American Society of
Agricultural and Biological Engineers

An ASABE Meeting Presentation

Paper Number: 061192

UIUL Number: 2006-7002

Development of a Yield Monitor for Citrus Fruits

Tony Grift PhD, Assistant professor, Dept of Agricultural & Biological Engineering, University of Illinois. 1304 W. Pennsylvania Ave. Urbana, IL 61801, grift@uiuc.edu

Reza Ehsani, PhD, Assistant professor, Citrus Research & Education Center, Lake Alfred, FL, ehsani@crec.ifas.ufl.edu

Kentaro Nishiwaki, Senior Researcher, National Agricultural Research Center, Japan, nisiwaki@affrc.go.jp

Catherine Crespi, PhD. Postdoctoral researcher, Dept of Biomathematics and Biostatistics, University of California at Los Angeles, ccrespi@ucla.edu

Min Min, PhD, Postdoctoral researcher, Citrus Research & Education Center, Lake Alfred, FL, minmin@ufl.edu

Written for presentation at the
2006 ASABE Annual International Meeting
Sponsored by ASABE
Portland Convention Center
Portland, Oregon
9 - 12 July 2006

Abstract

A method to measure the per-tree yield of citrus fruits was developed based on three approaches. Initially, the problem was viewed as a variation on the theme of particle counting as performed earlier on flows of fertilizer particles. The main difference was that 1) citrus fruits are much larger than fertilizer particles 2) they travel at a significantly lower speeds an 3) their flow density may be higher than that of fertilizer.

The earlier research for fertilizer mass flows used a dedicated time of flight device that employed optical magnification to measure small particles with relatively large sensors. A second advantage was that the image of the particles could be projected in focus on the sensor plane. In the case of citrus fruit counting there was no need for magnification, since citrus fruits are much larger than the sensors. However, for citrus fruits this led to defocus problems, when direct lighting was used.

The theory used in the fertilizer application was capable of dealing with multiple particles (clumps) interrupting the sensors simultaneously, which made the method non-intrusive and eliminated the need for particle singulation. Although this method was viewed to be transferable to the citrus counting problem, an easier and relatively low-tech method was assumed to be more reliable. A laser beam was projected just above the surface of a plate on which the fruits roll down into a collection bin. Theoretically, if the laser beam is placed low enough, each fruit will be counted individually, even if they are clumped together. Tests revealed that this assumption does not hold under high-density flow regimes. Therefore, an improved gutter system was designed which effectively transforms the flow of fruits into different streams of fruits in single file formation. This method, combined with low-placed laser beams, allowed for accurate counting of the fruits in the laboratory. In future research, the system will be tested on a canopy shaker machine.

1 Introduction

Citrus fruits in Florida represent the state's most valuable agricultural product worth 742 Million dollar in 2005. To date a very high percentage of fruits are still harvested by hand, but mechanization is a growing effort. Early on, there were yield-monitoring attempts for manual harvesting, by counting and locating containers in the field (Whitney *et al*., 2001, Schueller *et al*., 1999). Annamalai *et al*., (2004) used a machine vision approach to obtain an estimate of the yield of citrus fruits per tree.

Tree canopy shakers are a promising method of automated citrus fruit harvesting, but they lack a reliable method of yield monitoring. Canopy shakers have a conveyor mechanism that transports the fruits from the collection curtains upward to a level where the fruits roll along a ramp into the collection bin. A simple approach to yield monitoring would now be to count every fruit that is transported into the collection bin. However, since many fruits passing per second, forming a mass flow, two methods could be employed. Firstly, the mass flow could be considered as a sequence of clumps and spacings, and a dual photo-interrupter arrangement could be used to measure the lengths of the clumps and spacings among them. This method is completely non-intrusive, but an algorithm is needed to estimate the total number of fruits from the clump and spacing data. This algorithm was developed in earlier research by Grift, (2003). Secondly, a method could be employed that takes advantage of the flow being constrained by a ramp which allows counting using a single laser beam method. When the flow density becomes very high, an alternative method is to use singulation of the flow into single file gutters, which allows reliable counting using low-tech and low-cost hardware. The objectives of this research were to investigate three methods of fruit counting being:

1. Dual photo interruption based mass flow measurement using clump and spacing lengths as inputs for a mass flow estimation algorithm
2. Single laser beam fruit counting taking advantage of the rolling ramp
3. Single laser beam fruit counting using gutter singulation

2 Summary of previous mass flow measurement research

Grift, (2003) developed a generic method to measure the flow of fertilizer particles to be applied in aerial fertilizer application. The most important assumption was that the flow can be regarded as an arrival process, similar to classical examples of Poisson driven flows from queuing theory, people arriving at a helpdesk, emails/telephone calls arriving etc. Figure 1

shows the flow of identical particles falling from a funnel (left side). The image shows that the density of the flow due to acceleration is higher at the funnel side (left) and secondly, that the flow is intermittent and consists of clumps of particles separated by spacings. The challenge is now to estimate the number of particles per time unit (mass flow) while only knowing the lengths of the clumps and spacings.

