Timely and Accurate Crop Yield Forecasting and Estimation

History and Initial Gap Analysis

By

George A. Hanuschak Sr.

GOAL(s)

The stated project goals are to identify state of the art methods, and suggest improvements moving forward, to improve worldwide and national forecasts and estimates of crop yields. It sounds quite simple but is anything but simple. It is complex and is very demanding of both applying proper statistical methods, and perhaps even more importantly demands very high data quality and integrity of the input data sources. This includes very large crop field agriculture to very small and often intercropped areas. This includes the majority of rain-fed crops to irrigated crops and large holdings like plantations. This includes developed economies, developing economies and some un-developed economies. This includes all types of soils, pest infestations, crop diseases, weather and climate situations and harvest loss etc. The variation even within the same crop year could include bumper record setting yields to parts or entire fields in a nearly total crop loss situation. Thus, the methods used must allow for rather incredible variability from crop field to crop field or even portions of a field. The project consists of a thorough literature review, both theoretical and applied, and an analysis of the strengths and weaknesses of current applications of crop yield forecasting and estimation used throughout the world. The emphasis here is on ground gathered data approaches. The ag-met models with or without remotely sensed data are very well reviewed by Atkinson (2013). I also assume that UN-FAO wants to establish or update country systems by working with national governments that are stable and have the appropriate staffing of statisticians, agronomists, meteorologists, and also remote sensing experts. The UN-FAO is in the best position on where to conduct new research studies or to enhance or develop operational systems.

BRIEF HISTORY

Starting with the 1940's and 1950's, there were at least two major developments. In India, the development of initial crop cutting procedures based upon sound statistical sampling procedures were developed and implemented. A distinguished statistician, Dr. P. V. Sukatme, was essential in establishing those procedures. In fact in the 1950's, UN-FAO recommended these procedures as the desired world standard. The crop cutting method for final yields has been adopted by many nations. In a somewhat parallel effort in time, procedures were developed in the U.S. for the Florida citrus crop to forecast and estimate the yield of a very important tree crop. Dr. Ray Jensen was the statistician responsible for the majority of this development. In many respects, these two methods still stand today with some revisions based upon new or improved sources of information, such as weather data, soils data, vegetative index data most of which is geo-referenced. The power of geo-referenced data will be discussed last in this paper before the recommendations on GAP analysis. Those efforts and methods were focused initially on final crop yield estimation at or near harvest and then progressed over time into crop yield forecasting several months before harvest. The author would also include a third category referred to as yield assessments which include valuable and needed information, but often can not reliably be converted into a statistical forecast or estimate. The review here will include all three types of yield information, including strengths and weaknesses. Regardless of the system used, there should be an available public track record of the forecast at the time and the final yield estimates that can be independently calculated and verified.

A major development occurred in the U.S. Department of Agriculture in the 1960's and expanded further in the 1970's and beyond. Statisticians Harold F. Huddleston, William E. Kibler and Charles E. Caudill were among those responsible in developing a scientific program referred to as "objective yield" for crop yield forecasting and estimation for major crop production areas in the U.S. These procedures remain basically intact today, with revisions for improved or new sources of data. This program is thoroughly documented by Aune et.al (2006) and more recently by Abreu and Riberas (2008). NASS 's relatively small International Programs Office has worked with numerous countries in transferring some of this methodology on several continents. There were advancements around the world on every continent ever since using similar procedures, but revised for regional or national or even local conditions. UN-FAO has been a consistent guide to many nations on documenting and suggesting state of the art procedures to follow. Dr. Rene Gommes, a world recognized expert has published a considerable number of papers and documents for agri-meteorological models to forecast and estimate crop yields. In the area of agri-meteorological models and outputs, I completely defer to Dr. Gommes as the expert source of information and will generally exclude that category in this paper. In addition, Dale Atkinson (2013) provided a very detailed literature review of crop area and crop yield forecasting using primarily ag met models of some sort. This paper will generally refer to crop yield forecasting and estimation systems based upon some form of ground data collection, sometimes including additional sources of weather or remotely sensed data in addition.

At the national level, the state of the art varies considerable. For developed nations, most have an advanced form of crop forecasting estimation procedures similar to crop cuts and "objective yield" procedures, and often an advanced sample survey system of farmer reported yields as well. Some even have composite estimation, which is a statistical weighting of the two sources. For developing economy nations, there may be a majority portion of such a system being developed. For some undeveloped economies, there may yet be a smaller portion of a complete and reliable system. One assumes a major project goal is to narrow the gap between these systems where possible, realizing the system will likely need to be somewhat different based upon the nation, and it's agricultural portfolio and practices and budget staffing levels for agricultural statistics purposes..

Literature Review (Books and Papers and Operational Documentation)

After a thorough literature review, some of the most valuable recent documents and papers include The Yield Forecasting Program of USDA/NASS (2006,2012), Fermont and Benson (2011) from the International Food Policy Research Institute, Joint Research Center of the EU methodology descriptions and papers over the last decade, papers from the Indian Agricultural Statistics Research Institute (IASRI), UN-FAO documentation. I will base much of my writing based upon those excellent references and an individual GAP analysis pointing toward productive areas of research for the UN-FAO project's next phases. There will also be a section on appropriate statistical text books as well. There will also be a major emphasis on data quality and integrity, regardless of the mathematical/statistical/modeling approaches used as often sample sizes are restricted due to cost considerations. All the methods described and tried DEPEND on very high quality data inputs. I will give very specific recommendations on maintaining or improving data quality for future efforts. For historic data, the FAOSTAT data base is an incredibly detailed source of agricultural related data.

