FEASIBILITY REPORT OF COMPUTER MAPPING SYSTEM

AGRICULTURAL REHABILITATION AND DEVELOPMENT ADMINISTRATION

DEPARTMENT OF AGRICULTURE, GOVERNMENT OF CANADA

PROJECT 14007

-

FEASIBILITY REPORT OF COMPUTER MAPPING SYSTEM

PREPARED BY

SPARTAN AIR SERVICES LTD., OTTAWA, CANADA

INDEX

CHAPTER 1

INTRODUCTION

The ability of an electronic computer to store, compile and assess statistical data has been increasingly demonstrated in recent years. A computer mapping system is essentially a combination of instruments and techniques that allows this capability to be applied to areal data derived from maps as well as to statistical data.

The nature of data relevant to the assessment and evaluation of marginal land has necessitated its portrayal in the form of maps. These are many and varied, but the data from them must be combined with the related statistical data if an understanding of the marginal agricultural areas is to be gained.

The Agricultural Rehabilitation and Development Administration has undertaken the major task of co-ordinating the evaluation of marginal lands in Canada. This includes the accumulation and handling of pertinent data, its compilation in such a way that all factors regarding a specific area may be considered together, and finally, the accurate weighing or comparison of it so that sound judgements can be made. On this basis the Agricultural Rehabilitation and Development Administration contracted with Spartan Air Services Ltd. to undertake a study of the feasibility of a Computer Mapping System and application to the ARDA program.

The study was carried out between March and August 1963. The present report covers the research carried out in that period.

As the data available to the system has a direct bearing on the input methods used and the type of computer facility required, the first part of the report includes a preliminary assessment of the location, format and amount of the data involved, and an analysis of the format with regard to input design. It must, however, be emphasized that the selection of specific assessments and evaluations of land resources can only be determined by ARDA, and that this in turn decides what data is required in the system. This investigation was carried out to indicate the data available to the system in order to allow the design of the input system to proceed.

There follows an outline of the various types of input system that would allow the entry of graphic data, and a report of the input processing, the maintenance, and assessment processing that would be needed. Finally, a discussion is made of the available output units and a summary is given of the system feasibility with recommendations for its implementation. The Appendices, bound separately, provide supporting data for references made in the text.

The study was carried out by the staff of Spartan Air Services Ltd. Ottawa, with services of International Business Machines Limited engaged to provide the necessary technical data regarding electronic computing methods and equipment.

The study would not have been possible without the active co-operation of the staff of ARDA, of related branches of the Department of Agriculture, and also of the Dominion Bureau of Statistics, The Department of Transport, the Department of Defence Production, the National Research Council, the Geographical Branch, Department of Mines and Technical Surveys, the Army Survey Establishment, U. S. Army Map Service and the U. S. Army Corps. of Engineers (GIMRADA). Co-operation was also received from the commercial companies noted in the text.

CHAPTER 2

DATA AVAILABLE TO THE COMPUTER MAPPING SYSTEM

Map Data

The map data available to the system will primarily be that generated specifically for the Canadian Land Inventory, with the addition of maps indicating the census subdivisions prepared by the Dominion Bureau of Statistics, and the present land use maps prepared by the Geographical Branch, Mines and Technical Surveys. Maps not included above may be considered for use in specific areas. The maps prepared for the Canadian Land Inventory will be available in their final form at a scale of 1:250,000. However, the production of some of these sheets will utilize preliminary maps compiled at a scale of 1:50,000 or 1:100,000, from which the data will be reduced to the final 1:250,000 scale. It may be considered desirable, therefore, to extract and record the data at the optimum scale at which it is available. In this way the most accurate information developed in the survey would be stored in the computer system for subsequent assessment. It can be noted that the computer system requires only a manuscript copy of a map for input purposes, and printed editions of the larger scale working sheets need not be made available.

Canadian Inventory Area

In Canada the total of occupied agricultural land and productive forested land, discounting forest woodland, is approximately 1,080,000 square miles. This figure corresponds well with the projected coverage of 135 sheets of 1:250,000 scale maps which would take in an area of approximately 1,117,000 square miles. The occupied agricultural land is approximately 270,000 square miles. While total area of forest land and agricultural land actually surveyed is not known, it is estimated that the forest reserves of immediate interest would not more than equal the area of the agricultural land, bringing the figure to a Canadian Land Inventory total in the order of 500 – 600,000 square miles. It is realized that the maps produced will necessarily have some areas of little interest to the inventory, and it is thought that the conservative figure of 500,000 square miles will represent the area of main concern in the Canadian Inventory.