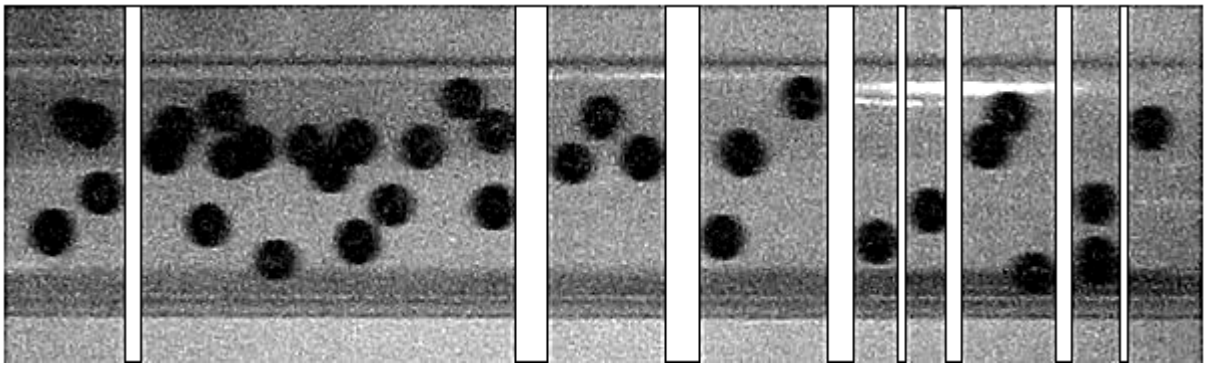

Figure 1. Flow of identical particles falling from a funnel (left side) showing clumps of particles separated by spacings.

The estimation of the total number of particles per time unit (mass flow) was based on the assumption that the locations of the particles in space and time were independent conforming to stationary (constant density) Poisson process, implying that the clumps and spacings among them are mutually independent. Although at the time this assumption was speculative, the results showed that it was possible to estimate the number of particles to within 1.5%, independent of the mass flow intensity. This led to the conclusion that the flow of particles in the fall tube was indeed near-Poisson driven and constitutes an elementary coverage process, also known as the 'simple linear Boolean model' process. Other applications of the simple linear Boolean model are found in particle and electronic counters (Hall, 1988; Takacs, 1962), diffusion of suspended particles (Bingham and Dunham, 1997), Markov/General/ ∞ queue (Kleinrock, 1975) and biomedical applications (Crespi, *et al*, 2005, Parthasarathy, 1997).

2.1 Flow rate estimation based on spacing lengths

According to a Poisson driven process assumption the lengths of the spacings must be exponential with a mean being the reciprocal value of the density. This method is seemingly ideal, since it merely needs a measurement of the spacing lengths in a time frame. Secondly, since the particle distribution only influences the clump lengths and since the clump and spacing lengths are independent, this method is valid for any particle distribution.

During experiments with identical particles, it was found that there was a consistent lack of short spacings and therefore any estimation based on spacing lengths such as the mass flow and the diameter estimate was inaccurate. In the estimation of flow density for instance, errors up to 25%, proportional to flow density were found.

2.2 Flow rate estimation based on clump lengths

Whereas the spacing distribution is simply exponential, the clump length distribution of particles is implicit and complex (Daley, 2001, Hall, 1988). Crespi and Lange (2006) developed methods to solve the equation for the clump lengths and used maximum likelihood estimation to

produce errors on their estimates. They were successful in estimating the mass flow, however since their methods include the spacing lengths (which were found in accurate) only after removing 6% of the data.

Grift (2001), used simulated clump lengths generated from a Poisson driven flow process to develop a mass flow estimation formula as given in equation (A1), (Appendix A).

$$\frac{E(N_T)^2}{E(N_1)} = \lambda \qquad (1)$$

where:

N_T: Number of clump arrivals per time unit

N_1: Number of "Singles" (clumps with length in interval [*D*, 2*D*] where *D* is the mean particle diameter

Equation 1 is only valid for identical particles measured without error. For distributed particle flow rate, Grift (2003) proposed a method where the distributed particle clump lengths are approximated by simulated identical particle clump lengths. This method was termed the Simulated Identical Particle Approximation (SIPA) method. The justification of this method is the existence of a common point in the clump lengths for distributed as well as identical particles as shown in figure 2 for a simulated Poisson process with Gaussian distributed particles.