The USDA/NASS objective yield forecasting system is well done and well documented by Aune et. Al (2006 and 2012), but it is somewhat expensive by world standards and aimed at large fields cropping. The models used are dependent on the stage of the crop development throughout the cropping season. The publicly available and independently verifiable track record is available on NASS's monthly crop report releases and occasionally in special reports as well. The data quality is strictly controlled by detailed field data collector's manuals and by small quality control samples. Probably, the greatest strength of the NASS program is its' highly trained field enumerator staff of thousands spread across the U.S. and also very familiar with the local agriculture practices. Since sample sizes and area samples are quite small in a relative sense, data quality is very essential. NASS also has a very extensive multiple frame based sample of farmer reported projections and final yields. The two major sources are then combined in a statistically based composite estimate. NASS also has a detailed Weekly Crop Stage and Condition report throughout the season composed of reports from several thousand agriculture experts. This report would fall into the category of a crop assessment. Even though it does not have a statistically measurable error term, it provides expert opinion knowledge throughout the season. Even though NASS does not recommend it for yield modeling, there are a number of private sector system forecasts that use the NASS weekly subjective data, often combined with weather data and even some remotely sensed data. In fact, the trading of agriculture commodities are affected by this data as well as the more scientific monthly crop reports of NASS, based upon a weighting of objective yield and farmer reported data.

Another very impressive system with a long history is that of India, based upon statistically defensible procedures for crop cutting and models for crop forecasting. The Indian system was the initial basis for the U.N. recommendation to use crop cutting procedures as the desired method for final crop yield estimation for decades starting from the 1950's. Undoubtedly, there are now hundreds of national systems across the continents with similar systems. The much more recent and thorough paper describing and comparing the various methods of crop yield forecasting and estimation by Fermont and Benson (2011) is quite impressive. In addition, it was done in Uganda and not a extremely large economy for agriculture such as the U.S., the EU, Brazil, Argentina etc. They provide an excellent description of the methods used throughout the world based on ground data collection, including strengths and weaknesses of the various methods. It was impressive to see a paper that compared the various methods in a similar environment. The methods included crop cuttings, farmer reported, agricultural expert reported, those with weather data involved, various forms of crop forecasting, et. al. Their conclusions were that crop cuttings were generally upward biased in some form, farmer reported were reasonably accurate, and other methods had some form of variation, explained or unexplained. The U.S. experience is somewhat similar, but with some differences over time about the different methods. One somewhat counter example experienced by the author over a career was when a soft red winter wheat crop in a major producing state had nearly eight months of nearly perfect weather conditions, but then a very late developing disease reduced yields rather dramatically, surprising even the farmers. To look at the fields and the crop, from the human eye everything looked great. However, a very late developing crop disease, well after plant senescence took a major toll on the crop yield. Even though the wheat plants and wheat heads looked great to the human eye, if one grabbed a wheat head and rolled it between your hands, the wheat kernels were miniscule in size. The objective yield data near harvest reflected this phenomenom, but the farmer reported data did not. So in this very complex environment, there seem to be no near perfect solutions. Yield modelers and forecasters often claim incredible accuracy, but then highly variable events occasionally surface to disprove such claims. In general, regardless of method used, the real track record usually shows that the confidence band of a forecast is larger as the time before harvest is. For one example Johnson (2011) independently verified NASS objective yield accuracy for corn 4 months from harvest to have root mean square error up to 6 percent, RMSE dropped to 3 percent two months from harvest and even the final has a sampling error around one percent. This track record is good but not overly impressive.

NASS has done research on many alternative yield models such as water balance, ag-met, markov chains, but has not revised it's preference for the objective yield approach based almost solely on plant and fruit counts and measurements. Using corn as an example, the philosophy is to directly measure for certain crop stages, variables like number of plants, number of ears and eventually final size and weight of ears. Based upon the laboratory work, the grain samples are dried to a standard moisture content as well. The final yield estimated also has a harvest loss component. NASS still does some research on alternatives, especially for small area, where objective yield approaches would be far too expensive, even in the U.S..

Several very recent examples were an early South American forecast of soybean yields (2012) that proved to be very over optimistic in nature and even caused political tension and the drought year of 2012 for the U.S. This are two mature national systems that had early forecasting problems in an extreme year. That is due to the FACT that unexplained future weather, disease, etc. cannot be fully explained before it's actual occurrence. Methods and data sources related to future predictions become better and better over time, however they are still far from perfect. On the other side, the further from harvest that a reliable forecast is available, the more valuable it is economically. However, need does not necessarily translate into accuracy. Thus there is room for and need for early warning or assessment programs, yield forecasting systems built on ground data and agri-met and other models and for farmer reported data and for final crop cuttings as well. Economics will likely determine how many approaches can be accomplished in any given nation.

Another excellent program on crop yield forecasting and estimation is that of the EU's Joint Reseach Center in Italy. It is well described by Genovese (1998, 2006) and by de Wit et. al. (2010), ---------------------------- and also on the JRC website.. It generally falls into the ag-met type category and is very well covered by Atkinson (2013) in his detailed literature review. The detailed literature review here is presented for both scientific papers and statistical text books related to crop yield forecasting and estimation based primarily on ground gathered data, such as crop cuttings and forms of objective yield. There are hundreds of papers listed, some for historical purposes. The topic of yield modeling is covered in detail by Atkinson (2013). There is considerably more documentation of yield modeling than there is for ground based systems for crop yield forecasting.

