A REVIEW OF AREA SAMPLING FRAME TECHNIQUES AND INTRODUCTION TO THE USE OF SPOT IMAGERY IN AREA FRAME CONSTRUCTION

Prepared by:

The Food Security Management Project, USAID
Agriculture Data Collection Component

In cooperation with:

Government of Pakistan
Statistics Division
Federal Bureau of Statistics
Agriculture Data Collection Cell

August 8, 1989
October 1989 (revised)

INTRODUCTION:

Area Sampling Frame (ASF) technology has become the foundation for the collection of agricultural data in much of the world. It is a proven method of monitoring agricultural outputs, providing decision makers with a near real time picture of the food and fiber situation. Knowledge of the current situation in agriculture may be the most important stone in the foundation of Food Security Management. Introducing Area Sampling Frame concepts and technology into Pakistan is the major thrust of the Food Security Management Project/Agriculture Data Collection Component sponsored by USAID and the Government of Pakistan. Training in Area Sampling Frame (ASF) construction began in November 1985. A pilot project of seven districts was specified in the project to test and demonstrate the workability of ASF technology in Pakistan. The project began conducting field surveys in April 1987 clearly demonstrating the effectiveness of this approach.

In July of 1989, permission was granted to begin extending the Area Frame into a National data collection system. It was decided to begin national coverage in the pilot project areas. This was done in order to validate new technologies employed in frame construction. The two major technical changes are: using SPOT (Systeme Probatoire D'Observation De La Terre) satellite imagery as the frame stratification media, and computerizing of frame area measurements and accounting.

The use of SPOT Satellite Imagery does not change any of the fundamental ASF concepts but raises the level of technology employed substantially. Initially a 'hands on' rather low technological approach was recommended. Experience has demonstrated that the Federal Bureau of Statistic (FBS) staff is able to absorb the necessary training and make use of newly developed and highly sophisticated equipment. The National Area Frame will be the constructed using 'state of the art' automated techniques unavailable at the time of project initiation. It is anticipated that the new area frame procedures introduced here will bring the FBS Area Frame Laboratory to World Class standards.

According to the Fifth Annual ADC Project Work Plan, area frame construction will be completed in about two years. This assumes the full support of USAID, The Government of Pakistan/ Federal Bureau of Statistics, and the continued dedication of the Agriculture Data Collection Cell staff. The USDA advisory team believes that this is an ambitious, but not unreasonable goal.

Concurrent with National Area Frame construction the ADC Cell will continue to conduct field surveys in the pilot area. Valuable insight will be gained in designing and managing major field surveys. The pilot area surveys are already proving to be a useful laboratory for research in survey instrument design, and objective yield survey training and

implementation. Research in many aspects of survey procedures will be undertaken in this manner before the first anticipated National Area Frame Survey to be conducted in January 1992.

A REVIEW OF AREA SAMPLING FRAME CONCEPTS

A Sample can be defined as an item(s) which show the quality of the whole from which it was taken. Statistical sampling refers to the science of selecting samples from a whole, or population, in such a manner that an valid inferences can be made about the whole from which it is drawn.

A Sampling Frame is list of items in a target population from which a statistical sample can be drawn. The sampling frame may or may not include all of the member items of the population, but an inference drawn will only be valid for the members included population. Frames which include all the elements of the population are called complete frames. In addition, to be a valid statistical sample the quality of the inference drawn must be measurable.

The Area Sampling Frame (ASF) is a list of parcels of land, termed segments, from which a statistical sample can be drawn. ASF has proven desirable for the collection of agricultural statistics for a number of reasons:

1. most attributes which are of concern in an Agricultural Statistics program can be associated with a unique parcel of land (Segment),

2. it is a complete frame, with no duplication or omission, that is any single element in the entire target population has some known probability of being included in a sample drawn from this frame;

3. the area frame produces statistically unbiased estimates with measurable precision;

4. it is long lived: the area frame can be used for many years without update;

5. ASF lends itself to subsampling for special purpose surveys, for example crop cutting/objective yield surveys, and some types of economic surveys;

6. because of relatively small samples necessary, area frame surveys can usually be conducted rather quickly, which helps to insure timely results; and

7. the area frame survey is relatively inexpensive to operate.

Research into area frame technology for agricultural purposes began at Iowa State University in 1938. The United States Department of Agriculture, National Agricultural Statistics Service (USDA/NASS) began using the area frame in 1954, and it is now the main stay of the area and production statistics program in 48 of the 50 states. It is not used in the far north State of Alaska with limited agriculture, nor in Hawaii where agricultural is very specialized.

ASF has been introduced in many developing countries, including Pakistan, Morocco, El Salvador, Costa Rica etc. More than two years of survey history in Pakistan has demonstrated the feasibility of the area frame for agricultural surveys. ASF was introduced in central American in the early 1970's. In every case where it has been introduced, except Egypt, it is believed that ASF has been incorporated as part of the statistical institutions of the country.

In simplest terms AREA SAMPLING FRAME technology requires an accurate base map on which to divide the target area (population) into ultimate sampling units, or segments. The area of the population must be accurately measured to calculate the probability of selection of each segment in the sample. As a minimum ONLY accurate base maps which can be used to define sampling units and measure their area is required. A more efficient frame can be constructed if the target area can be stratified by major land use categories. A properly stratified sample will produce an indication (estimate) with a higher level of precision, that is a lower sampling error. Stated still another way: A properly stratified sample is more efficient.

SAMPLING ERROR refers to the difference in the sample estimate and the true value which is due to sampling, or using a small portion of the population to make an inference about the total. This error is knowable, predictable, and measurable based on analysis of survey data.

The ability of a survey to precisely make an estimate is usually reported by the COEFFICIENT OF VARIATION (C.V. or CV), which is the standard error of the estimate divided by the estimate expressed as a percentage. For unbiased estimates with no non-sampling error, a cv of 5% means that if the survey was duplicated many times the resulting estimates would be within 5% of the real or true value 69% of the time; within 10%, or 2cv's, 95% of the time; and values within 3 cv's would occur more than 99% of the time. This variation from the true value is the result of sampling error.

The 'true' value of an estimate can never be known. The presence and magnitude of non-sampling error cannot be determined with certainty: there are no perfect surveys (or census'). It follows that ACCURACY or how close a survey estimate is to reality cannot be known. Statisticians normally refer to PRECISION, which is the expectation that the average estimate from a great many similar surveys will fall within a specified range. Precision assumes that bias and non-sampling error are consistent survey to survey. Precision can be estimated, and predicted, but the relationship between a precise estimate and the 'truth' can still not be known with certainty.

NON-SAMPLING ERROR includes all the errors in an estimate which come from all sources other than sampling error. It is non sampling error which renders the truth unknown. The total impact of non-sampling error can never be fully known. It includes such diverse sources as recording and measurement errors during frame construction, errors in crop identification during the field survey, and key entry errors in

survey processing. Only vigilance and constant effort will reduce non-sampling error. It [illegible] ver be completely eliminated, or if it is the statistician will never [illegible] ow. Most non-sampling errors are not measurable or predictable and are often unknown entirely.

EFFICIENCY refers to the cost of producing an estimate with a given CV. The more efficient survey design is one which produces a estimate with a lower CV for the same cost, or the same CV for a lower cost. Cost is often measured in terms of the money cost of conducting the survey or the man-hours involved. Sample sizes are often good indicators of relative survey cost for similar surveys, but are very misleading when comparing surveys of different designs.

Another cost/value which should be considered in survey design is the TIME VALUE of estimates. An otherwise very efficient survey which produces a result too late to be of value is worthless. For example a wheat yield forecast before harvest time might allow the government to take full advantage of international markets in making up a short fall, or disposing of a surplus. This early estimate could easily be worth millions and millions of Rupees in market advantage. The same estimate, available weeks after the harvest, might have no valve at all. This is TIME VALUE. It is one of the most critical and often ignored aspects of survey planning.

Another important aspect in survey design is the level of desired precision expected from a sample survey. A CV of ten or fifteen percent might be adequate at the district level, while provincial estimates could be in the range of five percent for major items. At the same time decision makers might expect National estimates to have a CV of 2 or 3 percent. The statistician can with reasonable success plan, predict, and adjust, in advance, the level of precision a survey will deliver. Remember the indications and variance from independent surveys are additive, but the CV must be recalculated. The effect is as similar survey results are combined the aggregate CV declines. It is fully reasonable that a series of district surveys producing estimate with 10 percent CV's will combine to a national estimate with a CV 2 or 3 percent or less.

Because the highest aggregate estimate is usually the one most important to the decision makers, sample planners should consider the precision of this estimate first. Usually this will be the National or Provincial estimate. District or lower level estimates can still be made, but some control of their precision may be lost. This 'top down' approach is the most efficient way to allocate resources, but is often difficult for local politicians, and sometimes statisticians to accept as it de-emphasize the importance of lower level estimates.

As a rule of thumb the CV of an estimate can be reduced by half by increasing the sample size four times, an inverse geometric ratio. If the 10% district CV in our previous example was obtained with a sample of 100 segments per district it could be reduced to 5% by sampling 400 segments per district. The number of enumerators and office support

staff required will increase by a similar amount. The resources devoted to training and quality assurance monitoring must also increase. The volume of paper which must be handled will expand substantially. If 6,000 forms per district are processed to produce the 10% district CV, it is likely that 1.8 million forms will be required to produce national estimates at the 5% per district CV level. If rate of non-sampling error did not change, this reduction in CV increases the opportunity for error geometrically. For many reasons, the rate of non-sampling error will increase as the work load increases further compounding the effect of increased error opportunity.

Survey planners and data users need to carefully consider the level of required precision in light of the resources and time required to complete the survey. Consider also possible effect on non-sampling error as a survey becomes larger. Both for money economy and survey operation economy, survey statisticians should encourage acceptance of the highest possible CV compatible with the data needs of the user at every level of summarization.

A term needs to clarified before continuing. The word ESTIMATE in agricultural statistic is used in two differences senses. 'Estimate' is a statistical term which refers to the final expansion of data from a sample survey, i.e. the area planted to wheat from the Rabi ASF survey. In common terms the reported valve after the results of one or more surveys are reviewed and made public is also termed an ESTIMATE, i.e. the yield of wheat per hectare reported by the Minister in the newspaper. These are both valid uses of the word, however confusing. It would be desirable to interchange the word INDICATION for estimate in the first sense, that is of the results of statistical surveys. As many survey INDICATIONS might be available to the statistician who will finally set one ESTIMATE for publication.

Earlier in the discussion it was noted that stratification is a tool used to increase survey efficiency. Stratification is the process of grouping similar sampling units together based on there similarity. It should be apparent that a smaller sample is necessary to measure a parameter if all of the units in the population have similar values. The grouping of similar items together for sampling (Stratification) is the most important action that can be taken to increase the efficiency of the area frame sample. This is an integral part of frame construction.

The stratification process includes the following well defined steps:

1. REVIEW and SELECT resource materials that will provide the information need to segregate the target area into meaningful categories. In the developing world where wide spread resource inventories may be few there is often little to chose from. Aerial photography and satellite imagery are most common. Satellite imagery is now universally available, while photography is not.

2. DEFINE the various strata to be used. Definition should be appropriate for the intended use of the frame. They should reflect what is possible, given the stratification medium used. Definitions should be as specific as possible, but worded in such away that every parcel of land in the target population fits into one and only one stratum.

3. Delineate the target area into STRATUM BLOCKS. A stratum block is the largest contiguous area in which all the included land is of the same strata type. The minimum size for a stratum block is one frame unit. The technicians should be careful that no area is duplicated in more than one stratum, and that no area is omitted (Remember the complete frame).