Map Types

Four main types of land capability maps are to be produced at the final scale of 1:250,000 within the Canadian inventory area. These are the agricultural land capability map, the forest land capability map,

the wildlife land use map and the recreational land use map.

Within the agricultural area there are approximately 400 existing soil maps at varying scales and of various ages. To a great extent it is planned to use these as a source of information for the agricultural land capability maps. If it is wished to derive data for computer input from as large a scale map as practicable, the number of map sheets available to the system for agricultural land capability would fall between 135 and 400. It could then perhaps be estimated that approximately 300 map sheets, either manuscript or final, at varied scales between 1:50,000 and 1:250,000 would be available for input of agricultural land classification data.

A similar analysis would be completed for the forest land capability map. The essential lack of forest cover in those parts of the prairie provinces included in the Canadian Inventory area would probably limit the number of maps with data for forest land capability to 200.

No estimate is available at this time for the method of production of the wildlife and recreation maps, though their number will probably not be less than 135, and additional detailed sources may raise the total of available sheets to 150 in each case.

The census subdivision boundaries are printed on a wide variety of topographic map scales. Perhaps the best basic series to cover the inventory area is at a scale of 1:500,000. This, however, would need to be amplified by 1:250,000 and 1:50,000 maps in areas wherever census subdivision size decreased, particularly in eastern Canada and in British Columbia. After consultation with the Dominion Bureau of Statistics, Cartographic Division, it is estimated that a selection of 500 maps would give the optimum representation of accurate information across Canada.

The climatological maps being produced by the Ontario Research Foundation will be at a scale of 1:2,000,000. If data at this scale is required in storage, two maps would cover the inventory area. If it is desired to read in the separate overlays of the individual factors combined in the final map, a total of 14 maps can be envisaged.

Present land use maps are produced by the Geographical Branch, Department of Mines and Technical Surveys, by some of the Provincial Governments of parts of the respective Provinces (usually conservation areas) and by Universities. The last named are mainly large scale investigations of local, or occasionally, regional interest.

Thirty-five maps at scales varying from 1:500,000 to 1:50,000 have been completed by the

Geographical Branch. Of these, 22 are at a scale of 1:50,000 and a total of 17 compilation sheets at 1:50,000 underly the printed maps at smaller scales. It is estimated that a further 167 x 2 working sheets at a scale of 1:50,000 are under present or projected development to portray the present land use in New Brunswick and parts of Quebec, Ontario and south-western B.C. The manuscript stage of compilation of all of these should be completed within the time projected for the Canadian Land Inventory. The Provincial Governments of Saskatchewan and Manitoba are planning production of present land use maps from existing assessment records to cover the agricultural areas in the provinces. It is estimated that the total number of land use maps available to the computer system from all sources will be between 400 and 500 sheets.

In summary, there will probably be between 1600 and 1700 map sheets available for assessment within the Canadian Inventory area. No account is taken of future map production, more detailed surveys, other sources of the same type of map, or other types of maps (e.g., water shed, conservation areas, etc.). On the basis of a three year programme for the Canadian Inventory, and on suitability for input and revision thereafter it is recommended that the input device have a capacity in the order of 1000 maps per year as a minimum.

Statistical Data

Most statistical data can readily be put into a data processing system. Frequently statistical data is already in the form of punch cards or may even be stored on magnetic tape. Manuscript or "raw" data needs to be put onto punch cards, using unit record equipment similar to that already being used by the Department of Agriculture, Data Processing Division. There are no technical feasibility problems in this area. The use of such information will depend on the cost of using existing equipment to translate the required raw data to computer format, or to extract pertinent data from existing cards or tapes.

The largest source of immediately available agricultural data is the Dominion Bureau of Statistics. 155 separate factors regarding each farm, taken from each agricultural questionnaire (Form 6) is recorded on magnetic tape, and tabulated by census subdivision (township). Given the map of census subdivisions in computer storage, this data can then be compared by area to map data. This data can be regarded as being readily available at this level of reduction.