THE RICH HISTORY OF RELATED TEXTBOOKS

There is a very rich history of related statistical textbooks to the subject, directly or indirectly of sample surveys for crop area and yield forecasting and estimation. In 1937, Jersey Neyman presented a paper introducing the concept of confidence bands to statistics which is quite useful to this day. The recognized father of crop cutting in India, P.V. Sukatme studied under J. Neyman, R.A. Fischer and E.S. Pearson in London in the late 1930's and then applied the theory to crop cutting experimants in India. In 1954, P.V. Sukatme published his textbook on Sampling Theory of Surveys with Applications. In 1953, U.S. Census Bureau statisticians M. Hansen, W. Hurwitz and G. Madow published a two volume series on Sampling Survey Methods and Theory. These texts remained the basis until the W.G. Cochran book on Sampling Techniques. the third and finalversion published in 1977. More recently, there have been some very valuable texts published on sample survey methods, and in particular some aimed at agricultural statistics. Two aimed specifically at data quality and integrity are Nonsampling Errors in Surveys by Lessler and Kalsbeek in 1992, and Measurement Errors in Surveys edited by Beimer et. al. in 1991. In 2003, Beimer and Lyberg published Introduction to Survey Quality. On the topic of spatial and geo-spatial statistics, in 1989 Isaaks and Scrivastava published An Introduction to Applied Geostatistics. In 1993, Noel Cressie published a classical text on Statistics for Spatial Data. In 2001, Webster and Oliver published Geostatistics for Environmental Scientists. One text written by many world recognized authors is Agricultural Survey Methods, edited by Beneditti et. al. in 2010. This is likely the detailed text aimed specifically at survey methods for agriculture. Some other sampling texts were by Lohr, in 1999, with the title, Sampling: Design and Analysis and one of the most recent by Fuller, entitled Sampling Statistics. On time series, the classic text was by Box and Jenkins in 19--. On applied regression analysis, a classic was by Draper and Smith in 19--. The related statistical textbooks for sampling and estimation, in general, or for agriculture are rich in history. I have not even included here topics used on some occasions in crop production models and statistics, such as marchov chains, non-linear models, stochastic processes, small area estimation and other related texts on agro-meteorology, soils, agronomy, precision agriculture etc.

NEW TYPES OF DATA TO ASSIST IN CROP YIELD FORECASTING AND ESTIMATION

One of the rather incredible developments, over the last several decades, is the steady improvements in weather data, soils data, cropland data layers, crop disease detections. Advanced DOPPLER radar systems on a geo-gridded basis is one of those new and important sources. Ground based systems were often developed before some of these developments and are still catching up to take advantage of these better data sources. One early example was in Canada where vegetative index data from satellite was used to stratify the total crop area into sub strata based on crop stress and this substantially improved the precision of the ground gathered data. I believe there is significant opportunity to reduce the cost or increase the sample sizes with the cost savings. This will be covered more in the next section, which is an initial GAP analysis and has suggestions for the research and operational arm of this UN-FAO project down the road.

INITIAL GAP ANALYSIS AND RESEARCH SUGGESTIONS

This, hopefully will be helpful to UN-FAO plans on how to proceed from here on ground gathered data systems for crop yields, with an emphasis on developing economies or less developed economies to improve national systems. There will be recommendations for each phase, such as early warnings and crop assessments, statistically defensible crop forecasting and last statistically final crop yield estimation. In addition, small area crop yield forecasting and estimation will also be addressed if that is a UN-FAO goal of their five year project.

For the topic of crop assessments, it would be good research to see for a specific country if a system of agricultural experts could be established, that is geographically dispersed and can report on crop stage and condition in a timely electronic fashion. Basically, it would be a system like the Weekly Crop Stage and Condition reports of USDA/NASS. It is not a statistical system but is based on local expert opinion. In the U.S., agricultural extension experts report to NASS for the 3,000 counties of the U.S. The data reported is very popular with the external data users of NASS. Obviously, the UN-FAO early warning system is scientific and covers global areas based on weather data, soils data, vegetative index data etc. There, I am sure UN-FAO staff can provide suggestions on how to improve or expand that system developed by Gommes et. al.. Early warnings can have very significant impacts, especially for nations that import some of their food and fiber. The warning does not need to provide an exact yield number, but the direction of an impending crop disaster alone is quite valuable, even if the exact magnitude is not known yet. An excellent current example is a early warning by the UN-FAO about a leaf rust disease in much of Central America's coffee growing area.

The next category is statistically based objective yield procedures, such as in the U.S..NASS designs a self weighting sample based on crop area. They have very strict procedures for entering a selected field and placing randomly selected plots for crop counting and measurements throughout a season, such as corn ears, row spacing, corn ear length and circumference , etc.. The procedures are somewhat expensive and sample sizes are relatively small but can provide state level estimates. NASS attempts to keep any nonsampling errors to a minimum by detailed training and procedural manuals for the field data collectors. In addition, they have a small quality control sample where supervisors revisit the field to see if procedures were well followed. Since cost is an issue, procedures need to be made more efficient to transfer this type of methodology , especially to developing or under-developed economies. NASS International Program Office has some experience in such a methodology transfer, but it has been limited experience in this endeavor. IASRI and staff would also seem to be an excellent source of staffing for future research studies fir this component, plus the many other national statistics offices focused on agriculture.

Recommendations for improving the cost efficiency of such an approach would be to stratify first based on geo-referenced vegetative index data and weather data, to make the sampling more efficient, such as was done in Canada and India. The locally available weather data for the sample sites could then be added to the models, which currently do not use weather data directly. Some of the cost savings may be used to expand the quality control and the rest could be saved for other purposes .The U.S. combines a multiple frame based sample of farmers who either forecast their yield during the season or report the final yield after harvest and the objective yield model results. A very good list frame is required for this approach. The survey data of farmers is mostly reported in the U.S. by telephone or Internet to keep costs down.