FRAME UNITS are formed when stratification is complete by dividing stratum blocks into uniform sized smaller subdivision. The subdivided stratum block, or frame unit, should be at a minimum equal to two segments, and on the average equal to ten segments.

The frame unit is the primary sampling unit (PSU) in the area frame sampling procedure. In locating a sample segment the frame unit is randomly selected first which contains the target segment. The SEGMENT which is selected from the frame unit in the second step of sampling and is some times referred to as the subunit or ultimate sampling unit in statistical texts.

The concept of frame units is important for one reason. By selecting a frame unit which will contain the sample segment and then dividing that frame unit into segments, frame work is greatly reduced. For example, a sample of 100 segments is to be selected from a frame with 100,000 potential segments, using single step simple random sampling. All 100,000 segments would have to be identified in advance of sample selection. By creating frame units of 10 segments each only 10,000 frame units would be required. If 100 of these were selected (assuming no frame unit contained two sample segments) only 1000 segments would be required to be drawn. A savings of 89% of the total work load is realized.

When frame unit construction is complete frame units must be numbered in logical and systematic fashion. Frame units should be number serially starting in the north east corner of the population and proceeding from east to west and working southward in a serpentine manner. While the starting point and the directions of numbering are arbitrary, the concept is to insure that as much as operationally possible, frame units which are most similar to one another are numbered consecutive. Even in the same strata two frame units which are adjacent are more likely to be similar than two which are separated.

The numbering pattern for frame units helps reduce sampling error inherent in the finished frame. This is accomplished by dividing the population into blocks of consecutively numbered frame units called PAPER STRATA. By associating neighboring units in a substrata, variance within that substrata and therefore the strata is reduced.

Paper strata are called this because this substratification is done only on paper (more recently in the computer). No work on the frame map itself is required. Sometimes they will be referred to in the literature as GEOGRAPHIC SUBSTRATA. This terms mean the same thing and may be more descriptive of the concept.

Another important advantage of using paper strata is that sampling with substrata allows that the sample can be divided between the various substrata. As the distribution of the substrata are somewhat uniform over the population it is assured that the distribution of the sample segments while random, will also have some uniformity in distribution.

As a point of clarification, COUNT UNIT is an older and less descriptive term for frame unit. In early area frame research stratification was based on the density of farm households. Count units were established by counting houses. This concept is no longer valid but the name persists. For our purposes it has exactly the same meaning as frame unit. Frame Unit is the preferred term.

The last major step for the frame construction technician in area frame construction involves the measuring of the area of each frame unit. This has been commonly accomplished using polar planimeter. With recent advances in micro computers it is now possible to digitize maps and compute areas in the computer. Digitizing is the process of translating spatial information from the map to a string of characters that the computer is able to process.

The area frame construction is complete when the assignment of numbers of potential segments to frame units is completed. This is a computer operation so that the frame technician does not usually 'see' this operation. The area of the frame unit is divided by the average target size for a segment in that particular strata. This number is then rounded to the nearest integer value for each frame unit. Rounding to whole numbers is required as it is not possible to deal with fractions of segments.

When a sample is selected, the area frame technician will be given a list of frame units from which specified number of segments must be drawn, usually one. This step: preparing segments and sampling material will vary as more than one process is available for segment identification. The final product is the same, a SEGMENT defined in such a way that it can be enumerated in a field survey.

In ASF the target population within a strata has been defined in terms of some fixed number of potential segments. The sampling procedure selects some number of these segments to be enumerated. This is a

critical concept in the understanding of Area Sampling Frame Theory. The population which is sampled is NOT THE AREA of the province or district, but the AGGREGATE OF THE SEGMENTS which were (potentially) created in the frame construction process. Potential segments is specified, because while the number is known, only a few segments are actually defined.

Another way to look at this important point is to consider: The expansion factor in an area frame sample is the number of potential segments in the frame divided by the number of actual segments in the sample. It is not the area in the frame divided by the area in the sample. Consider this carefully; misconceptions are common.

Note for mathematical statistician: Some statistical text books refer to ASF sampling design as multistage sampling. ASF is a special case of multistage sampling where the probability of selection of each ultimate sampling unit (segment) in a stratum is equal (Sampling Techniques Third Edition, Cochran, John Wiley & Sons, 1977, page 295). In this special case, formula for simple random sampling within each subunit (paper strata) apply. This sample design is termed multi-step sampling in the context of the ASF (Area Frame Design for Agricultural Surveys, Nealon and Cotter, USDA/NASS, August 1987).

USDA/NASS uses a replicated sample design, where multiple independent samples are drawn in order to facilitate sample rotation and sample size adjustments. This results in some reduction in sampling efficiency. The reduced efficiency is minor, particularly when compared to the added flexibility a replicated design offers.

THE PAKISTAN NATIONAL AREA FRAME

The Government of Pakistan ask for introduction of area frame sampling because current methods used in Pakistan were judged to be ineffective. Critical estimates were chronically late, and the precision of estimates was regularly called to question.

An area frame was developed in Pakistan on a pilot basis in the first phase of the Agricultural Data Collection project. Low level aerial photographs were used to stratify and divide the target area into sampling units. Standard 15 minute topographic maps were used to accurately delineate and measure the areas involved. This required the transfer of strata blocks and count from photos to maps. Frame unit area was then determined on the map. Survey material was prepared from enlargements of the associated photo.

It is important to note that the two uses of the photography; stratification and locating sample units for field enumeration are two very distinct and different operations. It was convenient that these two functions used the same basic input, aerial photography, but there is no requirement for this in area frame theory.

In National Area Frame construction the stratification medium will be SPOT Satellite imagery. The major advantage of SPOT is that it is current, and the color directly relates to the intensity of growing vegetation. With photography the intensity of land use was only implied from the presents of fields, irrigation canals, and other artifacts of cultivation. Usually stratification based on this kind of information is quite accurate, but problems did arise in saline soil regions.

The SPOT image used will be as current as possible. Usually the best timing for SPOT used for land use identification is during the peak of the most extensive growing season. Image dates of early September or January are probably most appropriate in Pakistan. Ideally the image used for the National Frame will be no older than from the previous September. Because the satellite only traverses a location about once every 16 days, one cannot precisely specify the required date of coverage. But by requesting a suitable range of dates one can obtain the needed imagery, weather permitting.

Stratification using SPOT is essentially the same process as with low level photography. The technician using his best judgment, as stratification involve subjective reasoning, divides the target area into 'blocks' of similar land use. The difference is the stratifier will not be able to locate precise boundaries for strata blocks on the SPOT. In the beginning he may have some difficulty fixing the location of the image on the map. This is because there are few man made landmarks which are large enough to appear on the image, and the topo map emphasizes only cultural features. However, with careful study and the use of topographic maps for reference one will soon discover many identifiable features.

Location of segments in National Area Frame will be done in a fashion similar [illegible] field located segments in the Pilot Frame. Where low level photog[illegible] is available the field process will be mimicked in the laboratory, all segments will be constructed around a randomly selected point.

The field location of sample segments has become commonly known as 'point sampling'. This is a misnomer as no points are enumerated. In geometry points have no area and this implies an infinite population. In ASF a known and discrete number of segments (potentially) exist. Quite different statistical assumptions apply in the two cases. A more accurate term to describe this type of segment preparation is 'field located segment', which does not have the implied assumptions.

The technicians involved in frame construction will become aware of a number of significant changes in frame construction techniques. Most obvious will be the addition of two pieces of high technology tools, the Zoom Transfer Scope (ZTS) and the Digitizer. The ZTS will be used to assist the stratifier in transferring data from the SPOT image to the base topographic map. The Digitizer is a tool which transfers map coordinates into digital data. The digital data can be input into a computer. The digitizer and associated computer programs will replace the planimeter. Much hand review and manual data entry will also be turned over to the digitizer and computer.

ZOOM TRANSFER SCOPE:

The Bausch and Lomb Zoom Transfer Scope, Model 53-05-30-10, will be used. This instrument has the capability of enlarging, rotating and translating photographs or other graphics material (SPOT). Spot may be viewed in superimposition with the image of a map or other output material, and significant data may be transferred. It also has the capability of stretching the input images to obtain a better match, as is required when the input material has distortion due to the angle of the 'camera'.

In pilot frame construction the integration of the photo and the base map was done in the framers mind. He studied the photo, remembering the location of the data (line and boundaries) to be transferred then moved to the map to record as transfer area for the data. All needed corrections was carried out, more or less unconsciously, by the framer.

The ZTS is a very sophisticated precision instrument, that performs one single simple task. It corrects and adjusts the scale of one image and allows it to be superimposed on to another. For our purposes this means we may look at the SPOT image overlaid onto a topographic map. This is only new in the sense that we are now using a machine to speedup and make more accurate a task that previously was done manually.

The ZTS consists of a frame or stand which houses and supports all of the electrical and optical components. Supported above the frame is a glass stage. Three illuminator lamps are attached to the frame, one for the base map placed below the frame, and two for the SPOT scene placed on the stage.

Electrical controls, found on the frame, include a master switch, and two flicker switches on either side of the optical housing, and three knobs which control the intensity of the illuminators. There is also a foot switch which duplicates the action of the flicker switches. Experience has shown that the foot switch will be used extensively.

All of the optics controls are located on the optics housing. A description of each its function and location follow:

1. INTERPUPILLARY DISTANCE SETTING: The distance between the pupils of the eyes varies from person to person. By gently separating or pushing the eyepieces of the ZTS together while looking at an image the field of view can be adjusted so that a full field can be seen with both eyes. There is a scale above the eyepieces which displays the interpupillary distance of the operator. Once a comfortable setting is found the operator may return to this setting latter.

2. EYEPIECE FOCUS CONTROLS: Each eyepiece tube has a knurled ring for focusing the eyepiece. To focus the ZTS for individual use both eyepieces should be turned all the way in and backed off about one quarter turn. The operator should focus each eye independently by closing one eye and making the needed adjustment. Then make small adjustments while keeping both eyes open until the sharpest image possible is obtained. The operator may or may not wear glasses as is most comfortable. Persons with serious astigmatism should wear their glasses, or may find that they are unable to properly align material in the ZTS.

3. ZOOM CONTROL KNOB: The zoom control knob, located on the lower right front of the optical housing, controls the magnification of the zoom system. When using SPOT images and standard topographic maps a zoom of about 4x will bring the spot into the same scale as the map.

4. STRETCH CONTROL: The stretch control, located to back right side of the optical housing, is used to stretch the image in such a manner that a square may appear to be a rectangle. The ratio of stretch is variable from 1:1 to 1:2.

5. STRETCH DIRECTION CONTROL: The stretch direction control is adjacent to the stretch control. It is used to set the direction of the stretch from 0 to 180 degrees.

6. FINE X AND Y ADJUSTMENTS: Coarse positioning of the image is done by moving the map beneath the instrument or by moving the SPOT image on the stage.

For fine adjustments, the map lens may be scanned over the map. The Fine x [illegible] Adjustment Screws are located above the map lens. If the end of [illegible] travel of the screw is reached and more movement is required, turn the screws into the middle of the range, reposition maps and then use the Fine x and y Adjustment Screws. When turning the screws outward do not continue turning after image has ceased moving, otherwise the screws will fall out.

7. ILLUMINATION CONTROLS: The illumination or light controls should be adjusted so that the intensity of the image and the map are satisfactory for the operator. The lights are controlled independently by dimmer switches on the left of the frame and a master switch on the right.

Illumination of the map or the SPOT image are controlled by flipper switches on either side of the optical unit and are duplicated in the foot control.

HANDLING OF SPOT IMAGES:

SPOT images will be received from SUPARCO in the form a color photograph negative. It will be on film about 10 inches square. It will be packaged in a clear plastic envelope. A contact print will be made from this negative. Contact prints are made be placing the negative directly on the print paper and exposing it to the light. The SPOT negative will then be careful stored. All stratification work will be done using the contact or positive print.

