be needed to confirm just how many samples would be necessary. To recalibrate, without selecting new histograms, fewer than 30 images should be required, and there should not be a need for a validation set to check for overfitting.

Conclusions

Image processing methods were developed to accurately count the number of pink bollworm eggs on oviposition pads. This technique may be useful for an online system to segment oviposition pads with optimal egg counts or as a quality control technique where rapid and non–destructive counts are required. Egg count prediction equations were derived by stepwise regression techniques which could rapidly and accurately count the number of eggs on a pad segment based on the pixel counts in just three to five bins or cumulative histogram bins. Equations using cumulative histograms were inherently more robust and produced lower RMSE's using the validation sets at all three (low, medium, and high) camera resolutions. Therefore, in practice, the lowest resolution (94 pixels/cm) could be used, reducing costs for imaging equipment. In the current commercial scale system for mass rearing of sterile pink bollworm moths, only a small number of egg–laden pad segments contains the ideal number of eggs. The initial premise of using a progressive scan CCD camera to count eggs on oviposition pad segments appears to be workable. The counting techniques may be applicable to counting insects in other stages of growth.

Acknowledgements

Appreciation is extended to Dr. Fred Stewart, Director, APHIS–PPQ–PBRF, Phoenix, Arizona for his continued support, and to the Cotton Pest Control Board, California Cotton Growers Association, for their financial support. The authors are indebted to Ms. Anna Lowe and Mr. Jose Gomez for the endless hours preparing and counting eggs on oviposition papers.

References

Han, Y. J. 1998. White speck detection on dyed fabric using image analysis. *J. Cotton Sci.* 2(2): 91–99.

Hruschka, W. R. 1987. Data analysis: wavelength selection methods. In *Near–Infrared Technology in the Agricultural and Food Industries*, eds. P. Williams and K. Norris. St. Paul, Minn.: American Association of Cereal Chemists, Inc.

Kim, Y. S. 1999. Determination of bran in contamination in wheat flours using ash content, color, and bran speck counts. *Cereal Chemistry* 76(6): 957–961.

Miller, E. 1999. Personal communication, USDA–APHIS, Phoenix, Ariz.

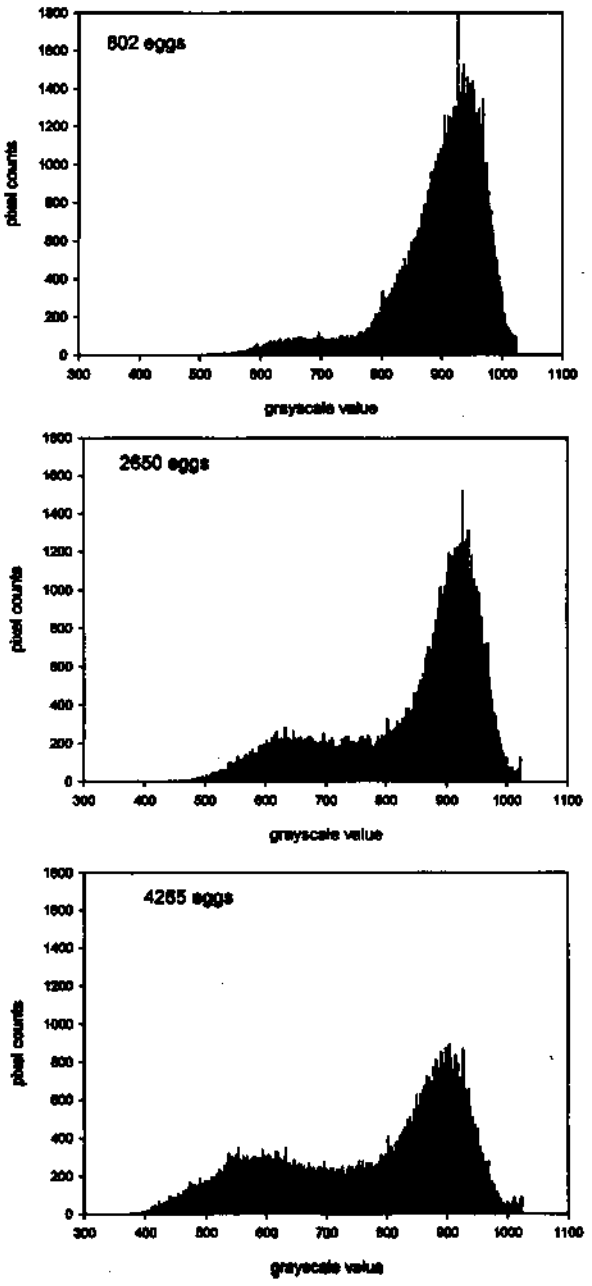

Figure 5. Example histograms from oviposition images with 802 eggs (top), 2650 eggs (middle), and 4265 eggs (bottom).

SAS Institute Inc. 1987. *SAS/STAT Guide for Personal Computers*, Ver. 6. Cary, N.C.: SAS Inst. Inc.

SAS Institute Inc. 1991. *SAS System for Regression*, 2nd Ed. Cary N.C.: SAS Inst., Inc.

Symons, S. J. 1996. Semolina speck counting using automated imaging system. *Cereal Chemistry* 73(5): 561–566.