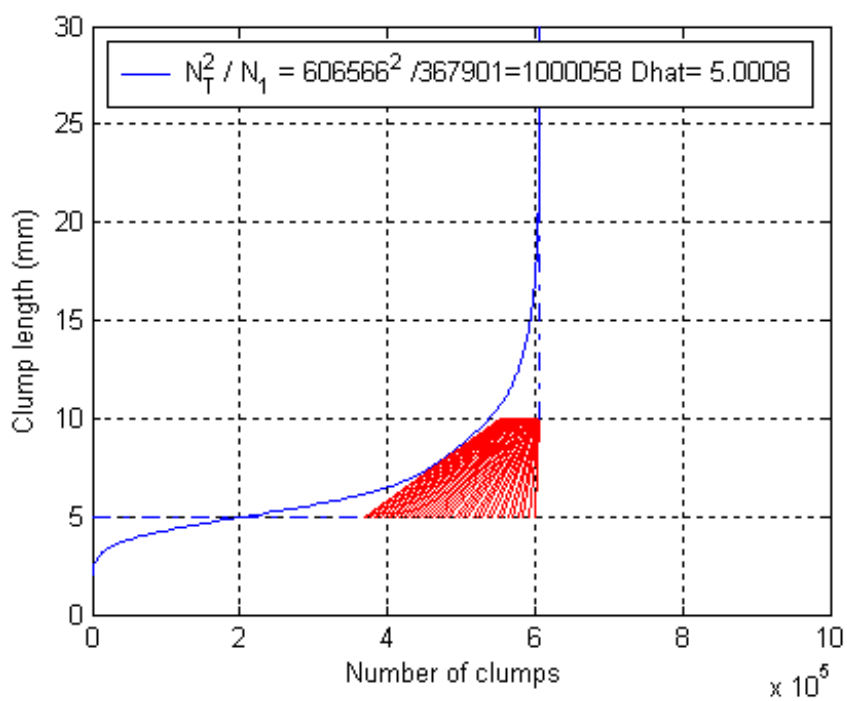

Figure 2. Simulated sorted clump lengths for distributed diameter particles (Gaussian, mean 5 mm, std 1 mm) and a flow density of 0.5. The straight lines indicate the Simulated Identical Particle Approximation (SIPA) method.

In the graph, the curved line represents the sorted clump lengths of Gaussian distributed particles (mean 5 mm, std 1 mm). The straight lines represent the Simulated Identical Particle clump lengths for increasing density (from right to left). When these straight lines intercept with the measured curved clump lengths, the number of Singles is obtained and equation 1 is used to

estimate the total number of particles. In this case, the total number of particles was 1 million and the estimate was $\frac{N_T{}^2}{N_0} = \frac{606{,}566^2}{367{,}901} = 1{,}000{,}058$. The diameter estimate was also accurate at 5.0008 for an input diameter of 5 mm.

Although the SIPA method worked for many simulated distributions (Gaussian, Uniform, Bi-modal, Multi-modal) there is currently no mathematical proof showing its validity for an arbitrary distribution. Estimation of the number of particles using the SIPA method in experiments led to an unbiased estimate (independent of the flow density) at an accuracy of 1.5%. It also implied that although in the flow the spacing lengths are not consistent with the Poisson assumption, the clump lengths are to a high extent. To date there is no valid scientific explanation for this phenomenon.

2.3 Estimation of the mean particle diameter

The mean diameter of the particles can be estimated using equation (2).

$$\hat{D} = \bar{S} \ln\left(\frac{\bar{C}}{\bar{S}} + 1\right) \tag{2}$$

where $\hat{D}$ is the estimator for the diameter, $\bar{C}$ is the mean of the clumps lengths and $\bar{S}$ is the mean of the spacing lengths. This equation does depend on both the clump lengths and spacing lengths, but is valid for particles with an arbitrary diameter distribution (for proof, see Appendix B). Since the spacing length distribution in experiments was unreliable, it was shown that the diameter estimate was inaccurate as well.

3 Materials and methods

3.1 Counting fruits using a time-of-flight device

The time-of-flight device clump length measurement principle is shown in figure 3. The time difference between the passings of the clump fronts between the two layers allows measurement of the velocity $v, (ms^{-1})$ by recording the time difference $\Delta t_f, (s)$ as in $v = \frac{b}{\Delta t_f}$ where $b, (m)$ is the distance between the light layers. The clump length $CL, (m)$ can be obtained from the total time in which the clump blocks either layer $\Delta t_p, (s)$ as: $CL = b\frac{\Delta t_p}{\Delta t_f}$. The spacing lengths were recorded as the time intervals among the clumps.