The emulsions, that is the chemicals which carry the color in photography, are extremely soft and scratch very easily. For normal use scratches on the surface of a photo are not very important. In stratification we will be viewing the SPOT positive images magnified many times. Even minor scratches will appear as major gashes in the landscape.

Because of the soft nature of the photography the negative should never be taken out of the protective cover, and the positive should be handled as carefully as possible. Particular care should be taken not to place it face down on any surface (except the viewing stage of the Zoom Transfer Scope). The face or picture surface should never be wiped or slid across any surfaces.

HANDLING OF MAPS:

Topographic maps will be the basis for the Pakistan National Area Frame, just as they were for the pilot area frame. More careful handling of the maps, however, will by required. The automation of frame construction requires that maps NOT be rolled folded, creased, smudged, tea spilt, or subjected to other abuses. Maps should only be handled with clean, dry, oil free hands.

Paste board file folders should be made for map storage. A hinge can be made using cloth tape. The frame topo maps should be stored in these folders. No more than about ten maps should be stored in one folder. This will help minimize the map handling while keeping an orderly file system. Of course the outside of the file folder should be anointed with its contents.

A common practice during pilot framing was for the framer to curl the map under the desk in front of him. This caused the paper to be slightly creased as it turned. This will not be tolerable during the new frame construction as the digitizer is very sensitive to small shifts in paper location on the digitizing tablet caused by these small folds.

Another practice that needs to be changed is the taping of the map edges. This was very desirable to help preserve the maps when they were handled over and over again. The goal now must be to handle a map the minimum number of times needed to get the job done. Also, the maps expand and shrink with changing humidity. The tape does not. For this reason on a very humid day the taped maps will not lay flat as they 'bubble' inside the taped edge. The slight change in map dimensions due to weather in itself will not effect our work as long as it can still be accurately aligned on the digitizing tablet. Therefore, the maps should not be taped so that they expand and contract uniformly.

The single overriding factor in all map handling is "How does it effect the digitizing process?". Digitizing is the most technologically sensitive operation in frame construction. Digitizing will greatly reduce the time involved in frame construction. It will greatly increase the accuracy of the finished frame, and it will do most of the error trapping that was previously done by hand. BUT, it requires precision products (maps) to work with. The digitizer normally works in an accuracy range of 0.1 millimeter or smaller. Even very small changes in the map surface will render a map unusable. It is imperative that maps be handled as little, and as carefully as possible. They should never be folded, creased, or rolled up. They should be handled only with clean, dry, oil free hands, and then as little as possible. Every flexure of the paper causes the structure of the paper to break down slightly. Paper and specially the paper used for maps is very durable, but for our purposes it must be treated with great respect. Any map which cannot be digitized will have to be recreated. All data processing will proceed from files which are automatically created in the digitizing process. There is no provision for manual input.

Remember the completed maps and the digitized record is the frame. It will have cost many millions of rupees to create and will have a life expectancy of twenty years or more. It should be treated accordingly.

A new procedure for topographic map numbering will be employed in National ASF construction. When the relatively small district was the target population in the pilot project all frame units in a district were numbered at one time. The topographic map numbers were used to maintain orderly files, but were not part of the frame identification

itself. In the National area frame this would require all stratification and frame unit construction completed before the next step could begin. This is clearly not possible and still maintain an orderly flow of work.

The first step in National ASF construction will be the numbering of the topographic maps in such a way that each map can be processed independently. All maps used in a province will be laid out in large 'mosaic' and number in a serpentine fashion. This will allow each map to be handled as an independent unit, but the completed frame will still be organized into a geographic distribution. The discussion of numbering of frame units which follows at a later stage in this paper applies equally here.

STRATIFICATION:

The first requirement of the stratification process is to develop stratification categories or stratum definitions. Each stratum type should be clearly recognizable from the material available. For National Area Frame Construction this is a combination of SPOT imagery and topological maps. As the primary use of this area frame is for agricultural surveys stratification should consider land use types which are important to agriculture. Area frames could be developed to measure nearly anything which can be associated with the land, but stratification would be much different if the distribution of coal deposits were the primary consideration.

The stratification definitions developed for the pilot frame are likely to usable in the development of the National Frame with little modification. One adjustment likely, is that it will be difficult to distinguish irrigated from unirrigated crop land with certainty. This should require the collapsing of some specific strata definitions into more general terms. The absolute detail will not be as apparent on the SPOT. This will cause some use types to be unrecognized when only small areas exist. This will not lead to difficulty as the smallest a strata block should ever be is about one half square kilometer.

It is beyond the purpose of this paper to address the interpretation of SPOT. SUPARCO will provide specialized training in image interpretation. However, satellite images use a false color scheme which is standardized to represent the intensity of the various light patterns the satellite receives. Generally growing plants appear red, bare soil is brown to yellow, and water is dark blue to black. Clouds will usually be white with a distinctive shadow. Recognizing these general pattern will aid the stratifier new to satellite imagery to begin to relate the image to the map or the ground.

The following outline sets out the general steps required to stratify a topographic map using the ZTS and SPOT imagery. No attempt is made here to set hard and fast procedures, as the best approaches will

evolve over time. These steps are offered as a starting point. Even before detail training in image interpretation the stratifier should practice scene alignment and the other skills outlined here.

1. Begin working as near to the center of the SPOT scene as possible for a particular map. This reduces the amount of edge distortion, which is significant along the outer edges of the SPOT

2. Locate a land mark on the SPOT in the area you intend to stratify. Center this area on the stage with the labels on the image pointing away from the operator. The rotation control should be set so that labels on the SPOT are correctly oriented in the viewer. Magnification should be set at 1X. This allows for the largest area possible in the field of view, which is oriented so that adjusting the position of the scene on the stage moves the image in similar directions. When the rotation control is not thus aligned it is hard to move the scene as it travel in the viewer is not as expected. Remember by using the foot switch or the finger controls it is possible to view the topo map or the SPOT scenes alone, and also to superimpose one on the other.

3. Center the topo map on the table, attempting to line up the landmark chosen on the SPOT with the same location on the topo in the center of the field of view. It is not necessary that the topo map be in the right side up. Place it in the most comfortable position to work. The SPOT scene can be 'rotated' by using the rotational control wheel. By using this control the scene can be given the same orientation as the map. Now slowly turn the magnification control to 4X, while watching the image in the viewer. The SPOT image will seem to move as it is enlarged. Adjust the position of the map as needed to maintain alignment. The SPOT and the topo should be now roughly aligned.

4. The near perfect alignment of the scene is critical to accurate stratification. By experimentation with the magnification control, stretch direction, and stretch distance control carefully align the central part of the field of view. Only the central two thirds of the field of view should be used as the outer edges have some addition distortion and is hard to see in any case. When making these alignment adjustments align as much of the entire field of view as possible. Again alignment is CRITICAL to a good stratification job.

A Hint: Sliding a sheet of heavy white paper over and the topo map has the same effect as turning off the table lamp. This has two advantage: 1) It blocks the map image completely, and 2) it allows for 'turning off' parts of the map only. This is useful in checking for alignment.

Remember the SPOT map is accurate and up to date. Things which are visible on the image are properly placed (given correct alignment of the scene), but things change and there are errors on the topo maps. When there is a difference the SPOT is right. In general roads, canals, and other man made features are the most accurately placed on the topo maps. Natural rivers change location and size over time. Mountain

peaks and ranges should be well located on the topo maps and of course move little. Keep in mind the possibility of placement errors (on the topo) when adjusting the alignment.

Stratification using SPOT has been tested in a number of areas in Pakistan by the ADC Staff using the Zoom Transfer Scope. Field checking was not done, but areas the ADC Staff were familiar with were used. The major difference the stratifier will notice is that the SPOT image carry much less information than the comparable photo. He should remember that the photo had far more information available than was actually used in the stratification process. Very accurate information on the intensity of current land use in available on SPOT. This is the criteria for stratification, all other data available on the photo was secondary to the stratification process.

5. Begin stratification from the center of the field of view working outward. It is necessary to work while watching your pencil through the viewer, but be careful. The field of view is magnified. Control and coordination of the pencil does take practice. When stratifying, decide what is the least represented strata in the field of view and block this area into the strata blocks. Remember to clearly identify the strata defined, i.e. 12, 31, etc. In no case should a strata block be smaller than one half square kilometer. The red grid squares on the map are one full kilometer square.

Note: This means that very few villages will be delineated as separate strata blocks. It is evident from the pilot frame construction that incorporating the village strata greatly increased the work load in frame construction, but did not significantly improve frame performance.

The Stratification procedure is identical whether low level photography or field located points will be used for sampling in the district. Procedures for defining the segment will be such that a segment is constructed around a random point in the laboratory on photography if available.

6. When each field of view is complete the stratification should be reviewed by a second technician. This review should be to verify and assure agreement on basic image interpretation. It should not be a review of the exact location of each line drawn on the map.

A note to the reviewer: Remember stratification is at least as much Art as Science. This means that there is no single 'right' way, and no two stratifier will do the job in the same way. Be very critical of stratification which is wrong (Remember non-sampling error). Be careful to distinguish from different.

7. When review is complete the scene and topo map should be shifted about 2/3 of the field of view and the process of alignment and stratification completed again. This process should be continued in a spiraling fashion around the first field of view completed until the entire map is stratified.

The stratifier should record the strata from one map on to the the edge of adjacent maps. This will help insure a smooth transition in strata from map to map, and help reduce the personal bias in map stratification. All strata blocks should end at the map boundary: there should be no overlap on to adjacent maps.

8. Completion of the stratification of a topo map should require about one day. The last operation that the stratifier should do is to review the map for completeness, and to correct obvious errors. There should be no incomplete boundaries. All lines drawn should terminates at the edge of the map or at another line. There should be no unstratified areas also, no holes.

With SPOT procedure it is not necessary that boundaries be observable on the ground, so the most convenient boundaries possible should be used to form frame units of the appropriate size. These will usually be the red grid lines that have been preprinted on the topographic map. Using these lines will help in digitizing and insure uniform sized frame units.

Frame units should next be numbered. Numbering should begin in the northern east corner of the map and precede in a serpentine manner right to left until all frame units are numbered. Each topo graphic map will be numbered independently. The numbering sequence should include all frame units regardless of the strata. Note: This is a deviation from the pilot procedure. Frame unit numbers will be a three element number. The elements include the ASF map number, the strata identification number, and serial number of the frame unit on the map. The number will be in the form of xx-yyyzzz where xx is the strata number, yyy is the ASF map designation, and zzz is the frame unit serial number. Only in a special cases will it be necessary to record the ASF map number on the frame unit as it is number. The map number will be normally added by the digitizer.

Each topographic map will stand alone. No count unit will overlap on to any other map. This will add an additional measure of quality assurance as the sum of all areas in all strata in a map will equal the map area (except where a district boundary transits a map). Although each map stands alone, the stratifier should review adjacent completed maps in an effort to maintain a smooth transition from map to map.

Every effort should be made to insure accuracy in frame numbering so as not to overlook a frame unit, or a number, or to duplicate numbers, extensive review is not required. The digitizing operation checks for all of the common numbering problems.

When frame unit numbering is complete the map should be dotted and otherwise prepared for digitizing. Dotting is a process of marking each end of all straight lines which make up a frame unit. Also dotted are direction marks each line intersection. A common problem associated with digitizing is that the dotter begins to make dots that are in reality small line segments. This makes it particularly difficult for the digitizer

to know exactly where the dot is intended. It is a common practice for the technician who digitizes to do his own dotting. This helps him know what was intended. There are usually more technicians than digitizing equipment so this does not slow the process.