The best report I found on ground gathered data for final crop yields was by Fermont and Benson (2011) having a literature review and well designed research and information comparing all the known types of ground based crop forecasting systems. They report on the strengths and weaknesses of the different approaches they used in Uganda. Crop cutting has been the recommended standard for decades. It has many strengths, but can also be costly. Farmer reported data came out relatively well in their study. The only caution is that on rare occasion, even the farmers are not aware of hard to observe crop diseases, especially if they occur late in a season. Composite estimation could be added, if the approaches are statistically measuring the desired population parameter and there are not significant nonsampling errors of one or more of the approaches. Strict quality control procedures are needed to ensure the data collected by field data collectors is correct. Other possible additions to research might be miniature weather stations near sample plot locations, and yield meter data for large field agriculture areas, and field technologies such as tablet computers with GPS. Close cooperation between a multi-disciplinary teams should define the staffing needs. It is very important regardless of the approach to be aware of the crop stage at the time of a forecast. Are there destructive sampling methods for the fruit weight measures that could lead to more accurate forecasts. Using corn as an example, the plant population and the corn ear population is known fairly in the season based upon counts. However, the final corn ear weights are the most difficult to forecast. Could it be possible to also count developed corn kernels or get better proxies to final kernel or ear weights. Statisticians need to work very closely with agronomists to make sure the forecasting methods are up to date, especially as varieties are constantly changing.

For the topic of small area crop yield forecasting and estimation, there is substantial opportunity for improvements over a five year period, mostly because of geo=referenced data on soils, weather such as precipitation, temperature, wind velocity etc,, cropland data layers and maps, vegetative indexes, etc. Prior to these developments, statistically based methods were used to disaggregate an estimate that was not designed for smaller breakouts. Many impressive statistical methods were developed and used but were hampered by the lack of reliable small area data. Now, a completely different era seems to be upon us. There is considerable geo-referenced data pertinent to crop yields that can be used to supplement good ground based systems. In areas where ground systems are not feasible due to civil wars and violence, the geo-referenced data related to crop yields may be the only feasible approach. It would not be as reliable as a combination of ground gathered systems from objective yield or farmer reported and the new related geo-referenced data sources. UN-FAO staff are in the best position to decide where ground gathered data systems are not feasible presently because of inherent dangers.

Whatever the method used for crop yield forecasting and estimation at any level of aggregation, it is important to distinguish between accuracy claims and a publicly available and independently verifiable track record between the forecast and the final yield estimate after harvest. A common measure is the root mean square error between the forecast and the official final. Because of the difficult nature of the task, these track records even for expensive and mature systems is not overly impressive. What is a realistic goal for a forecast four, three, two and one month from harvest. The further away from harvest, the larger the forecast error should be. Goals for forecast accuracy should be established, but in any case it is the true track record that eventually needs to be improved. Another major challenge is the timeliness of a forecast. The future on timeliness looks more promising as tablet computers, geospatial data access of all types more common, global positioning devices, computer efficiency in all types of yield forecasting methods.

More Extended GAP Analysis for Yield

On the topic of accuracy, the state of the art is judged by two major methods currently. First, is the statistical fit measures of the yield model used, such as r squared and others. This does not provide the second method of a track record of published forecasts versus the final yield. The most common method for the latter is root mean square error of a forecast versus final yield over a 10-20 year period. Usually, the track record is worse than other statistical fit measures but is probably the most realistic. Currently, in the U.S. with an expensive system, the track record for corn is not that impressive. At four months before harvest, it is 6.3 percent and drops to 3 percent two months before harvest. Perhaps a desired goal five to ten years from now would be 5 percent four months before harvest and 2-3 percent two months before harvest. Regardless of which nations, the track record has potential areas for improvement. There seem to be two special needs for accuracy improvement. One would be to conduct studies similar to Fermont and Benson (2011) in more countries and continents. A second one would add nonsampling error studies in addition aimed at potential reductions in nonsampling errors, regardless of method used. For farmer reported data, reinterview surveys can be helpful in identifying potential nonsampling errors. Cognitive principles are used to see how the farmer thought and information process goes when a farmer provides their original response. Because of response burden concerns, the sample sizes for reinterview surveys are usually held to a minimum.

When I was Director of Research and Development for the U.S., I used a matrix method for prioritizing research aimed at the ultimate goals. I prioritized potential research projects into a nine cell matrix where one column was anticipated payoff (High, Medium and Low) and the other was the expected cost of conducting the research and/or the cost of operational implementation (High,Medium and Low). Then I selected from the cells with a potentially high payoff with medium or low costs or with medium payoff with low costs. It was my crude form of GAP analysis.

The major GAP needs that I foresee in the future research efforts for this project for crop yield are the following. For topic of timeliness, small studies aimed at minimizing the time from data collection to the time of publishing results would seem to be in order. For the topic of cost, there need to be studies that that compare a baseline method (Ag-Met, Objective Yield, Farmer Reported, Administrative Data, etc.) with proposed streamlined alternatives aimed at cost reduction and to compare the precision or accuracy of the methods.

In order to attain a major goal of more uniformity across nations, it seems that processes that are somewhat streamlined for both cost and complexity of implementation are desirable. Certainly, available staffing with proper training in each nation is rather critical. Another potential effort is to examine training programs for both professional staff and any field data collectors as well. One good example provided is the NASS 2013 Corn Objective Yield Enumerator's Manual.

Another area that needs addressing is the national staffing available or the more regional or global staffing available. Perhaps, more can be accomplished nationally than global. If global is the goal, then perhaps some combination of staffing from UN-FAO, JRC. NASS, IASRI and other impressive national systems such as Brazil, Argentina, Europe, Australia, etc. is the goal. If the goal is national systems and staffing, then the bar changes.