The maximum use from census statistical data would be gained if the data for each farm could be applied to the locale of the particular farm rather than merged in the average of a census subdivision. Enumerator area maps exist as raw data to which the individual census returns could be related. This in itself would be a significant breakdown from the census subdivision level. Farm numbers and locations can also be manually taken

from the enumerator maps in many areas of Canada and their co-ordinate position determined. This would allow any farm data to be recorded at a specific point, which could then be noted on any area of assessment regardless of size. In many areas of western Canada the range and township for each farm is known and recorded on magnetic tape, and farm data can be related to a much more precise area than a census subdivision. The handling of such detailed data is well within the capacity of a computer. If the computer mapping facility is developed, it is recommended that the reduction of census data to smaller units than census subdivisions be carefully considered, and steps taken to use this data in conjunction with the similarly detailed land capability studies.

The types of data which may be extracted from the census reports either at census subdivision level, or with manual data reduction at a more detailed level, include: – size of farm, net farm income, acreage of improved pasture, acreage of unimproved pasture, irrigated acreage, number of livestock (separated into cattle, horses, hogs and poultry), and type of land tenure.

The combination of statistics whether prior to entry into the computer mapping system or within the system itself could produce data which may be pertinent to ARDA requirements. The relationship between gross incomes of the classified groups (dairy, wheat, etc.) could be produced in a classification of type of farming by census subdivision, or, again with manual reduction of existing data by smaller subdivisions. The area of land used for various purposes could be combined to form a present land use rating which might substantiate the present land use maps. An economic classification of farm types could result from a combination of total sales of farm products, amounts of off-farm work and age of operator. A capital investment rating may be produced by combining data regarding real estate, livestock and farm equipment. Population data, subdivided into total population, rural population and farm population is available, already tabulated by enumerator area.

It can be emphasized again that the above comments regarding agricultural data are intended only to illustrate the types of data available, their present accessibility, and the steps necessary if a more detailed breakdown of the existing data is required. The selection of statistical data input will depend entirely on the output required for ARDA purposes.

Statistics regarding forestry are available from both the Dominion Bureau of Statistics and from Provincial Governments. The Dominion Bureau of Statistics has data covering forest products from on-farm woodlots, not in computer format but tabulated by township. Primary data, regarding sawmills, pulp and paper mills and veneer and plywood mills is available on punch cards with more detailed data available in raw format. Tabulation of this group is by municipality. A sample census of logging of 7000 operators is taken and has been extrapolated

manually into provincial estimates, broken down by municipality. The Dominion Bureau of Statistics also carries out sampling of privately owned woodlands. The results from this sampling is available on punch cards for New Brunswick, Nova Scotia, Prince Edward Island, approximately half of Ontario and selected areas in Quebec. The sampling is broken down into forest districts within each province.

The Provincial Governments have the forest data regarding crown lands within their provinces. The breakdown of this data is not known, though the forest production by product and by species will be recorded on punch cards in 1963 in the provinces of Ontario and Manitoba.

As climatological maps are being produced for the Canadian Inventory and may be available to the computer system, statistical data for climatological assessment will not be given in detail. The maps will, however, be at a relatively small scale (1:2,000,000 approximately) and it may be of value to note the data source. The data will come from the first and second class weather stations, comprising 650 stations using data from 1921 - 50, and an additional 80 stations with records from 1951 - 60 which will be adjusted to the 1921 - 50 period. This data is available on punch cards. It could be put into computer storage with regard to the co-ordinate position of the station if desired. In the same way any climatological data from any other station could be inserted into computer storage with regard to position in Canada.

Statistical data from many other sources, such as Farm Loan Agencies, Crop Insurance Agencies, The Tobacco Growers Association, The Sugar Beet Company of Winnipeg, CPR, CNR and other transportation agencies have not been investigated, but are available, and can be put into computer storage and compared with other data whenever pertinent to ARDA requirements. A considerable amount of this data is known to be on punch cards.

The excellent facility for handling raw data and transcribing it onto punch cards already existing in the Department of Agriculture Data Processing Centre would, of course, be a valuable asset in the input of statistical data to the computer mapping system.