Digitizing is done by placing the printed topo map security on the digitization table and mark each line segment, defined by the dots. A cursor with a cross hair sight is used. Following the 'prompts' displayed by the digitizing software on micro computer lead the technician through the process. The computer will ask the operator to 'mark' the dots which identify the frame units on the map.

No attempt is made here to explain or teach digitizer operation. Only a few observations are made. Digitizing requires a fairly high level skill. A number of operators should become proficient in its operation, but it is not cost effective to train everyone. Also, digitizing is a one person operation. Frustration levels rise, and data can be lost too quickly to justify having more than one person pushing the buttons at a time. Finally, proper file handling habits should be carefully cultivated so that data is never 'misplaced'. It should go without saying that important data file should be backedup on tape at frequent intervals.

When the digitizing is complete and all errors reconciled the digital file (in the computer) merged' with all of the other digital map files for the province. Area frame construction is complete. The combination of the maps and the digital data files, properly stored and backedup, is the area frame.

Sample allocation and selection is the next step. These topics will be covered in separate papers.

DEFINITIONS AND ACRONYMS

The following list of words and acronyms are presented to assist in the understanding of some of the jargon developed around remote sensing in agricultural research. Not all of these terms are used in this paper, but may be found commonly in remote sensing literature.

AgRISTARS — Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing. A USDA and NASA joint project.

cursor — An aiming device, such as a lens with cross hairs used in digitizing

digital computer — A computer that operates on the principle of counting as opposed to measuring

digital image — An image having numeric values represent color intensity

digitizer — A machine that changes graphic cartographic information into numeric form

false color — A color image prepared by assigning blue, green and red colors to the three or four frequent bands which the satellite receives producing colors similar infrared photographic film

FOV — Field of view

LACIE — Large Area Crop Inventory Experiment. An early USDA/NASA exercise.

LANDSAT(ERTS) — Land Satellite, Earth Resources Technology Satellite

grain — One of the discrete silver particles resulting from the development of light sensitive film

gray scale — A monochrome representation of a color image, a monochrome strip of shades ranging from black to white used to balance colors in color photography

image The recorded representation of an object produced by optical or other means when light (or other emitted energy) is not directly recorded on film

MSS Multispectral Scanner

negative A photograph on film, plate, or paper in which the tones are reversed

NOAA National Oceanic and Atmospheric Administration

pixel Picture Element; The area on the ground represented by a data value; the smallest area 'visible' on a satellite image

positive A photograph on film, plate, or paper having approximately the same rendition of light and shade as the original subject

resolution The ability of an optical system to distinguish between signals that are spatially near or spectrally similar

scene everything that occurs before the satellite sensor; the image recorded by the satellite sensors

signature Any characteristic or series of characteristics by which a material may be recognized

spatial pertaining to the location of, proximity to, or orientation of objects in respect to one another

SPOT Systeme Probatoire D'Observation De La Terre

thematic map, TM The graphic display of data from a geographical data base

PN-ABW-341
95462

PAKISTAN'S

AREA SAMPLING FRAME

NEED FOR IMPROVING PAKISTAN'S CURRENT AGRICULTURAL STATISTICS

1. Pakistan's current method of estimating area and production, the Village Master Sample, employs sampling procedures based on the Land Revenue System.

2. The current system of collecting agricultural statistics does not fully meet the requirements of planners and policy makers.

   a) crop estimates are not available on a timely basis

   b) reliability of the estimates is not known

3. As a result, the Government of Pakistan asked USAID to provide technical assistance in improving the existing system of agricultural statistics .

4. A U.S. team recommended a new technology known as the "Area Sampling Frame" which has been successfully used in the U.S as the primary method of collecting acreage, yields and other kinds of data since the early 1950's.

WHAT IS AN AREA SAMPLING FRAME ?

An Area Sampling Frame, such as the one being implemented in Pakistan, includes all of the land area in the country without overlap or omission.

This land is stratified into categories, according to its use, by using spot satellite imagery and topographical maps.

Once the land is stratified by it's use for a province or the entire country, probability samples of land areas (called segments) can be drawn from each land use stratum to conduct various surveys and arrive at estimates for a province or the entire country.

WHAT ARE THE OBJECTIVES OF THE AREA SAMPLING FRAME?

To modify / replace the VMS data collection system by:

1. Providing statistically reliable estimates of acreage and production for major crops on a timely and cost effective basis.

2. Making possible the use of growth models for forecasting yields of major crops before harvest, so early production estimates can be made.

WHAT ARE THE ADVANTAGES OF THE AREA SAMPLING FRAME COMPARED TO THE EXISTING SYSTEM?

1. It is completely independent of all other sources such as revenue data.

2. It is more timely for both acreage and production estimates.

3. It is cost effective.

4. It is statistically reliable

5. It is more flexible

PROGRESS TO DATE

1. **Phase I**

   a. 1976 aerial photography was obtained in 1985 and used to stratify four districts in the Punjab and three districts in the Sindh Province.

   b. acreage and objective yield surveys have been conducted continuously in these seven districts since 1987-1988.

   c. couldn't get more (or recent) aerial photography

2. **Phase II**

   a. permission received to obtain spot satellite imagery for the entire country.

   b. stratification got underway for the entire country in 1989 using the satellite imagery.

   c. surveys are continuing in the seven districts until all land in Pakistan has been stratified using the satellite imagery.

THE STEPS INVOLVED IN THE CONSTRUCTION OF AN AREA SAMPLING FRAME

STEP 1

The target population is first identified -- in Pakistan it is the province.

STEP 2

Homogeneous land use categories (strata) are defined and identified.

(see the following strata definitions)

STRATA DEFINITIONS FOR PAKISTAN'S AREA SAMPLING FRAME

Stratum 10 - Intensive agriculture, 60 to 100% cultivated.

Stratum 20 - Extensive agriculture, 30 to 59% cultivated.

Stratum 31 - 1 to 29% cultivated, and/or pasture.

Stratum 32 - River bed

Stratum 41 - Land not suitable for cultivation or pasture

Stratum 42 - Land reserved for National security

Stratum 51 - Urban areas

Stratum 52 - Villages and other agri-urban areas

Stratum 61 - Forests or concentrations of large trees

Stratum 62 - Orchards or other concentrations of tree grown crops

Stratum 70 - Water

STEP 3

Homogeneous land use areas (strata blocks) are delineated on topographical maps using satellite imagery.

(see Figure 1)

PROVINCE

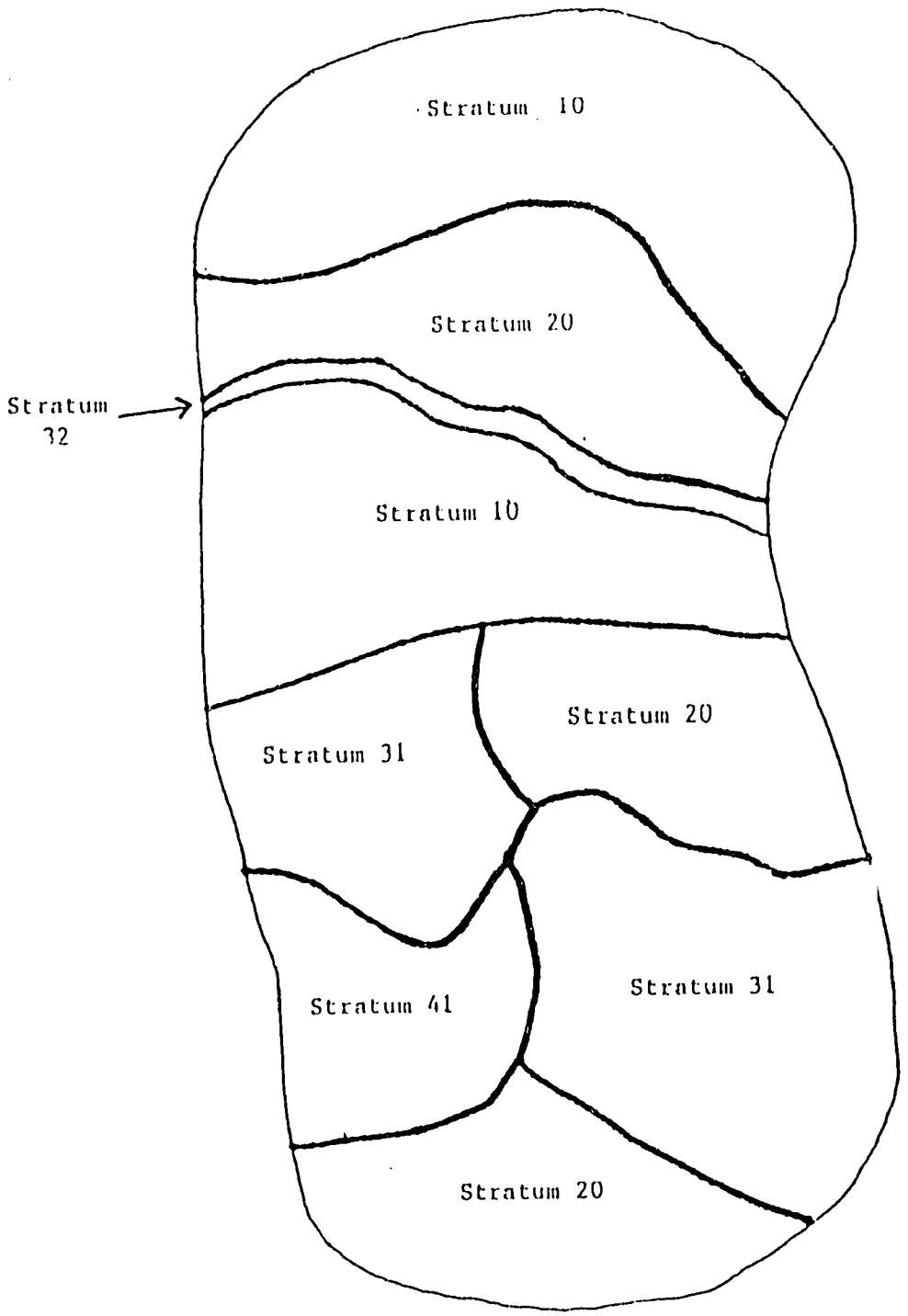

Figure 1- Example of delineating homogeneous land use areas (strata blocks)

Stratum 10 - Intensive agriculture, 60 - 100% cultivated
Stratum 20 - Extensive agriculture, 30 - 59% cultivated
Stratum 31 - 1 - 29% cultivated and or/pasture
Stratum 32 - River bed
Stratum 41 - Land not suitable for cultivation or pasture
(Note: There are several other strata not shown here)

NOTE: In theory, all land in each stratum is next divided into small areas (called segments) on topographical maps.