Another important goal ls the definition of yield. The recommended definition is biological yield minus harvest loss in the field (standardized to a moisture yield content).

References

Abreu, D. and Riberas, Z. (2008) "General overview of the NASS objective yield and objective measurement programs", RDD Research Report Number RDD-08-10, USDA NASS, Fairfax, VA.

Adrian, D.,(2012) "A model-based approach to forecasting corn and soybean yields", USDA, National Agricultural Statistics Service, Research & Development Division. 2012.

Agrawal, Ranjana and Mehta, SC (2007). Weather based forecasting of crop yields, pests and diseases - IASRI models. J. Ind. Soc. Agril. Statist., 61(2), 255-263.

Ahmad, T and Kathuria, OP (2010). Estimation of crop yield at block level. Adv. Appl. Res., 2(2), 164-172.

Allen, Rich; Hanuschak, George; and Craig, Michael. (1994) "Forecasting crop acreages and yields in the face of and in spite of floods", presented at the United Nations FAO / European Union Expert Consultation on Crop Yield Forecasting Methods, Villefranche-Sur-Mer, France, October 24-27, 1994.

Agrawal, Ranjana, IASRI. 2005. Forecasting Techniques in Crops.

Aune, Dave and Vogel, Fred (2006) "The yield forecasting program of NASS", Statistical Methods Branch, Estimates Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C., May 2006. NASS Staff Report No. SMB 06-01.

Baruth, Bettina (2012) "Wiki MARS crop yield forecasting systems", JRC Technical Reports JRC77958, European Commission Joint Research Centre, Via E. Fermi, 2749, I-21027 (VA) Italy.

Baruth, B., Royer, A., and Klisch, A., EC - Joint Research Centre, IPSC, Agriculture –Unit, 21027 Ispra (Italy), G. Genovese, EC- Directorate General for Agriculture and Rural Development, 1049 Bruxelles (Belgium). The Use of Remote Sensing Within the MARS Crop Yield Monitoring System of the European Commission. 2008.

Berg, Emily and Chandra, Hukum (2012) "Small Area Prediction for a Unit-Level Lognormal Model", FCSM IV-A, 2012.

Bhatia, VK (2011) "Vision 2030", Indian Agricultural Statistical Research Institute (IASRI), *New Delhi,* India.

Bhatia, VK; Sud, UC; Gupta, VK; Singh, Man; Sharma, DP; and Singh, DP (2011). Study to Determination of Optimum Sample Size for Yield Estimation at Gram Panchayat Level, IASRI, New Delhi.

Birkett, Thomas (1993) "Yield Models for Corn and Soybeans Based on Survey Data", talk presented at the Washington Statistical Society Meetings 10/27/1993, Washington DC.

Bojanowski J; Vreiling, A; and Skidmore, A. (2012) "Towards long-term time series of global solar radiation data for regional agro-meteorological modeling", EUMETSAT Meteorological Satellite Conference, Sopot, Poland, September 6th , 2012.

Boken, V.K., 2000. Forecasting spring wheat yield using time series analysis: a case study for the Canadian Prairies, Agronomy Journal, 92(6):1047-1053.

Boogaard, Hendrik; Wolf, Joost; Supit, Iwan; Niemeyerd, Stefan; and van Ittersuma, Martin (2013) “A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union”, Elsevier, *Field Crops Research* 143 (2013) 130–142.

Cantelaube, P. and Terres, J. European Commission Joint Research Centre/Institute for Environment and Sustainability. 2004. Seasonal weather forecasts for crop yield modeling in Europe.

Carfagna, E. and Gallego F. J. (1997) “Yield Estimates from Area Frame at European Level” in Prévisions de rendement agricole, Luxembourg, Office for Official Publications of the European Communities, 1997, pp. 197-202

Caudill, Charles and McArdle, Richard. (1979) “Research Evaluation Considerations for Agristars”, Third Conference on Economics of Remote Sensing, Lake Tahoe. November 1979.

Caudill, Charles (1976) "Current Methods and Policies of the Statistical Reporting Service," Third Symposium on Machine Processing of Remotely Sensed Data, LARS 76, Purdue University. June 1976.

CSO (2008) “Standard Report on Methods and Quality (v1) for Area, Yield and Production of Crops”, Central Statistics Office, Skehard Road, Cork 021 4535000, edited July 11, 2008.

Chandrahas; Agrawal, Ranjana; and Walia, SS (2010). Use of discriminant function and principal component techniques for weather based crop yield forecast. IASRI/PR-09/2010, IASRI, New Delhi.

Chandran, KP and Prajneshu (2005). Nonparametric regression with jump points methodology for describing country’s oilseed yield data. J. Ind. Soc. Agril. Statist., 59(2), 126-130.

De Wit, Allard; Duveiller,Gregory; and Defourny, Pierre. (2012) “Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations”, Elsevier, *Agricultural and Forest Meteorology* , 164 (2012) 39– 52.

De Wit, Allard; Baruth, Bettina; Boogaard, Hendrik; van Diepen, Kees; van Kraalingen, Daniel; Micale, Fabio; te Roller, Johnny; Supit, Iwan; and van den Wijngaart, Raymond. (2010) “Using ERA-INTERIM for regional crop yield forecasting in Europe”, *Climate Research*, Vol. 44, pg. 41-53, published online October 28, 2010.

Dong (2011) “Crop monitoring as an E-agriculture tool in developing countries (E-AGRI)”, Presentation for European Union, FP7 STREP Project (GA 270351).

Doraiswamy, P.C., S. Moulin, P.W. Cook, and A. Stern. 2003. Crop yield assessment from remote sensing, Photogrammetric Engineering and Remote Sensing, 69, 665– 674.