CHAPTER 3

ANALYSIS OF AVAILABLE DATA FOR INPUT SYSTEM DESIGN

Statistical Data

As indicated above, no technical feasibility problems are presented by the statistical data available to the program. Much of it is already in machine readable form. The reduction of raw data to machine-readable form can be handled with existing instruments and methods if necessary.

Map Data

There is not, at this time, any existing instrument suitable for directly reading map data into computer storage. The design of an input system capable of fulfilling this requirement is naturally dependant on the nature and format of the material to be handled. The maps available to the system have been indicated above. The following notes record the analysis of the format of these maps as it effects input system design.

Physical Format

1. The base maps on which the working data will be plotted will primarily be of paper. Subsequent over-lays produced from this paper will not remove the inherent distortions due to shrinkage, though they would provide a constant base for later reference.

2. 99% of the maps will be smaller than 48" x 48". No map will be larger than 48" x 96".

3. Maps may be folded, though every effort will be made to produce the original work on flat sheets, and to transport them rolled rather than folded.

4. Annotations will be either in the form of a printed line (in the case where final maps are used as source material) or, on manuscript maps, will be in the form of pencil or ink line over a printed line, or pencil or ink line over coloured paper, or pencil or ink line on blank paper. In cases where a pencil or ink line is annotated over a printed line, and deviates no more than 0.1" from the printed line, then the printed line is to be used as the appropriate boundary, unless otherwise indicated.

Data Format

"Map data" consist of lines (referred to hereafter as boundaries) demarking areas, marks which relate the sheet to the earth's surface, and identifiers which relate each area to its set of properties. Other con-

Note: Definition of Terms - Base Maps - Either topographic or other map on which data pertinent to the ARDA program is to be overdrawn. Manuscript Map - Base Map with data overdrawn. Overlay - tracing on stable material of data for computer input. Final Map - printed map - often at reduced scale from manuscript map - possibly with change in classification.

nected data consists of all information not presented graphically which must be identified with some area on the map sheet. The following parameters apply to the map data occurring on the sheets available to the ARDA programme:

1. Each map has marks, either in the form of co-ordinate points or co-ordinate lines, which allow its position on the earth's surface to be fixed.

2. Each map is at a known scale which is constant over the sheet.

3. There are six principle sets of maps.

4. Approximately 65% of the maps have a "normal density data format". This has the following characteristics:

    (a) The average parcel of land on the map sheets is approximately 2 square miles.

    (b) The average perimeter of this area is approximately 6.5 miles.

    (c) No area of less than 1 square mile is drawn at any scale. This is an area that can be drawn with approximately 1/4" (6 mm) of cross section at a scale of 1:250,000.

5. "Low density data format" maps have a data density of approximately one tenth of normal density data format maps.

6. "High density data format" maps will comprise less than 3% of map total, and will have a data concentration of no more than a factor of x5 compared with normal density data format maps.

7. The number of divisions of a particular classification on any sheet is less than 64.

8. In any one square mile less than 8 distinct areas are present.

9. As a base for system accuracy it is assumed that all data shown on the map sheet is positionally accurate.

CHAPTER 4

INPUT OF DATA

The key to the application of computer capability to the Canadian Land Inventory is the ability to transcribe the graphic data presented on maps to a digital form suitable for machine processing. It has been noted that the statistical data already in digital form can be processed with existing methods. If the map data can be similarly represented then both map data and statistical data can be combined and analyzed with existing computers and the resulting data can be reproduced numerically or graphically with existing instrumentation.

Based on the foregoing assessment of available information and the analysis of the map format, the following summary of input system criteria is possible:-

1.   The instrument must be capable of accepting 48" x 48" paper maps.

2.   The physical operation of input should be as simple as possible with minimal pre-machine preparation of the maps to be transcribed.

3.   The input system should provide data in a form suitable for machine processing with a minimum of programming problems.

4.   The system should be capable of processing 1000 maps per year, which in practical terms is an input speed of 1 map per working hour.

5.   The input system should maintain the accuracy of the data presented on the given map.

An input system will be composed of three linked units. The first of these, referred to as the 'transcription device' accepts the maps and converts the required data to digital form. The second part of the system is a 'buffer unit' which, accepts data at high speed, sorts the significant digits and sends it to the storage unit. The final unit is the 'storage unit', which can be a card punch, but in most cases is a magnetic tape drive. Both the buffer unit and the storage unit are readily available, so this report concerns itself mainly with the transcription device.