(see Figure 2)

PROVINCE

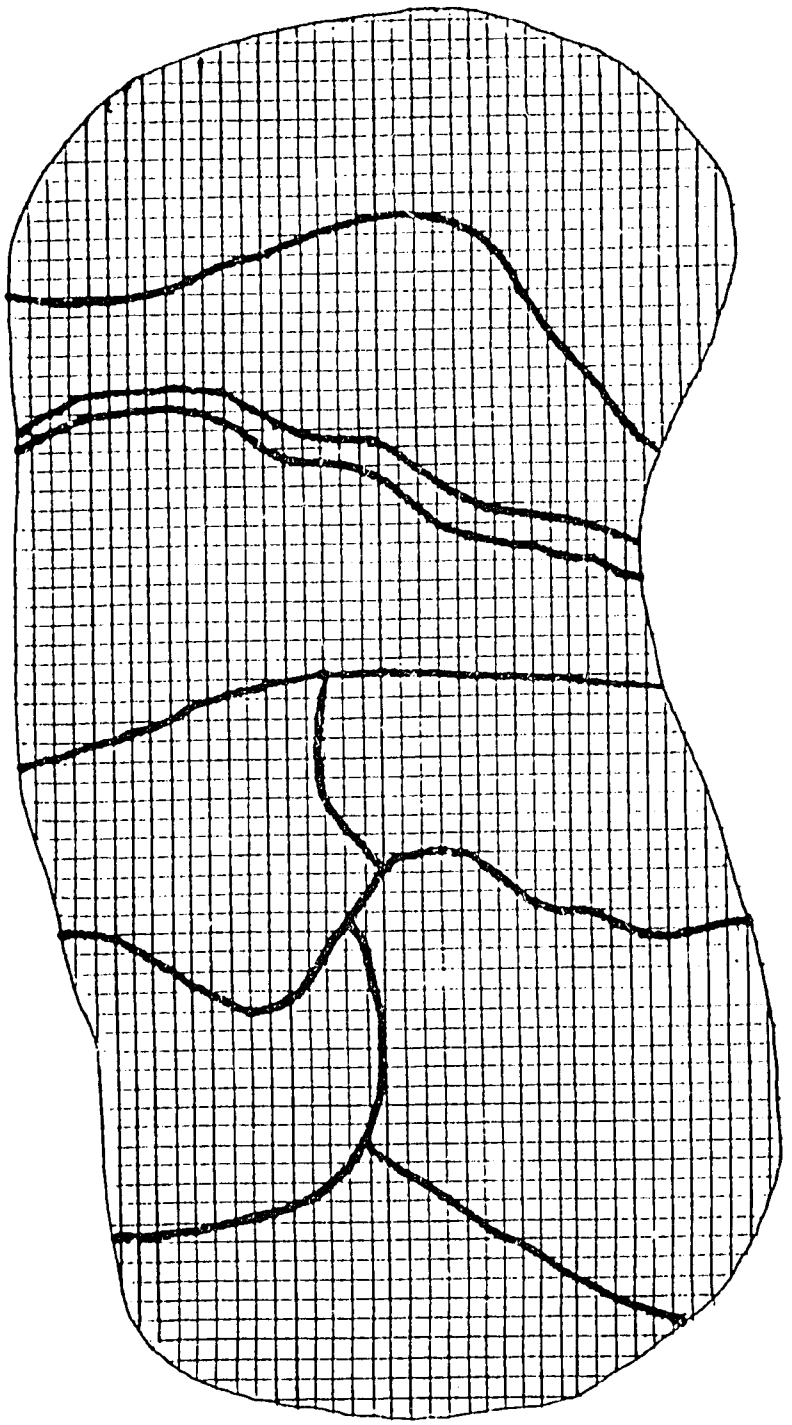

Figure 2- Example of delineating entire land area in a province into segments.

(Note: This takes too much time since each small segment must follow boundaries that can be found by enumerators -- such as roads, water courses, etc.)

STEP 4

Since dividing the land in each stratum into small units (segments) takes too much time, we divide it into larger units (called count units).

(see Figure 3)

PROVINCE

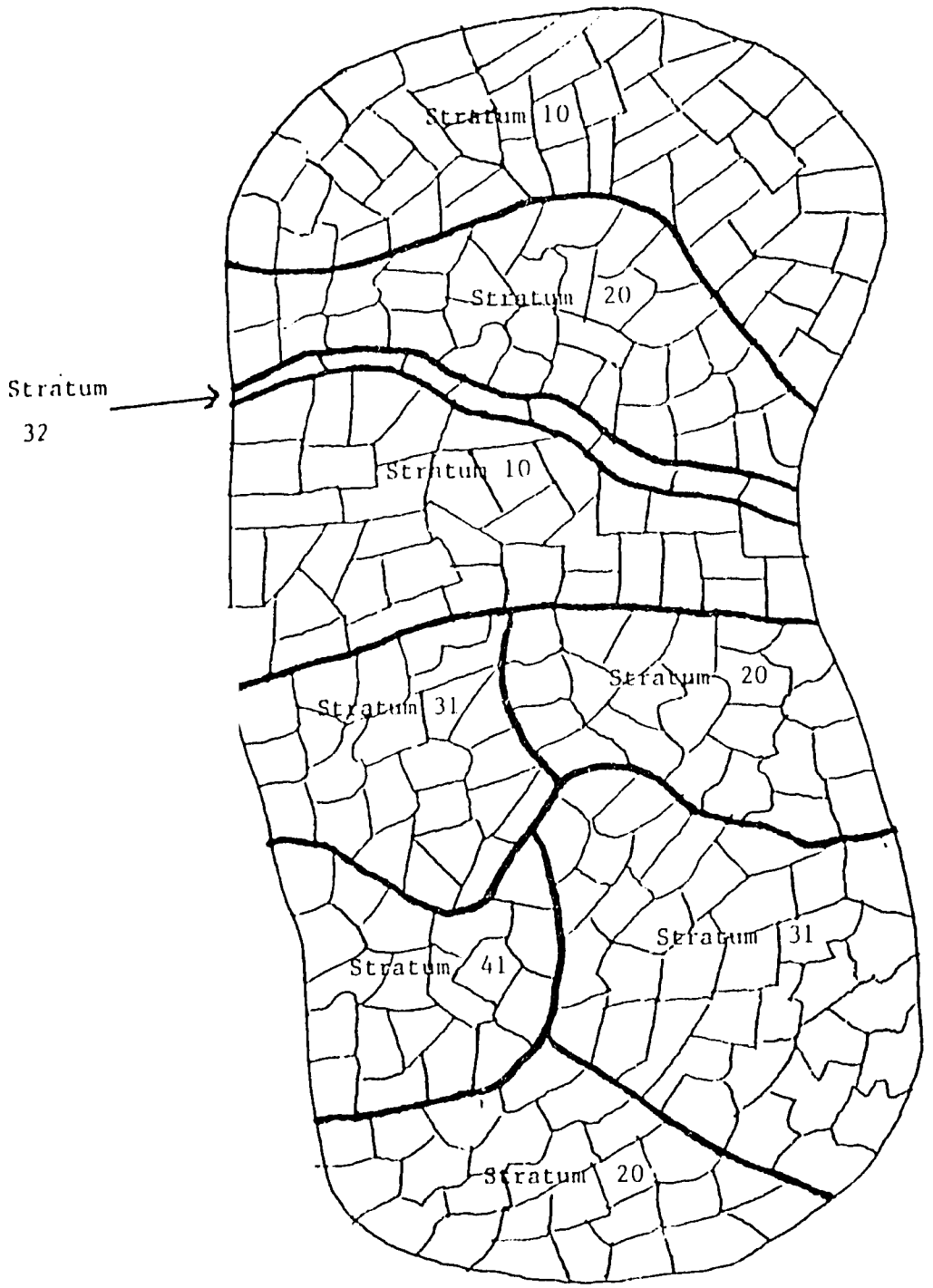

Figure 3 - Example of delineating strata blocks into count units.

(Note: Since it takes too much time to delineate all of the segments in a province, count units are delineated instead. In this example, each count unit contains 10 segments).

STEP 5

A sample of count units is randomly selected into the sample.

(see Figure 4)

PROVINCE

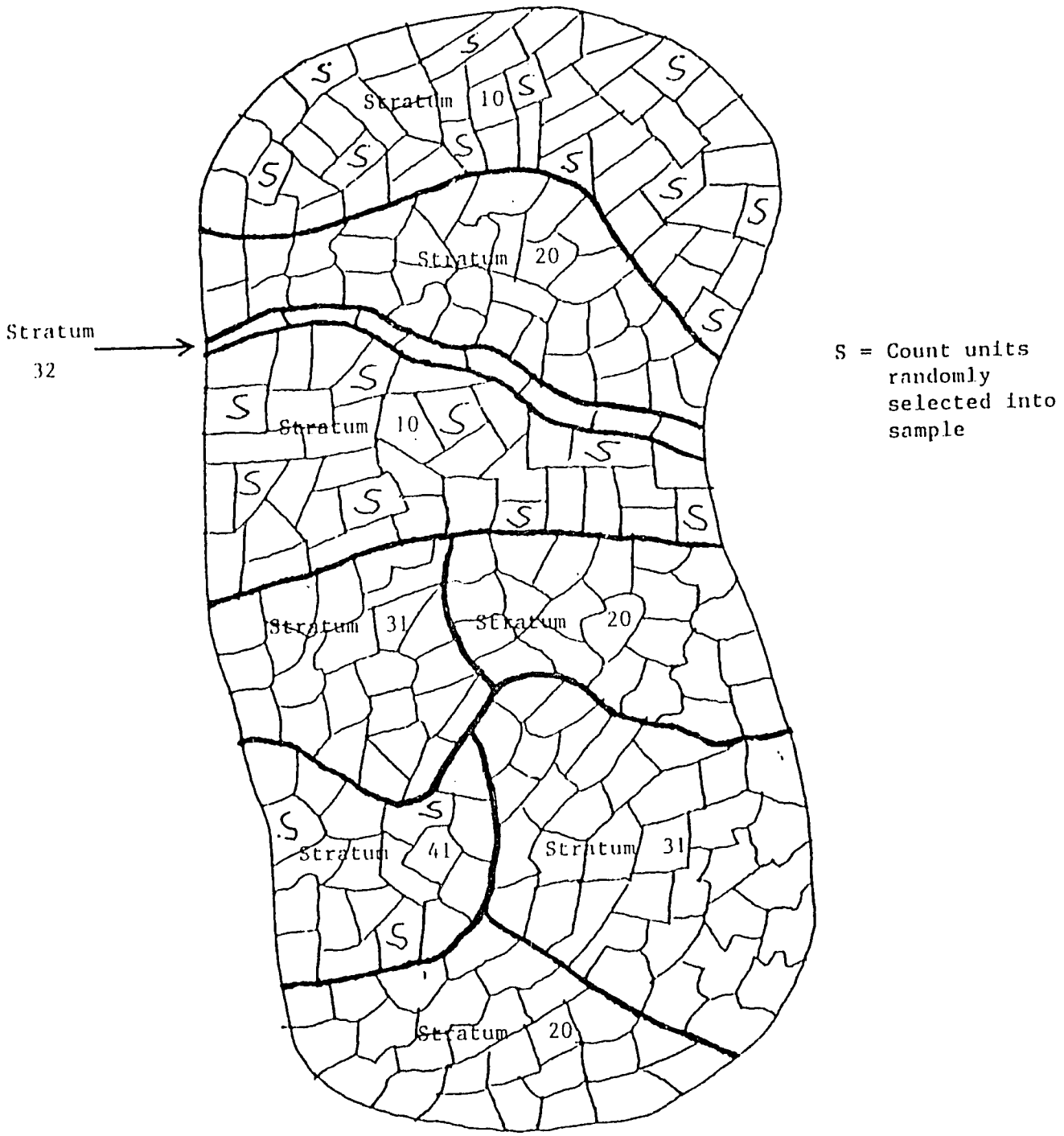

Figure 4 - Example of randomly selecting a sample of count units

STEP 6

The count units are then delineated into segments

(see Figure 5)