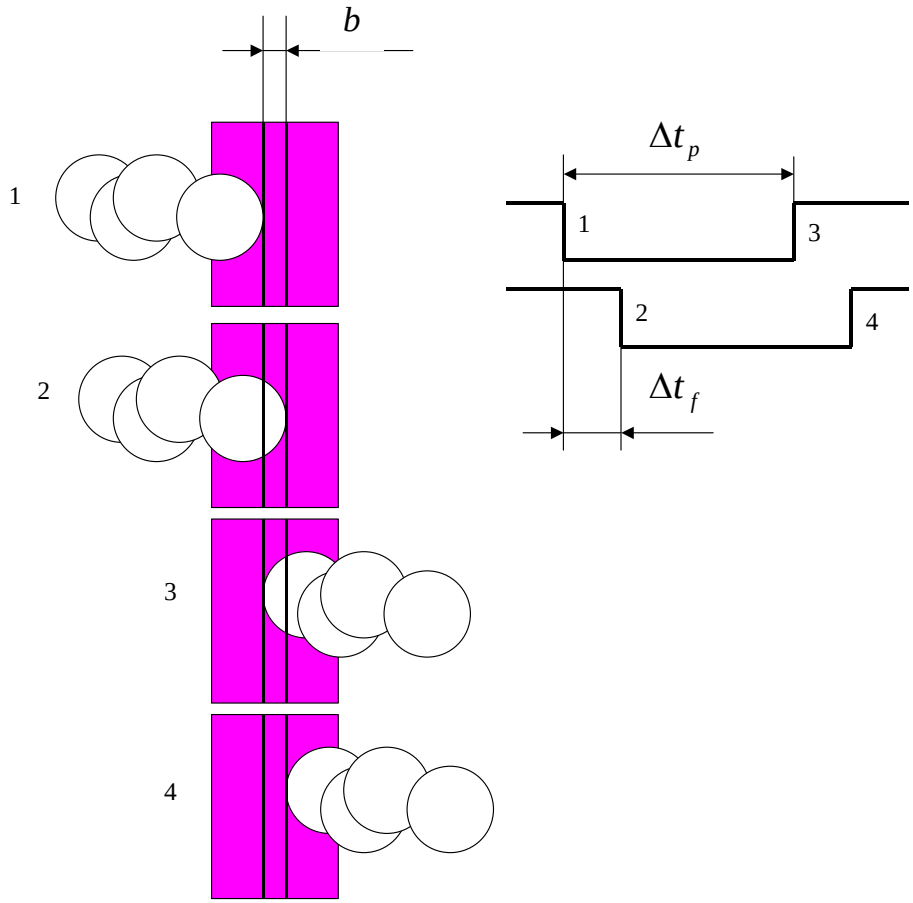

Figure 3. Time-of-flight clump length measurement principle with timing diagram.

To measure the clump and spacing lengths of flows of citrus fruits, a Large Sensor Array was constructed using a total of 240 optical switches (OptoSchmitt, Honeywell SDP8600). Figure 4 shows a portion of the Large Sensor Array, with a measurement width of approximately 60 cm.

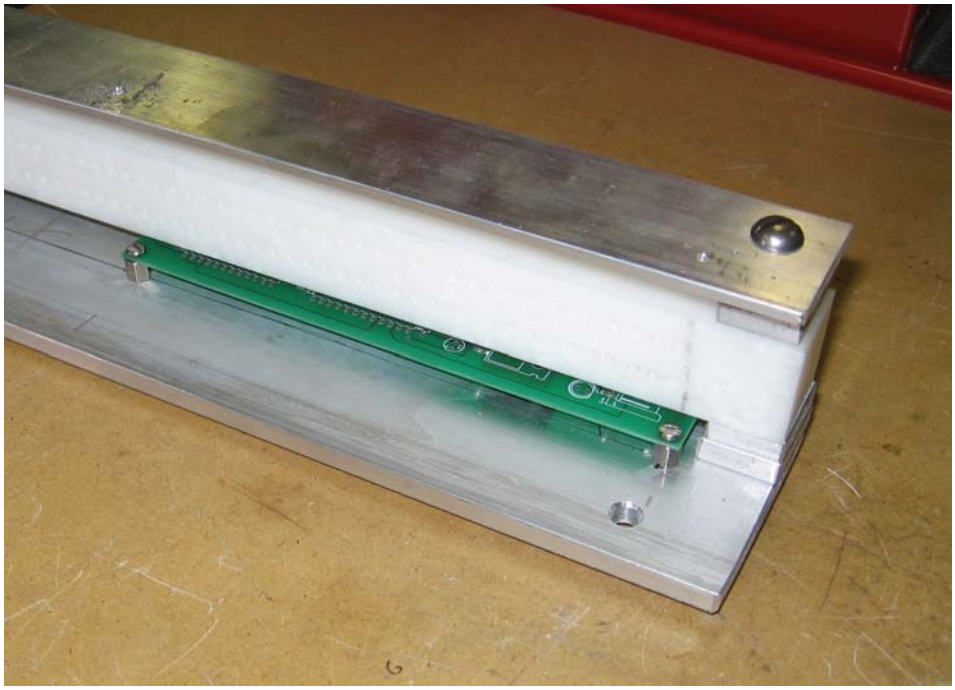

Figure 4. Portion of the Large Sensor Array, containing eight sensor boards with 30 OptoSchmitts each.

The width of the OptoSchmitts is approx. 5 mm, and for this reason in the case of fertilizer mass flows, it was necessary to optically magnify the image of clumps to accurately measure their lengths. This method had an added benefit in that the shadows that the clumps cast on the sensors were sharp due to the focusing mechanism of lenses. In the case of the much larger citrus fruits (10 cm average) there was no need to magnify the image of the clumps, and direct lighting became an option. A problem that emerged with direct lighting is defocus, which causes the shadows of the clumps to become 'soft' with gradual edges instead of square edges. This problem was alleviated in two ways. Firstly a bar was mounted in front of the sensors, which had small holes in it (1.5 mm) making the sensors directionally sensitive. Secondly, the light source was limited through a small grid to create a narrow slit (figure 5).