Doraiswamy, P.C., J.L. Hatfield, T.J. Jackson, J.H., B. Akhmedov, and A.J. Stern. 2004. Crop condition and yield simulations using Landsat and MODIS imagery, Remote Sensing of Environment, 92: 548– 559.

Doraiswamy, P.C., T.R. Sinclair, S. Hollinger, B. Akhmedov, A. Stern, and J. Prueger. 2005. Application of MODIS derived parameters for regional yield assessment, Remote Sensing of Environment. 97(2), 192-202.

Doraiswamy, P.C., B. Akhmedov and A.J. Stern. 2006a. Improved techniques for crop classification using MODIS Imagery. Proceedings of the International Geoscience and Remote Sensing Symposium, July 31 – August 4, 2006b, Denver, Colorado. CD ROM.

Doraiswamy, P., Akhmedov, B., Beard, L., Alan Sterna, A and Mueller, R. 2006b. Operational Prediction of Crop Yields Using MODIS Data and Products.

Doraiswamy, P.C., B. Akhmedov and A.J. Stern. 2006c. MODIS time Series data applications in Agriculture. A CEOS Land Product Validation topical workshop: Validation of global vegetation indices and their time series. University of Montana Missoula, MT. August 7, 2006.

Duveiller, Grégory; López-Lozano, Raúl; Seguini, Lorenzo; Bojanowski, Jędrzej S.; and Baruth, Bettina. (2013) "Optical Remote Sensing Requirements for Operational Crop Monitoring and Yield Forecasting in Europe", European Commission, Joint Research Centre, Via E. Fermi, 2749, I-21027 (VA) Italy.

Duveiller, Grégory; López-Lozano, Raúl; and Baruth, Bettina. (2013) "Enhanced Processing of 1-km Spatial Resolution fAPAR Time Series for Sugarcane Yield Forecasting and Monitoring", *Remote Sens.* 2013, *5*, 1091-1116; doi:10.3390/rs5031091.

Duveiller, G.; Baret, F.; and Defourny,P. (2012) "Remotely sensed green area index for winter wheat crop monitoring: 10-Year assessment at regional scale over a fragmented landscape", *Agricultural and Forest Meteorology*, Elsevier, vol. 166–167, pp. 156–168, Dec. 2012.

Fermont, Anneke and Todd Benson. (2011) "Estimating Yield of Food Crops Grown by Smallholder Farmers: A Review in the Uganda Context", International Food Policy Research Institute, IFPRI Discussion Paper 01097, June 2011.

Gallego, J; Carfagna, E; and Baruth, B. (2010) "Accuracy, Objectivity and Efficiency of Remote Sensing for Agricultural Statistics", Chapter 12, *Agricultural Survey Methods*, pg. 193-211, Wiley & Sons, United Kingdom.

Genovese, G.P., 1998: The methodology, the results and the evaluation of the MARS crop yield forecasting system. In: Agro-meteorological Applications for Regional Crop Monitoring and Production Assessment (D. Rijks, J.M. Terres and P. Vossen, eds). Luxembourg, Office for Official Publications of the European Communities.

Genovese, G., Fritz, S., Bettio, M., EC - Joint Research Centre, IPSC, Agriculture. 2006. A Comparison and Evaluation of Performances Among Crop Yield Forecasting Models Based on Remote Sensing: Results from the GEOLAND Observatory of Food Monitoring.

Genovese, G; Confalonieri, R.; and Micale, Fabio . (2007) "The Rice yield forecasting system of the EC", Presentation at the 4th International Temparate Rice Conference, June 25-28, Novare, Italy.

Ghosh, H; Singh, RK; and Prajneshu (2008). Possibilistic linear regression analysis with fuzzy response variable for crop yield estimation. Cal. Stat. Assoc. Bull., 60, 123-44.

Hale, Robert; Hanuschak, G.; Craig, M. (1999) "Appropriate Role of Remote Sensing in U.S. Agricultural Statistics", FAO Seminar on Remote Sensing for Agricultural Statistics, Bangkok, Thailand, 6/99

Hanuschak, G. and Delincé, J. 2005. Utilization of Remotely Sensed Data and Geographic Information Systems (GIS) for Agricultural Statistics in the United States and the European Union. 2005.

House, Carol (1977) "A Within-Year Growth Model Approach to Forecasting Corn Yields", Unpublished Manuscript, Crop Reporting Board, USDA ESCS, now USDA NASS, Washington, DC, September, 1977.

Houseman, E. E., and Huddleston, H. F. (1966). Forecasting and Estimating Crop Yields from Plant Measurements. Unpublished Manuscript. Monthly Bulletin of Agricultural Economics and Statistics, Vol. 15 #10.

Huddleston, H. F. (1978). "Sampling Techniques for Measuring and Forecasting Crop Yields", Unpublished Manuscript, Research Report # ESCS-09, Economics, Statistics and Cooperatives Service, (now NASS) USDA.

Huddleston, H.F. (1975) "Research in Yield Forecasting – Update", Unpublished Manuscript, USDA NASS RDD.

Huddleston, H.F. (1972) "Research in Yield Forecasting", Unpublished Manuscript, USDA NASS RDD.

Jaggi, Seema; Gill, AS; Varghese, Cini; Sharma, VK; and Singh, NP (2009). Impact assessment of agroforestry system on yield of associated barley (Hordeum vulgare) and gram (Cicer arietinum) crops. Current Adv. Agril. Sci., 1(2),101-105.

Kibler, William (1964) "Predicting Yields from Objective Counts and Measurements", Unpublished manuscript, presented at the Annual Convention of the Association of Southern Agricultural Workers, February 1964, USDA NASS, Washington, DC.