PROVINCE

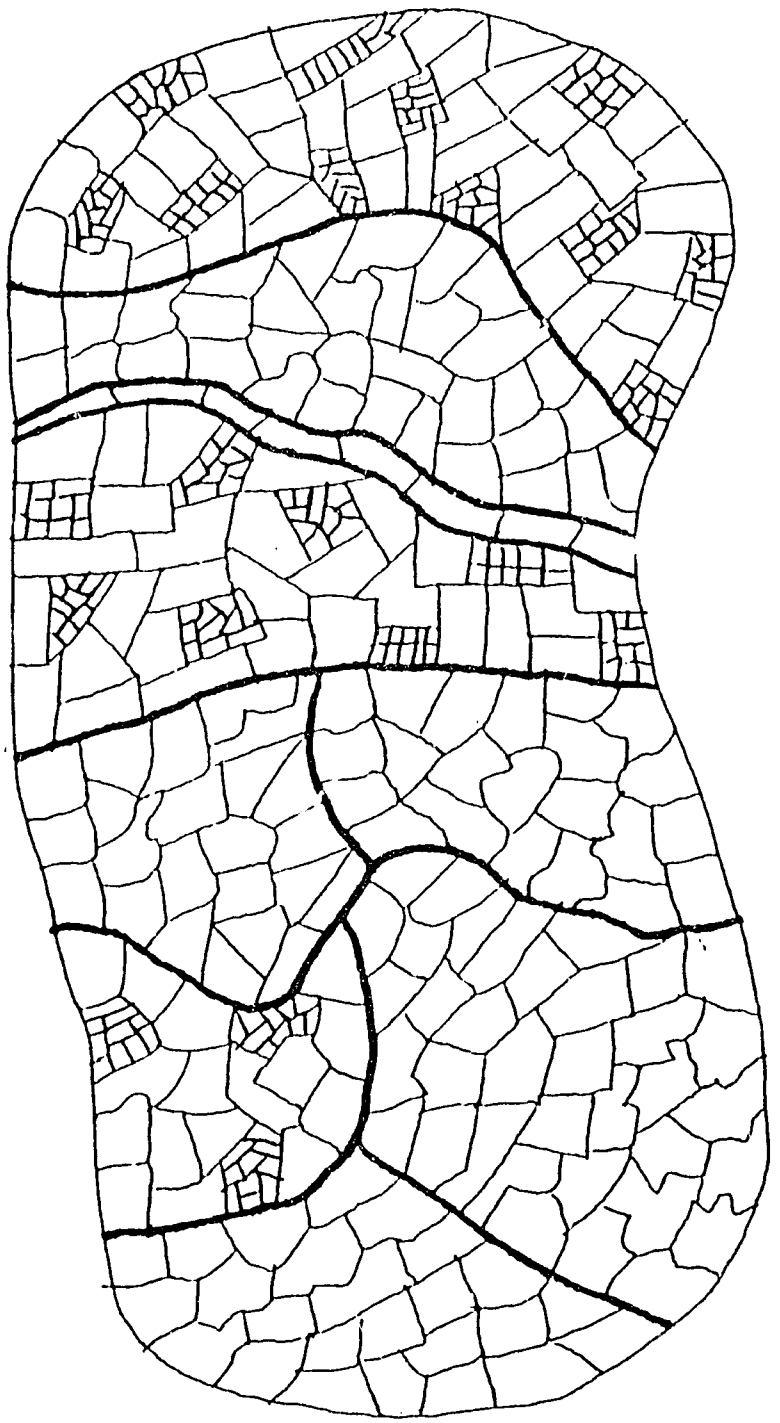

Figure 5 - Example of delineating selected count units into segments

(Note: In this example, each count unit contains 10 segments)

STEP 7

One segment in each selected count unit is randomly selected as the sample segment

(see Figure 6)

PROVINCE

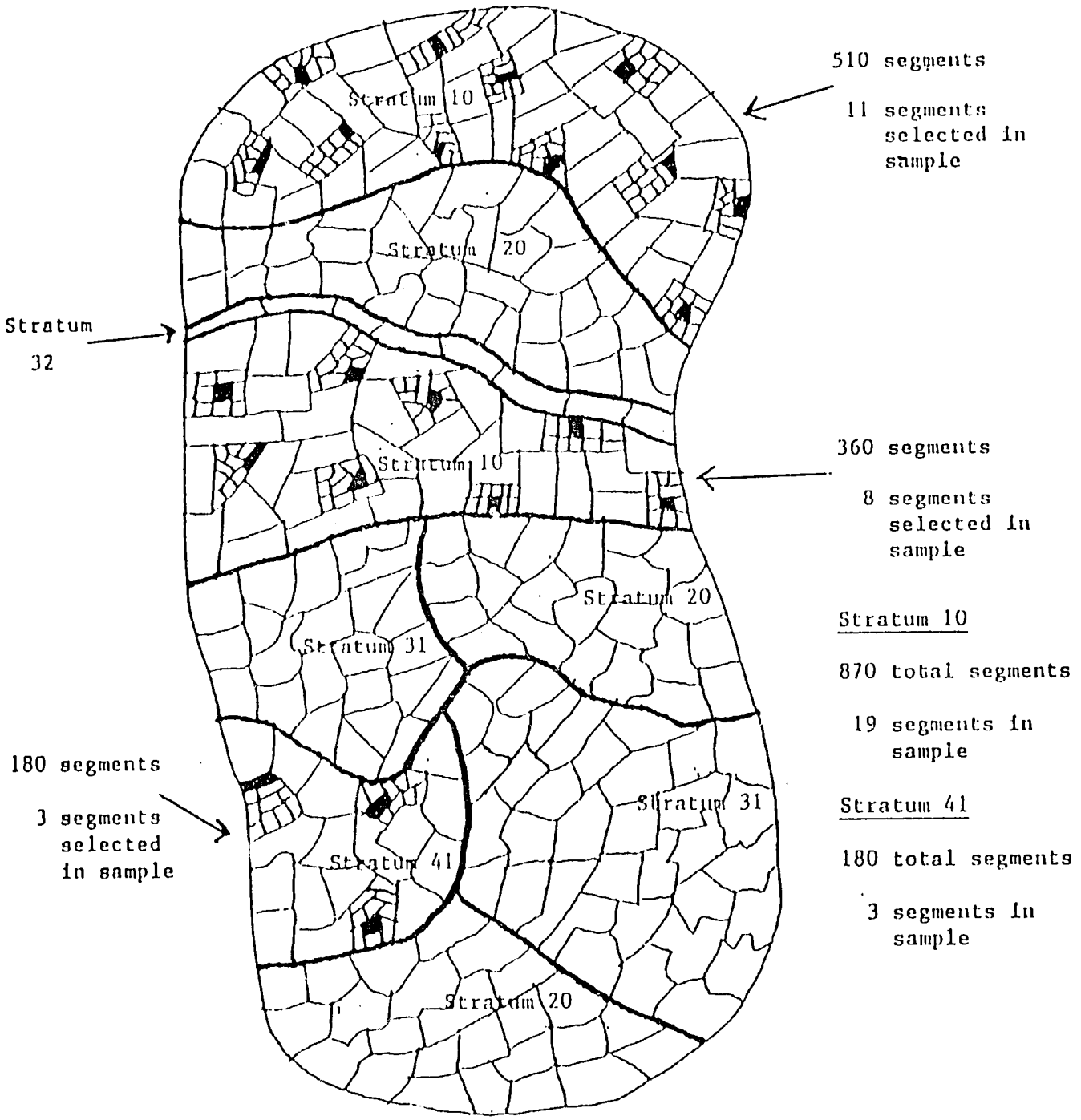

Figure 6 - Example of randomly selecting one sample segment in each randomly selected count unit.

STEP 8

These sample segments are then used for conducting various surveys.

HOW ARE THE SAMPLES DRAWN?

1. The best available estimates of variance for the major crops are used for computing sample sizes.

2. When the sample size for a stratum is determined, several separate simple random samples of segments are drawn from that stratum which will sum to the predetermined sample size.

   This method of sampling is called replicated sampling and has two major advantages:

   a. ease of sample rotation

   b. easy computation of variances for the data collected.

3. Samples will be drawn at the province level since the target population is defined as the province.

HOW ARE THE DATA COLLECTED?

1. Enumerators in each province will collect the data from farmers by personal interview when possible.

2. Enumerators from each province will be trained by FBS before each survey.

3. Supervisors in each province will conduct quality control visits to the same segments visited by enumerators and make checks to see if farmers were interviewed and the data accurately recorded by the enumerators.

HOW ARE THE SAMPLE DATA EXPANDED?

The sample data collected from the segments in each stratum can be expanded to the stratum level. The expanded data for each land use stratum can then be summed to arrive at estimates for a province, or the entire country.

(see Figure 7)

PROVINCE

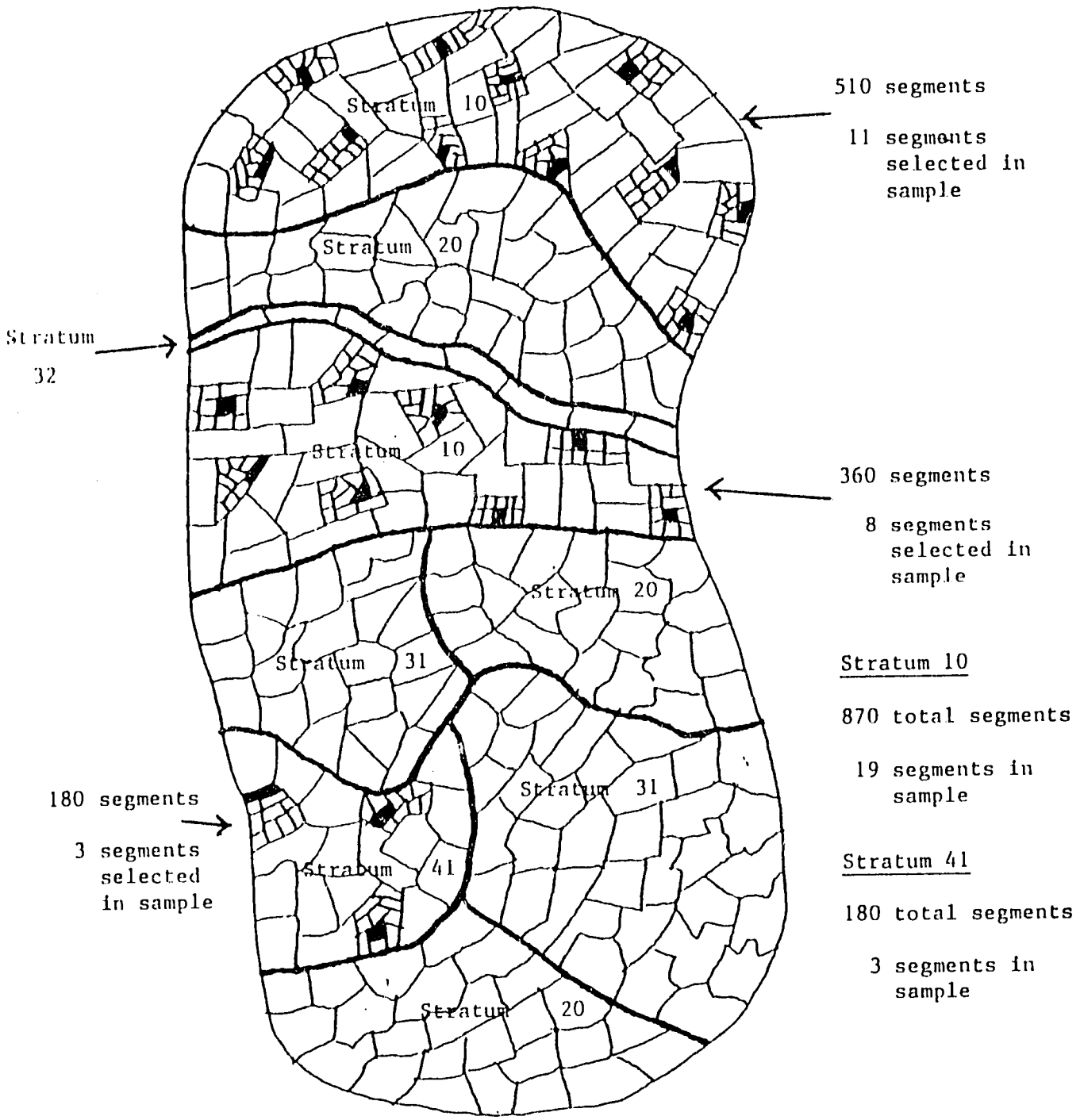

Figure 7 - Example of expanding wheat acreage found in sample segments by enumerators to stratum levels for province.