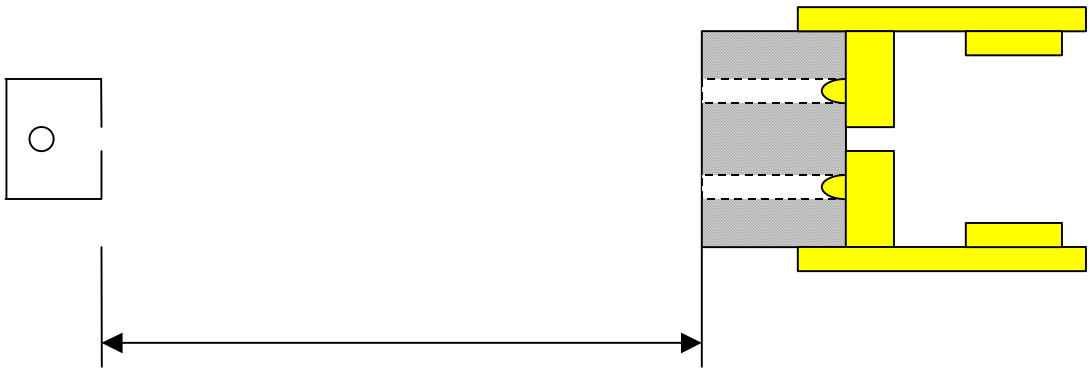

Figure 5. Dual receiver units mounted behind channeling tubes, excited by a single light source.

In experiments, the Large Sensor Array did not give reliable results due to defocus effects, and in future research an improved lighting system will be needed to make it perform properly.

3.2 Counting fruits using laser beams

After being harvested, the citrus fruits drop on collection skirts after which an elevator brings them to a height of approximately 4m. Here they drop on a slanted plate and they roll downward into the collection bin. Video clips of the mechanism show that some fruits bounce once but at the end of the plate they all roll. As an alternative to the bulky Large Sensor Array mounted under the edge of the plate, a dual laser beam interruption scheme could be employed which would allow measuring the lengths of the clumps and spacings among them, and the theory as developed for granular fertilizer particles would apply. However a second method was considered, which is simpler. Figure 6 shows an abstraction of fruits, spherical objects of varying diameter rolling down a plate. The side view (bottom) shows that even though the objects overlap, they are separated by a small gap among them, which allows counting them individually by projecting a small laser beam across the plate at the lowest possible location. Theoretically, since the objects have a point contact with the plate, there is always a separating gap close to the plate surface, since it is highly unlikely that two objects are exactly side-by-side. The laser beam is not required to have an infinitesimally small diameter, but it is helpful if the receiver is very light sensitive, meaning that even the smallest gap is detectable by a minute amount of light passing through. In reality the fruits are not exactly spherical, and when the density is high, there are many objects near side-by-side and some fruits in close proximity to each other may be interpreted as a single fruit.

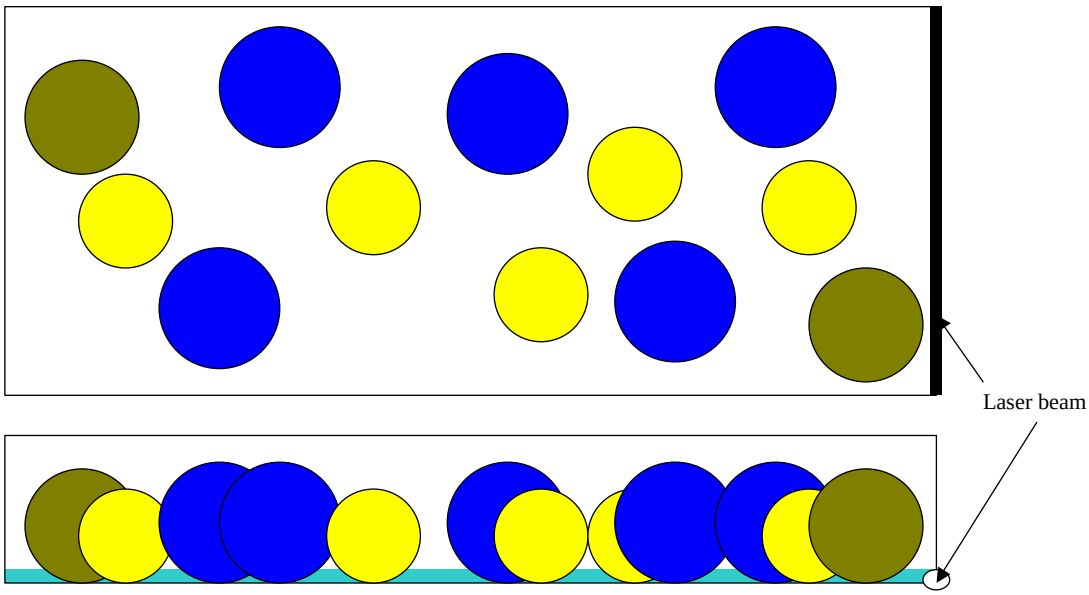

Figure 6. Top: Visualization of multiple spherical objects of varying diameter rolling down a slanted plate. Bottom: Side view of the objects showing a gap among each individual object.