Prior to this study, one type of transcription device had been suggested. Limitations have become apparent in the original concept, however, so further investigation has produced both developments of the original and a new approach. The result is four types of transcription device which may be characterized as follows:-

1.   Tracing Type

2.   Optical Tracing Type

3. Line Following Type

4. Scanning Type

Tracing Type

Experience with high precision photogrammetric plotters led to the suggestion that a transcription device could be produced which would allow any map sheet to be placed on a plotting table and the data traced directly into the buffer and storage unit. This would be done by aligning the map co-ordinates with the edges of the tracing table so that each movement of the plotter pointer would be related to geographical co-ordinates. As the pointer was moved around a type boundary, the changes in its co-ordinates would be automatically recorded in the computer. On the photogrammetric plotters, a movement of .01 mm represents a one unit change in a co-ordinate so it can be seen that an area could be very accurately described by this method. As the area was plotted, so the character of the area could be put into the computer's memory by a key punch method.

The problem of minor distortion in any map sheet could be resolved by placing the pointer on the corner co-ordinates at the beginning of the plot-in. This would allow the computer to record the known co-ordinate for a physical position of the pointer. Subsequent movements of the pointer would then be related to the known co-ordinate position rather than the edges of the plotting table and thus the information would be described by its true co-ordinate positions. The plotting by co-ordinates also resolves the problem of scale, as the storage, comparison and retrieval can be done entirely from the co-ordinate specifications.

The disadvantage of this type of system is the inherent human factor of the tracing process. This includes the possibility of missing a boundary completely, the tracing instability of the operator, and the length of time that the whole input system is in use for the tracing of one map.

Two types of tracing device may be recognized. The first of these is a non-automatic device where the pointer is placed on the map and a button is depressed whenever the co-ordinate position is required to be read into storage. It can be seen that this process would be extremely laborious for the tracing input described above, yet it could have an application to the input of spot information, particularly if incorporated with a boundary scanning type of input device. The second type is the fully automatic system which can operate both in a button pushing mode and in the automatic mode described above, where any movement of the pointer is instantly digitized.
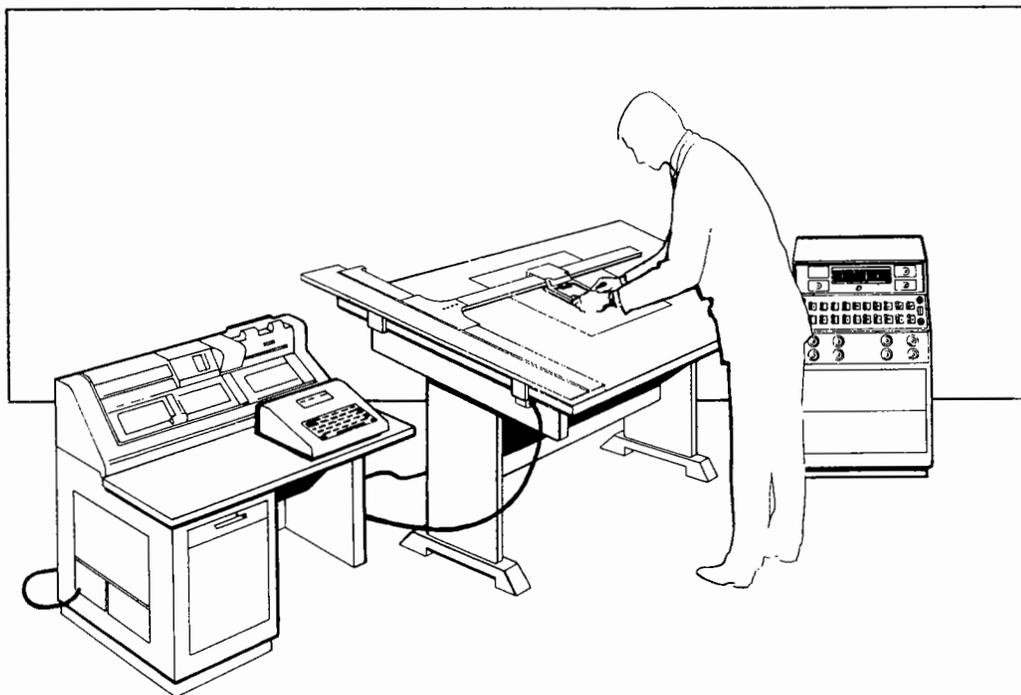
Typical Non-Automatic Tracing Type Transcription Device
coupled to an I.B.M. 526 Card Punch
(Courtesy Keuffel & Esser Co.)