Kouadioa, Louis; Duveiller, Grégory; Djabya,Bakary; El Jarroudia, Moussa; Defournyb, Pierre; and Tychona, Bernard. (2012) "Estimating regional wheat yield from the shape of decreasing curves of green area index temporal profiles retrieved from MODIS data", *International Journal of Applied Earth Observation and Geoinformation* 18 (2012) 111–118, Elsevier.

Kumar, Anil; Panwar, Sanjeev; Kumar, Vipin Choudhary; Sanat Kumar; Kumar, Pankaj; and Singh, Prem (2009). Yield estimation in rice-rice cropping system under long term fertility experiments. J. Farm. Sys. Res. Dev., 15(1&2), 170-174.

Kumar, A; Ramasubramanian, V; and Agrawal, R (2007). Forecasting rice yield using neural networks. Proceedings of the 3rd Indian International Conference on Artificial Intelligence (IICAI-07), Pune, 1510-1516.

Kumar, Amrender and Bhar, LM (2005). Forecasting model for yield of Indian mustard (Brassica juncea) using weather parameter. Ind. J. Agril. Sci., 75(10), 688-690.

Kumar, Anil; Kaur, Ajit; Kaur, Rajinder; Sharma, GC; and Gangwar, B.(2005) Combined analysis of experiments on long range effect of continuous cropping and manuring on soil fertility and yield stability. (AP Cess Funded)

Lazăr, Cătălin; Baruth, Bettina; and Micale, Fabio. (2009) "Winter wheat yield estimation for romania, based on normalized difference vegetation index data available on marsop site", AN. I.N.C.D.A. Fundulea, Vol. LXXVII, Fiziologia Plantelor, National Agricultural Research and Development Institute Fundulea, 915200 Fundulea, Călăraşi County, Romania, 2009.

Matis, J.H., Perry, C.R., Boudreaux, D.E., and D.J. Aune (1989). *Markov Chain Forecasts of Cotton Objective Yield,* National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250, Research Report No. SRB 89-11.

Mehta, SC; Pal, Satya; and Kumar, Vinod. (2010) Weather based models for forecasting potato yield in Uttar Pradesh. Project Report No. IASRI/P.R.-01/2010, IASRI, New Delhi.

Meroni, Michele; Marinho, Eduardo; Sghaier, Nabil; Verstrate, Michel M; and Leo, Olivier (2013) "Remote Sensing Based Yield Estimation in a Stochastic Framework — Case Study of Durum Wheat in Tunisia", ", *Remote Sens.* 2013, *5*, 539-557; doi:10.3390/rs5020539.

Mohan, Madan; Sharma, Rajbir; Rai, T; and Agrawal, Ranjana. (2007) Forecasting of yield loss due to weeds. Research Project Reports 07-08

NASS-ASB (2013) "2013 Corn Objective Yield Survey – Interviewers Manual", Manual prepared by the Agricultural Statistics Board (ASB), USDA NASS, Washington, DC.

NASS-SMB (2012) "The Yield Forecasting and Estimating Program of NASS", Statistical Methods Branch (SMB), Statistics Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C., NASS Staff Report No. SMB 12-01. 2012.

Nath, S.K. (2008) "Manual on Area and Crop Production Statistics", CSO-M-AG-01, Central Statistical Organization, Indian Agricultural Statistical Research Institute (IASRI), *New Delhi,* India.

Neas, Kathy; Molina, Jubal; Hardegree, Jason; Gerling, Michael (2006) "Using Personal Digital Assistants for the 2004 Cotton Objective Yield Survey", Paper RDD-06-01, Research and Development Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C.

Nègre, Thierry (2012) " Crop Monitoring and Food Security: the JRC action and prospect", Presentation to the Open Session on "Space for Agriculture and Food Security", EC JRC, Rome, 09 March 2012.

Rai, Anil. (2003) "Crop Area and Production Surveys", Indian Agricultural Statistical Research Institute (IASRI), *New Delhi,* India.

Rao, AR and Prabhakaran, VT (2005). Use of AMMI in simultaneous selection of genotypes for yield and stability. J. Ind. Soc. Agril. Statist., 59(1), 76-82.

Rembold, Felix; Atzberger, Clement; Savin, Igor; and Rojas, Oscar. (2013) "Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection", *Remote Sens.* 2013, *5*, 1704-1733; doi:10.3390/rs5041704.
Rojas, O. (2007) "Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya", *International Journal of RemoteSensing*, 28:17, 3775-3793, Taylor and Francis Online.

Sakellariou-Makrantonaki, M and Vagenas, I.N. (2006) "Mapping Crop Evapotranspiration and Total Water Requirements Estimation in Central Greece", European Water 13/14: 3-13, 2006, E.W. Publications.

Singh, D., Directorate of Economics & Statistics, Ministry of Agriculture, New Delhi, India. 2008. Forecasting Techniques for Agricultural Output.

Singh, RK and Prajneshu (2008). Artificial neural network methodology for modeling and forecasting maize crop yield. Agril. Eco. Res. Rev., 21, 5-10.

Sethi, SC; Mathur, DC; and Pandey, PS (2007). Estimation of wheat yield at block level of Haryana using small area method. J. Farm. Sys. Res. Dev., 13(1), 138-141.

Sethi, SC; Sood, RM; Dass, Bhagwan; and Mathur, DC (2006). A study on block level estimators of crop yield and comparison of estimates for important crops in the Sirsa district of Haryana State. Hry. Eco. J., 26(1-2), 103-105.