For experiments, a ramp was built, which allowed dropping many fruits (here simulated using tennis balls). For the laser beams, low cost visible red laser beams (640 nm) were mounted along the side of the ramp as shown in figure 7, indicated by the straight lines. The angle of the ramp was 30 degrees with respect to a horizontal plane. Although in theory a single laser beam would suffice to count the fruits, two laser beams were used to experiment with detection at varying densities. The receivers used were standard 3 mm photo-transistors (Jameco, Part Nr.112176), which have peak sensitivity in the 940 infrared range, but due to their highly sensitive nature, work well as detectors for 640 nm laser light. For the counting hardware, microcontrollers (PIC18F458) were used and the data were output on a serial bus.

Figure 7. Slanted ramp used for counting experiments of clumped fruits.

Although the principle of counting clumped fruits while avoiding singulation may be elegant, a gutter system, which splits the flow into 5 or more separate single file flows as shown in figure 8, was considered more practical and reliable. Since no fruits are allowed to "overtake" each other, the width of the gutters should be smaller than twice the average fruit diameter. Overtaking is already unlikely since due to acceleration, the fruits are moving away from each other in the gutters.

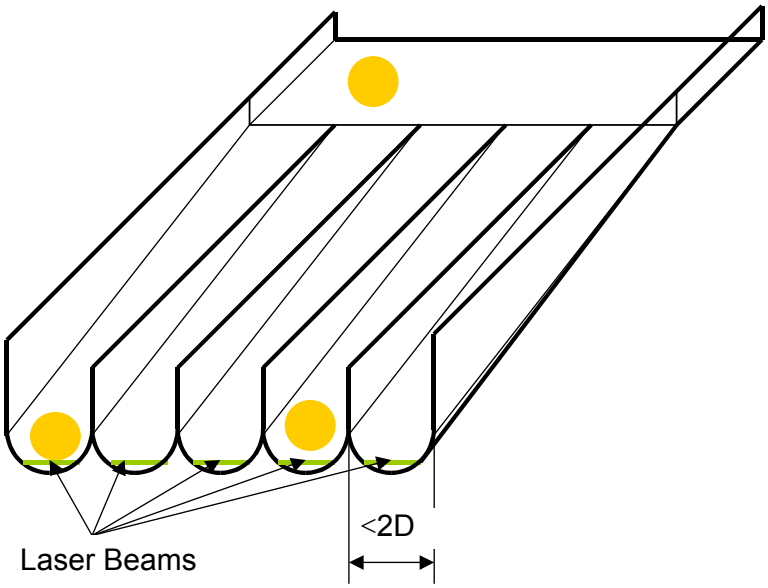

Figure 8. Singulation of citrus fruits using a gutter system. Each gutter contains a low placed laser beam, connected to a microcontroller, which counts fruits individually.

As before, the number of particles per gutter are measured using the same low laser beam arrangement as shown in figure 6, but here no overlapping fruits are expected. Figure 9 shows and experimental arrangement, where a single gutter is used and tennis balls are dropped into the tube to assess the reliability of the gutter based counting system.

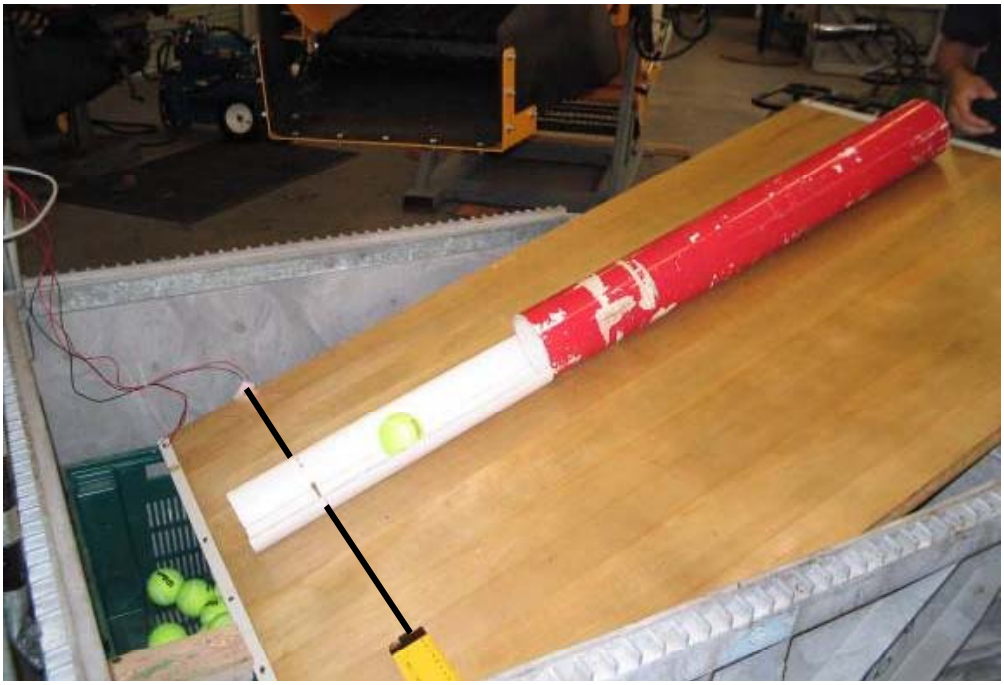

Figure 9. Experimental arrangement with single gutter to test the reliability of the counting method.