Stratum 10 expansion

Number of wheat acres found by enumerators in sample segments multiplied by 870/19 = total wheat acres for stratum 10 in province.

Stratum 41 expansion

Number of wheat acres found by enumerators in sample segments multiplied by 180/3 = total wheat acres for stratum 41 in province.

HOW ARE THE DATA TO BE PROCESSED AND TRANSMITTED?

1. Acreage data collected by enumerators will be manually edited, keyentered by computer specialists, re-edited, and summarized at the province level.

   Each province will have a laboratory for processing objective measurement raw data collected by enumerators. This will be manually edited, keyentered, re-edited and summarized at the province level to produce yield indications from which production estimates can be made.

2. Provincial estimates of acreage and yield will be prepared from the summarized data at the province level and approved at the province level. The estimates will then be transmitted to FBS in Islamabad.

3. All provinces will transmit the data to FBS according to a crops calendar schedule.

4. FBS will analyze the information submitted by the provinces and prepare estimates at the National level.

FINAL IMPLEMENTATION

1. The ASF has been used in seven districts of the Punjab and Sindh Provinces for several years. Existing ASF surveys will continue in these seven districts until the frame work for the entire area of the country, using spot satellite imagery, is completed.

2. Stratification and sampling will be completed in the Balochistan Province by April 30, 1991. The first survey will be conducted in this province as soon as field located segment work is complete.

3. The ASF project, based on spot satellite imagery, should be complete and operational for the entire country of Pakistan by June 1994.

PN-ABW 341
95462

TRAINING DOCUMENT FOR AREA FRAME CONSTRUCTION USING SPOT IMAGERY

Prepared by:

The Food Security Management Project, USAID
Agriculture Data Collection Component

November 30, 1989

51

TRAINING DOCUMENT FOR AREA FRAME CONSTRUCTION USING SPOT-IMAGERY
30 November, 1989

Area Frame Construction With Spot Satellite Imagery

Area sampling frame (ASF) construction has traditionally been completed with aid of aerial photography and topographical maps. This was true until the mid 1970's in the United States and has been true until now in the case of Pakistan. Beginning in the mid 1970's the Area Frame Section of the National Agricultural Statistics Service (NASS) began using satellite imagery as an aid in the ASF construction process. The imagery was used to outline strata blocks which were later delineated on the aerial photography and then transferred to the frame maps. This was all the imagery could be used for due to the 80 square meter resolution of the LANDSAT satellite.

Starting in the early 1980's a second sensor was added to the LANDSAT satellite with 30 square meter resolution. The 30 meter imagery caused a change in the NASS area frame construction process. Currently the satellite imagery is used as an integral part of the frame construction process. Photo mosaics are no longer built. All strata blocks and primary sampling unit delineation is made directly on the frame map. The 30 square meter resolution satellite imagery is used for delineation of the strata blocks and many of the primary sampling unit boundaries. Current aerial photography is then reviewed for areas where primary sampling unit boundary delineation was not possible using the combination of the satellite imagery and the frame map. The previous sentence needs some clarification for the sake of understanding the complete use of the satellite imagery. The frame maps (county road maps) are used extensively for boundary selection. That is, if a road is on the map the person stratifying is free to use that road for a permanent boundary. Therefore the aerial photography is used to identify boundaries of the primary sampling units where roads cannot be found on the map.

This brings us to The PAKISTAN AREA SAMPLING FRAME and the use of SPOT satellite imagery. Two differences should be pointed out between the United States area frame and the Pakistan area sampling frame currently under construction. The two differences are 1) the frame maps (topographical maps for Pakistan vs. county maps for the United States) and 2) the satellite imagery (SPOT for Pakistan vs. LANDSAT for the United States). Both of these items have a bearing on the frame construction process.

The topographical maps used by Pakistan give much more detail than the county maps used in the United States. This detail gives the person stratifying a much greater ability to identify permanent boundaries needed for the area sampling frame. The county maps give only roads as permanent boundaries (including railroads) which can be used without question in the frame con-

52

struction process. Other boundaries such as streams and rivers (natural waterways) can not be trusted due to changes in waterway courses over time. The topographical maps are superior over the county maps for the boundary identification process. Pakistan has many man made waterways, canals and feeder canals, as well as other permanent boundaries identified on the topographical maps which can be used in the boundary delineation process for strata blocks and primary sampling units. With Pakistan's long history of irrigation and land ownership systems, we can trust the topographical map boundaries. The satellite imagery will play a key roll in the detection of land use changes. As an example, a satellite image was reviewed and compared against the topographical map. A forest was noted on the map, while we identified the area as barren soil on the SPOT imagery. The imagery has shown a land use change since the topographical map was published.

The second item noted above as being different from the United States was the type of satellite imagery. As noted, Pakistan will use SPOT imagery with 20 square meter resolution compared to the 30 square meter LANDSAT imagery used in the United States. The SPOT imagery's increased resolution will give the person stratifying more information as compared to LANDSAT. This increased resolution will enable the person stratifying do a better job identifying the intensity of cultivation, resulting in a statistically more efficient frame.

This brings us to the changes in the ASF construction process from PHASE I to PHASE II. The strata boundaries and primary sampling unit boundaries will be delineated on the frame map without first being marked on aerial photography or satellite imagery. Boundary selection will be the same as in Phase I for the strata blocks and primary sampling units. The difference is that the topographical map will be relied on heavily for selection of the boundaries. It should be noted that the satellite imagery should still be reviewed for as many boundaries as possible. However the satellite imagery's first purpose is identification of the correct strata for the primary sampling units.

The last point that needs to be addressed is scale of imagery. What scale of satellite imagery should be used? The answer is the scale which the Statistical Officers feel most comfortable using. This brings us to the point that there are three operational methods and one method being researched in the United States currently for working with the satellite imagery. The three operational methods are: 1) A direct side by side viewing of the satellite imagery and the frame map. 2) The frame map is made into a transparency and is overlaid on the imagery. Boundaries are drawn on the transparency and latter transferred to the frame

53

map. 3) The satellite imagery is projected on the frame map using the zoom transfer scope. Each of the three methods have advantages and disadvantages. Each of the three methods require different scale imagery. The method being researched displays the SPOT digital data on a screen and stratification is completed on the screen.

The direct side by side viewing is the "traditional" method of frame construction and requires as large of scale imagery as possible.

The second method of creating an overlay and placing it on the frame map is used primarily for delineating the strata blocks and some of the primary sampling unit boundaries using roads on county maps. This method uses scales of 1:100,000 and 1:250,000. If 1:250,000 or 1:100,000 scale imagery is used grids should be developed at the scale of the satellite imagery to measure the amount of cultivated land in any questionable primary sampling units to ensure the primary sampling unit is assigned to the correct strata.

The zoom transfer scope can be used during delineation of strata blocks and primary sampling units. One of the most important advantages of the zoom transfer scope is the ability to accurately record changes directly onto the frame map. Since the zoom transfer scope projects the satellite image onto the map, changes in river courses, new highways, etc. can be placed accurately onto the map ensuring correct measurements for primary sampling units. Note, the measurement of primary sampling units affect the probability assigned to the primary sampling units and thus affect the statistical properties of the frame.

Review of SPOT Imagery

General

The SPOT images are printed in the standard false colors. The SPOT band 1 is printed in blue, band 2 is printed in green, and band 3 is printed in red. Band 2 is best for showing roads and bare soils. Band 2 heightens the contrast between vegetated and non-vegetated areas. Band 3 is the near-infrared band which is printed in red. This band, band 3, helps evaluate plant matter quantity and vigor of vegetation. It separates water bodies from vegetation. It is not as effective as band 2 for road detection.

Several items about the images should be kept in mind. The color tones in each image are made up of two different components.

54

First, the colors are a product of the three different bands. Since the bands are different shades of blue, green, and red, these three colors interact to produce the color on the negative. The second item to remember is that the printing process greatly affects the final image color hues. This point was brought out from the test scenes which SUPARCO provided. The two scenes were from the same location, same sensor and same date, that is, only different prints of the same image. One image was almost impossible to work with because of the blue tint of the scene, while the second scene had the red hue and was easy to work with.

There are many items that can make a difference on the scene appearance. A list of items would include time of year (different sun elevations), amount of moisture in the soil and stage of crop development. The scenes have been ordered to insure the crops are in the maximum growth stage. The other items are not controllable, however they do affect the appearance of the scene.

As noted above the actual printing of the images affect the usefulness of the prints. Also it should be noted that the printing process is somewhat of an art. Suparco should receive feedback on the usefulness of the prints and be apprised of any necessary changes in printing.

I. BASIC IMAGE CHARACTERISTICS - Color

Keeping in mind the above discussion on how printing affects the color hues the following information is given.

WATER

Water appears blue to black. The cleaner and deeper the water, the darker the blue color. Lighter colors are usually the results of high turbidity, not due to temperature.

CITIES AND OTHER URBAN AREAS

Cities and other urban areas appear light blue - gray to bright blue. Concrete is usually white or very light blue. Asphalt is very dark blue or black. Therefore, roads can appear from black to white depending if the road is concrete or asphalt. It should be noted that the cities will probably appear a lighter blue - gray in Pakistan due to the concrete construction. Homes with asphalt shingles as in the U.S. would appear a darker color.

CLOUDS

Clouds are white. The clouds shadows should be black, and have

55

the same shape as the clouds they are next to. Distance from the clouds to their respective shadows should be similar for all clouds in the scene. The sun location for Pakistan in relation to the satellite should be in the east to southeast. The cloud shadows should be located on the west to northwest side of the cloud. Be carefully in identifying clouds, they should appear similar to the salted out (saline) soil areas.

BARE SOILS

Soil colors vary greatly. They can be shades of green, gray, blue, brown, or white. Moist soils are darker than dry soils. Bare rock is often brighter than more developed soils. The salted out soils will appear almost white, with a blue - gray hue.

VEGETATION

As noted earlier growing vegetation reflects in the near infrared band which is printed in red. Hence, growing vegetation appears as red. The brighter the red the hardier (more dense and vigorous) the vegetation.

COLOR VARIATIONS

In addition to the color variations noted in the printing process, color variation can occur due to recent rainfall. Recent rainfalls in a dry period will make soils darker (moister) and vegetation redder (greener). This means no single color key will work for all scenes, for all dates.

II. Native Vegetation

Very large areas lacking obvious field patterns are likely to be native vegetation. Color and texture will be similar throughout the area, but close inspection will usually reveal a somewhat mottled appearance.

A. Marshes - Marshes are found in lowlands, along streams and rivers. They generally appear deep, dark red. A red vegetation signature is mixed with a dark water signature to produce a darker color than that of dry land vegetation. Areas with a dense tree canopy will appear brighter red (like a forest), mottled with deeper red marsh signatures. Salt marshes and land frequently inundated (wetlands) will be very dark, with very little red.

B. Forests and Brush Land - Forests also appear red in summer. In areas of sufficient rainfall, forests will be found in

56

uplands and mountains. Terrain often makes these areas unsuitable for agriculture, unless population pressure is large enough to press marginal land into cultivation. In more arid regions, forests are limited to stream valleys where moisture is available. Brush land is usually found in sub-climax ecosystems (like abandoned fields) or at the fringes of forests in arid regions, where moisture is not sufficient for trees.