Sharma, S.D.; Srivastava, A.K.; and Sud, D.C. "Small Area Crop Estimation Methodology for Crop Yield Estimates at Gram Panchayat Level", *Jour. Ind. Soc. Ag. Statistics 57 (Special Volume), 2004* : 26-37, *Indian Agricultural Statistics Research Institute. New Delhi, India.*

Singh, Rhandhir (2012) "Crop Yield Estimation and Forecasting Using Remote Sensing", Chapter 19, *Indian Agricultural Statistics Research Institute.* New Delhi, India.

Singh, Rhandhir ?? (2005) Application of Multiple Frame Sampling on Multi Resolution Satellite Data for Crop Yield Estimation", ssrn-id956398, *Indian Agricultural Statistics Research Institute.* New Delhi, India.

Singh, Rhandhir (2004) Application of Remote Sensing Technology For Crop Yield Estimation. J. Ind. Soc. Agril. Statist., 57, Special Issue, 226-246

Singh, Rhandhir ; Sahoo, Prachi Misra; and Anil Rai (2003) "Use of Remote Sensing and GIS Technology in Agricultural Surveys", Map India Conference 2003 Agriculture, Indian Agricultural Statistics Research Institute (ICAR), New Delhi, India.

Singh, Rhandhir (2003) Use of Satellite and Farmers Eye Estimate For Crop Yield Modeling. J. Ind. Soc. Agril. Statist., 56(2), 166-176

Steele, Ronald (1986) "Corn Objective Yield: An Empirical Evaluation of the Use of 3, 4 or 5 Years Data to Develop Forecast Equations", NASS Staff Report YRB-86-05, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C.

Sud, UC. (2012) “Small Area Estimation Technique For Crop Yield Estimation At District Level”, 6: Small Area Estimation Technique, , Indian Agricultural Statistics Research Institute (ICAR), New Delhi, India.

Sud, U C; Bhatia, V K; Chandra, Hukum; and Srivastava, A K. (2011) “Crop Yield Estimation at District Level by Combining Improvement of Crop Statistics Scheme Data and Census Data”, Wye City Group on Statistics on Rural Development and Agricultural Household Income, 4th Meeting, Rio de Janeiro, Brazil, November, 2011.

Sud, UC; Bathla, HVL; Khatri, RS; Mahajan, VK; Mathur, DC; and Chandra, Hukum (2009). Pilot study on small area crop estimation approach for crop yield estimates at the gram panchayat level. Project Report, IASRI, New Delhi.

Sud, UC; Mathur, DC; Jha, GK;, Sethi. SC; and Bhasin, RM (2006) Crop yield estimation at small area level using farmers estimates Research Project Reports 06-07

Vogel, Fred (1971) “Using Corn Plant Vegetative Characteristics to Forecast Production of Grain per Plant”, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C.

Wade, Gail; Hanuschak, G. (2000) "Visualization of a Crop Season - The Integration of Remotely Sensed Data and Survey Data", Proceedings of the ASPRS 2000 Conference, Washington DC.

Waldhaus, E., Oaks, E., and Steiner, M. (1993). “History and Procedures of Objective Yield Surveys in the United States”, Unpublished Manuscript, USDA National Agricultural Statistics Service, talk presented at the Washington Statistical Society Meetings 10/27/1993, Washington DC.

Warren, F. B. (1985). “Corn Yield Validation Studies, 1953-83”, Unpublished Manuscript, Research Report #YRB-85- 07, Statistical Research Division, Statistical Reporting Service (now NASS), USDA, Washington DC.

Williams, S.R. (1969) “Forecasting Florida Citrus Production Methodology and Development”, Florida Crop and Livestock Reporting Service, Unpublished Manuscript, Statistical Reporting Service (now NASS), USDA, Washington DC.

Wilson, Wendell (1985) “Evaluation of Crop Yield Models”, Unpublished Manuscript, Research Report, Statistical Research Division, Statistical Reporting Service (now NASS), USDA, Washington DC.

APPENDIX A—Crop Yield Forecasting and Methods Approaches

- Ag-Met Models (most are quite sophisticated and complex in nature
- Water Balance and Soil Based Models (again complex in nature)
- Farmer Reported Data after Harvest from Sample Surveys
- Farmer Reported Data before Harvest from Sample Surveys
- Methods based upon crop and fruit counts and measures from sample plots
- Crop cuttings at harvest
- Many statistically based models such as those based on:

  Single and multiple independent variables
  Time series analysis
  Composite Estimation
  Quality control and replicated sampling
  Sampling theory and application
  Marchov Chains
  Cognitively designed questionnaires and forms for farmers and for field data collectors

Appendix B--Organizations with experience

- National Agricultural Statistics Offices from every continent
- United Nations Food and Agriculture Organization
- Indian Agricultural Statistics Research Institute
- Joint Research Center of the European Union
- World Meteorological Organization
- Agronomic Research Centers around the Globe

Appendix C—Multi-Disciplinary Team for Crop Yield Forecasting and Estimation

- Statisticians
- Agronomists for Crops of Interest
- Agricultural Meteorologists
- Remote Sensing and Geo-Spatial Data Experts
- Information Technology and Statistical Software Experts
- Field Data Collectors, where required
- Questionnaire and Forms Designers (using cognitive principles)
- Soil scientists, if required
- Budget Officers

Appendix D—Proxy Yield Variables for Corn

- Leaf Area Index (LAI)
- Normalized Difference Vegetative Index (NDVI)
- Temperature
- Precipitation
- Wind Velocity
- Growing Degree Days (GDD)
- Plant Counts
- Ear Counts
- Over the Husk Ear Length
- Over the Husk Ear Circumference
- Preliminary Lab Ear Weights (Standardized Moisture Content)
- Final Lab Ear Weights (Standardized Moisture Content)