Eventually, the counting hardware will have to be in 5 fold and the data outputs synchronized. The counters used in this research have built-in Controller Area Network functionality, which simplifies the development of a multi-processor based counting system.

4 Results and discussion

The first method based on the Large Sensor Array showed to be inaccurate in measuring the lengths of clumps and spacings due to defocus problems. The second method based on the interruption of a low placed single laser beam showed promising, however under higher mass flow densities, many objects in close proximity were misinterpreted as single objects. A third method that forced particles in single file in gutters was deemed most reliable and robust for application on a tree canopy shaker machine.

5 Conclusions and further research

Three methods were investigated as candidates for a yield monitoring system for citrus fruits. The first method used a time-of-flight based photo interruption sensor similar to one that was used earlier to measure the mass flow of fertilizer particles. The method is non-intrusive and allows multiple particles to interrupt simultaneously (clumps separated by spacings) and a sophisticated algorithm was available to estimate the number of fruits from clump and spacing lengths. The sensor needed for fruits is much larger than the one for fertilizers particles and secondly, there was no mechanism to focus the shadow of the fruits on the sensor. This method was deemed not mature enough to perform an accurate mass flow measurement.

A second method relied on the fact that fruits can be assumed to roll over a surface while accelerating downward to a collection bin. Even though from a top view fruits may form clumps, when viewed from the side, just above the surface, there are always gaps among the fruits. This allowed using a single laser beam placed just above the surface on which the fruits roll to count them individually. Since this premise is only valid for an infinitesimally small laser beam, and perfectly spherical particles, in reality often two or more in close proximity fruits were falsely interpreted as one. A method to detect whether an individual fruit was present is to observe the total interruption times, which are larger for a clump of two or more fruits. This method requires a threshold to distinguish among individual and clumped fruits, which must be adaptive to accommodate for fruit diameter changes.

Both previous methods aimed to be non-intrusive which implies that the fruits are allowed to clump together, avoiding singulation. Although this approach is elegant, for a reliable citrus fruit counting system on a harvester, a third method was assumed more feasible which consists of singulating fruits into gutters. The width of the gutters must be such that it is not possible for two fruits to pass the sensor simultaneously, an event that is already unlikely since the fruits are accelerating in the slanted gutters. Each gutter contains a low placed laser beam connected to its own microcontroller that counts the number of fruits passing per time frame and reports its counts through a Controller Area Network bus.

Further research is needed in two areas. The lighting system employed in the Large Sensor Array did not allow for accurate counting of the fruits and needs to be revisited. Secondly, it needs to be compared to the gutter system to determine which method is most reliable by testing on a real harvester.

References

Annamalai, P., W.S. Lee, and T.F. Burks. 2004. Color Vision System for Estimating Citrus Yield in Real-time. ASAE Paper No. 043054. St. Joseph, Mich.:ASABE.

Bingham, N. H., and B. Dunham. 1997. Estimating diffusion coefficients from count data: Einstein-Smoluchowski theory revisited. *Annals of the Institute of Statistical Mathematics* (49): 667-679.

Florida Agricultural Statistics Service. 2005. Citrus summary 2004-05. Orlando, FL. Available at http://www.nass.usda.gov/Statistics_by_State/Florida/Publications/Citrus/cs/2004-05/cs0405all.pdf. Accessed May 9 2006.

Crespi, C. M., W.G. Cumberland, and S. Blower. 2005. A queueing model for chronic recurrent conditions under panel observation *Biometrics* (61):194-199.

Crespi, C. M., and K. Lange (2006). Estimation for the Simple Linear Boolean Model. *Submitted to Statistics and Probability Letters*.

Daley, D. J. 2001. The busy periods of the M/GI/∞ queue. *Queuing Syst*. (38), 195-204.

Grift, T. E., and J.W. Hofstee. 1997. Measurement of velocity and diameter of individual fertilizer particles by an optical method. *Journal of Agricultural Engineering Research*, 66(3), 235-238, doi:10.1006/jaer.1996.0128.

Grift, T.E. 2001. Mass Flow Measurement of Granular Materials in Aerial Application, Part 1: Simulation and Modeling. Transactions of the ASAE. 44(1):19-26.

Grift, T. E. 2003. Fundamental mass flow measurement of solid particles. *Particulate science and technology*, 21(2):177-193.