Forests appear a deep red due to shadows from tall trees. Evergreen forests may be distinguished from deciduous forests by looking for "green-up" (turning red) of the deciduous trees in the spring. Evergreen trees should appear red throughout the winter, deciduous trees will be gray-brown after leaves drop in the fall. NOTE: If an evergreen undergrowth is present (mountain laurel, etc.), there will still be some red in the winter. However, it should not be as bright as a summer red.

Natural stands of trees should have a mottled appearance. Shade from large trees suppresses growth of new trees, resulting in a large crown, surrounded by less dense vegetation. As large crowned trees die, growth of surrounding vegetation increases. This creates a mottled appearance.

Extremely dense, homogeneous red forests in smaller areas are possible tree crops. It is unlikely such a homogeneous texture would occur without some kind of cultivation. These trees may either shade crops below (coffee), or produce fruit (bananas).

C. Grasslands - Prairies and grasslands are usually flat or rolling areas in regions too arid for a climax forest to develop. Grass should also appear red in summer, but not as dark as forests (due to forest shade). Native grasses should have a appearance similar to cultivated crops such as wheat, oats, and barley. These cultivated crops will be brighter due to fertilization, irrigation, and pesticide use to increase productivity. They will also show a soil signature sometime during the year, in preparation for planting. Unimproved pasture will appear much the same as native grasses.

Improved pasture includes those fields fertilized or irrigated, and those planted to hay crops for forage. These improved pastures will be a brighter red than unimproved pasture and native grassland. Hay may be planted for forage, or it may be cut and baled. Hay is green when cut, it should have no ripening signature before harvest.

D. Beaches - Beach sand appears bright white. Dune vegetation (dune grass, sea oats, etc.) appear mottled browns and reds. Few

51

cultivated crops are sufficiently salt tolerant to grow in these areas. Areas further in from the ocean will support growth of pines, appearing red and mottled red brown.

III. Crops

A. Basic Crop vs. Non-Crop - Crops can be distinguished from non-crops by the following:

1. Field Patterns - Cultivated areas shows field patterns extremely well, compared to native vegetation. However, field sizes can be deceiving, many small fields separated by narrow, dirt trails may appear similar to large fields.

2. Planting - A soil signature should be present at least once during the year, if an annual crop is planted. Soil signature should be visible from field preparation before planting, through planting, until plants have emerged sufficiently to obscure the soil signature.

3. Crop Development - Annual crops have different signatures as they grow, ripen, and are harvested.

B. Crop Development

Most annual crops go through the same 7 stages from soil preparation to harvest. The colors for the stages very from crop to crop; due to physical characteristics of the plant, cultural practices, soil types, physical and climatic conditions. Also, colors for each stage very for the same crop in different locations on the same scene. How healthy a crop is will affect its physical appearance, giving a different signature.

Those factors which affect the signatures of soil and native vegetation can also very the signatures of crops in different stages of their development.

These 7 stages are:

1. Plowing: Field work in preparation for planting results in a dark soil signature. Freshly plowed soil will be moisture than soil which has been exposed to the sun for several days. Most fields are plowed sometime before planting.

2. Pre-Emergence: The plant is beginning to emerge from the soil. The appearance is still a soil signature, but lighter than that of fresh plowing. The plant has not developed enough leaves to obscure the soil signature.

3. Emergence - The plants are now getting taller, spreading out. The result is a mixture of soil and plant signatures. The color is usually a pink or tan. As leaves spread out into a canopy over the soil, the soil signature will disappear.

4. Growing - Healthy, growing plants reflect photo-infrared. All crops are some shade of red when in their prime. The bright reds are most common. Variations in color can be due to row spacing, crop height (shadows cause a darker red), health of the plant, rainfall, etc.

5. Ripening - As crops ripen, growth begins to decrease. Some crops yellow as they ripen, some drop their leaves; the result is a decrease in the red color. In addition, the dead leaves add a bright tan color, and soil signatures may also become visible again. A mottled appearance develops as the red color decreases. The ripe crop signature will vary depending on the underlying soil and the amount of soil showing.

6. Dead Ripe - When the crop is standing dead ripe in the field, it no longer has a red signature. It has lost all its green leaves, and reflects no photo-infrared. The appearance is a mixture of soil and dead vegetation signatures.

7. Harvest - The color of the harvest and post-harvest signatures depend a great deal on the local cultural harvesting practices. A light tan to gray or white signature will result if stubble and stalks or leaves are left strewn over fields after harvest. Wheat is a good example of a crop with a white post - harvest signature (in the U.S.) because the straw is scattered over the soil and the soils are usually a lighter color soil. If a minimal amount of dead plant material is left in the fields. The soil signatures will be more dominant.

After harvest, fields may be replanted with another crop planted to hay or other ground cover, or simply left idle Post-harvest stubble may be plowed under, or the stubble may be left standing to be grazed by cattle.

IV. Some Specific Crops

A. Wheat - The following are "Bob Payne's Colors for Wheat".

Plowing - Brown (Soil)

Pre-Emergence- Light green to white

Growing - Deep red

Ripening - A yellowing green, growing greener

Dead Ripe - Dark green or olive

Harvested - white or light tan to gray

Winter wheat is easy to separate from summer crops by its harvest date. Partially harvested fields are easily identified during the harvest season.

It is not possible to separate wheat from barely with visual analysis of LANDSAT. Oats will also be very similar in appearance.

Because of their respective growing seasons, wheat can be rotated with soybeans in the same field during one year. This is practiced a great deal in the major soybean area of Brazil. In this area, the post harvest stubble of the wheat crop is not plowed under after the harvest. Soybeans are planted directly into the wheat stubble. Since some dead wheat straw is still on the fields during planting, the soil signature may not appear as dark as would be expected.

B. Soybeans - Soybeans are grown on bushy plants. The beans form in pods hanging from the plant. They are grown in larger, flatter fields like those of wheat. Soybeans are harvested by combine.

C. Corn - Corn signatures are also much like those of wheat, and of soybeans. Because corn plants are taller than soy plants, the shadows give Corn a darker red during the growing season.

Post-Harvest signatures for corn can vary. If the stubble is gathered for silage, the post-harvest signature should be a soil signature. If the stubble is left in the field for grazing, the signature will be lighter, getting progressively more like that of a soil signature as the stubble is eaten. Looking at harvest dates may help identify corn.

D. Rice -Rice requires fresh water for irrigation to flood the fields. Fields must be flat (or terraced) to hold the water. Good growth requires high temperatures in the growing season. Flooded fields are drained before harvest.

Rice fields should appear dark (water) in pre-emergence. Dark signatures from water will be subdued when a canopy of rice plants covers the water. As the flooded fields are drained before

60

harvest, the plants should take on a typical ripe crop signature. Typical harvesting and post-harvest signatures can also be expected.

F. Tobacco - Tobacco is limited to small field sizes. Of all tobacco fields, 95% are planted 4 rows, skip 1 row. In addition, tobacco can also be in alternating strips with other crops in the same field. The result will be mottled appearance. Flue cured and cigar wrapper tobaccos are harvested by cutting the leaves individually as they ripen. The red growing signature should become fainter as the leaves individually ripen and are picked. Other tobacco are harvested by cutting the stalk and removing all the leaves at once. Post-harvest will show a soil signature, since there is no dead vegetation covering the soil.

F. Cotton - The size of cotton plants depends upon the climate. In tropical conditions, cotton plants grow into large bushes. At the northern edge of the Cotton Belt, the plants are much smaller.

G. Bananas, and Other Tree Crops - Bananas, and other broadleaf evergreen tree crops, will appear much like that of coffee grown under shade. An extremely even texture and homogeneous color usually indicate some form of cultivation, as native vegetation would have a more mottled appearance. Such tree crop would be very difficult to separate from one another.

It may be possible to locate deciduous orchards and citrus crops by looking for rows of trees. If grass is planted between rows of trees, winter scenes should still show some pink signatures, with leaf less trees appearing as darker spots in rows in the pink fields.

V. On Being An Analyst

A. Think like an analyst. (Make others think you know what you're talking about).

When trying to identify land use, think about what those crops look like on the ground. How dense is the canopy over the ground? Is the soil visible between rows? Consider factors which may alter the typical signatures for the different stages of the crop. Gather as much ancillary data as possible. Such data should include:

- Cultural information on how the crops are planted and harvested.

- Country crop data telling how much of which crops are grown

61

in an area.

- Crop calendar showing the planting and harvesting dates of all crops in the area.

- General climatic data for the dates of scenes being studied. This will tell you which way to shift the crop calendar dates as a result of local weather.

- Topographic data showing general elevations and relief.

Suspected "soil" signatures often appear as small villages on topographical maps. The satellite imagery will help you identify changes in the topographical maps.

62

STRATA DEFINITIONS FOR PAKISTAN AREA SAMPLING FRAME

STRATUM 10 — Intensive Agriculture, 60 to 100% Cultivated.

This stratum includes land areas which are 60 to 100% cultivated at any time during the year. The minimum block size delineated will be 1 square kilometer. Delineation and shading will be in PURPLE.

STRATUM 20 — Extensive Agriculture, 30 to 59% Cultivated.

This stratum includes land areas which are 30 to 59% cultivated. Include land that is cultivated at any time during the year. The minimum size block delineated will be 2 square kilometers. Delineation and shading will be in SKY BLUE.

STRATUM 31. — 1 to 29% Cultivated, and/or Pasture.

This stratum includes land areas which are 1 to 29% cultivated. Include land that is cultivated at any time during the year. This stratum also includes pasture or grazing lands. The minimum block size is 6 square kilometers. Delineation and shading will be in ORANGE.

STRATUM 32. — River Bed

This stratum is the land area where the water course of the river has changed over time. This area of land is subject to flooding. The stratum soil patterns are clearly distinct from the srounding area when viewed on the satellite imagery. Often an embankment is present on both sides of the stratum. The minimum block size is 6 square kilimeters. Delineation is in ORANGE and shading is in DARK BLUE.

STRATUM 41 — Land Not Suitable for Cultivation or Pasture

This stratum includes areas which are not suitable for agriculture or pasture, such as deserts, high mountains, national parks, barren escarpments, glaciers, or any other totally barren areas. The minimum block size is 2 square kilometers. Delineation and shading will be in BROWN.

63

STRATUM 42 Land Reserved for National Security

This stratum includes areas which have been reserved for security purposes that include military installations, restricted border areas, or any other areas where sampling is prohibited. The minimum size block is 2 square kilometers. Delineation will be in BROWN, and shading will be BRIGHT GREEN.

STRATUM 51 Urban Areas

This stratum includes city centers, commercial centers, office complexes, and other urban areas where no agriculture or livestock is expected. The minimum block size will be 0.2 square kilometer. Delineation and shading will be in GREEN.

STRATUM 52 Villages and Other Agri-Urban areas

This stratum includes urban areas that may have agriculture or livestock present. Also included is the perimeter area of larger cities, villages, and concentrations of ten or more residences which may not be associated with a village. The minimum block size is 0.2 square kilometer. Delineation will be GREEN and shading will be in ORANGE.

STRATUM 61 Forests or Concentrations of Large Trees

This stratum includes forests, woods, or any other large concentrations of trees or timber. Orchards are excluded. The minimum block size will be 2 square kilometers. Delineation will be in GREEN, and shading will be in YELLOW.

STRATUM 62 Orchards or Other Concentrations of Tree Grown Crops

This stratum includes land occupied by orchards where tree grown fruit or nuts are produced. Timber plantations should be excluded. The minimum block size will be 1 square kilometer. Delineation will be in GREEN, and shading will be in BRIGHT GREEN.

STRATUM 70 Water

This stratum includes any body of water which is completely surrounded by land. Such as lakes, reservoirs. Large rivers will be included, but smaller streams will be excluded. The minimum block size will be 1 square kilometer. Delineation and shading will be in BLUE.

65