Hall, P. 1988. Introduction to the Theory of Coverage Processes (Wiley, New York, 1988).

Kleinrock, L. 1975. Queueing Systems Volume I: Theory, John Wiley and Sons, New York.

Parthasarathy, P. R. 1997. The effect of superinfection on the distribution of the infectious period a continued fraction approximation, *IMA Journal of Mathematics Applied in Medicine and Biology* (14):113-123.

Schueller, J.K., J.D. Whitney, T.A. Wheaton, W.M. Miller and A.E. Turner. 1999. Low-cost automatic yield mapping in hand-harvested citrus. *Computers and Electronics in Agriculture* 23(1999) 145-153.

Takacs, L. 1962. Introduction to the theory of queues. Oxford University Press, New York.

Taylor, H. M., and S. Karlin. 1998. An Introduction to Stochastic Modeling. Academic Press, San Diego, 3rd ed.

Whitney, J. D., Q. Ling, W. M. Miller, and T. A. Wheaton. 2001. A DGPS Yield Monitoring Systems for Florida Citrus. *Transactions of the ASAE.* 17(2): 115-119.

6 Appendix A

Proof flow rate estimation formula $\frac{E(N_T)^2}{E(N_1)} = \lambda$.

Consider a simple linear Boolean model with arrivals at rate λ and mean particle size *D*. Let N_i denote the number of clumps (on a unit interval) with *i* arrivals, and let $N_T = \sum_{i=1}^{\infty} N_i$

Proposition.

$$\frac{E(N_T)^2}{E(N_1)} = \lambda \qquad \text{(A1)}$$

Proof.

Let $\overline{S}$ and $\overline{C}$ denote the expected length of spacings and clumps, respectively. Then the expected length of one clump/spacing pair is $\overline{S}+\overline{C}$. Thus the expected number of clumps per unit time is:

$$E(N_T)=\frac{1}{\overline{S}+\overline{C}}=\frac{1}{\lambda^{-1}+\lambda^{-1}\left(e^{\lambda D}-1\right)}=\lambda e^{-\lambda D} \qquad \text{(A2)}$$

Let N denote the number of particles comprising a clump. Since particles arrivals are Poisson with intensity λ, the number of particles arriving during an interval of length D is distributed as Poisson with mean λD. Thus $P(N=1)=e^{-\lambda D}$, and thus $E(N_1)=E(N_T)P(N=1)=E(N_T)e^{-\lambda D}$ and the result follows.

7 Appendix B

Proof Diameter estimation formula $\hat{D}=\overline{S}\ln\left(\frac{\overline{C}}{\overline{S}}+1\right)$, (any distribution).

Consider a mass flow particle counter in which particles flow by a sensor, giving rise to a signal that is "on" as any particle passes the sensor and "off" otherwise. Denote the mean size of the particles by *D*, where *D* is a finite quantity (that is, *D* < ∞), and assume that particles pass the sensor according to a Poisson process. Denote the mean particle flow rate as λ, measured in particles per unit time. The passage times of particles may overlap; thus an "on" signal corresponds to a clump with an unknown number of particles. The "off" signals correspond to the spacings between clumps. Let $C_1, C_2, \ldots$ and $S_1, S_2, \ldots$ denote the lengths of successive clumps and spacings, respectively, and let and $\overline{C}$ and $\overline{S}$ denote the mean clump and spacing lengths.

Proposition. Unbiased estimates of the mean particle size *D* and the mean particle flow rate for any distribution of particle size with finite mean *D* may be obtained as:

$$\hat{\lambda}=1/\overline{S} \qquad \text{(B1)}$$

$$\hat{D}=\overline{S}\ln\left(\frac{\overline{C}}{\overline{S}}+1\right) \qquad \text{(B2)}$$

Proof. Due to the Poisson nature of the particle arrival process, the spacing lengths $S_1, S_2, \ldots$ are exponentially distributed with mean $E(S)=1/\lambda$ (Hall, 1988). Thus an unbiased estimate of the Poisson arrival rate may be obtained as $\hat{\lambda}=1/\overline{S}$. To prove equation (B2), we

note that the number of particles passing the sensor at an arbitrary time *t* is a stationary process whose marginal distribution is Poisson with mean λD (Taylor and Karlin, 1998). Thus the probability that no particles are passing by the sensor at time *t* equals $e^{-\lambda D}$. Denoting the expected value of the clump lengths as $\overline{C}$, we must have:

$$e^{-\lambda D} = \frac{E(S)}{E(S) + E(C)} \tag{B3}$$

Solving for *D*, we obtain

$$D = E(S)\ln\left(\frac{E(C)}{E(S)} + 1\right) \tag{B4}$$

Substituting means for expectations yields

$$\hat{D} = \overline{S}\ln\left(\frac{\overline{C}}{\overline{S}} + 1\right) \tag{B5}$$

A proof that this estimator is unbiased is provided in Bingham and Dunham (1997). Note that no assumption has been made regarding the distribution of particle size other than that a mean particle size exists and is finite.