Co-operation from the Department of Defence Production resulted in a preliminary list of suppliers who had experience in either the production of photogrammetric plotters or had completed research in the fields of electronics related to the input systems. Preliminary specifications were drawn up for each type of input device. Specification A for the automatic unit is shown in Appendix 2. Specification B for the non-automatic unit is shown in Appendix 3. Of the twelve companies contacted, six (Gerber Scientific Instrument Company, Instronics Limited, Keuffel & Esser Company, Olkon Company of Canada, James W. Stevenson Company and Wild of Canada Limited) indicated that they had equipment which with slight modification would fulfil Specification B. This simple type of instrumentation ahs other applications which have allowed its production before this time. Operating models can be seen at a variety of locations. The modifications required are primarily to accept a larger format of input material that had previously been used. While the suppliers were only asked to indicate their interest in this field, to outline the nature of their existing instrumentation and to give an area of cost of this type of device, the six replies indicated that this equipment would be readily available and that the price range was in the order of $30,000 plus the cost of an IBM 526 punch card unit. The lowest cost received was in the order of $20,000 plus an IBM 526; the highest, $70,000. The last figure was essentially for a full
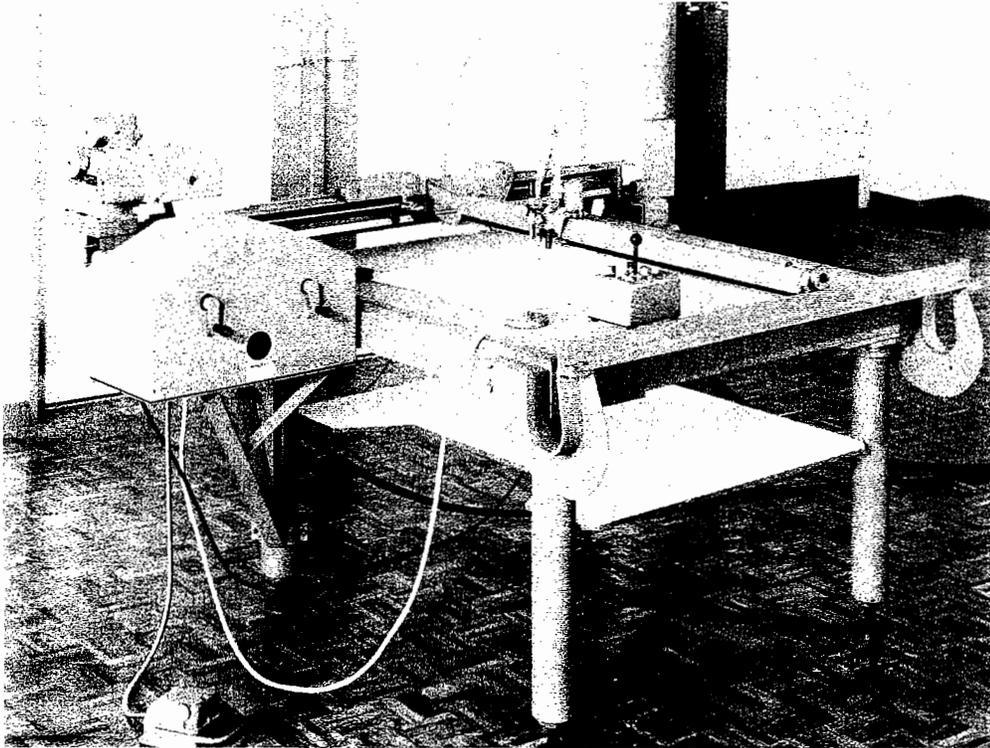
NON-AUTOMATIC TRACING TYPE TRANSCRIPTION DEVICE
(Courtesy Wild Company of Canada).

scale output plotting unit which could be used in reverse. It must be emphasized that the costs given are not tender prices and only indicate an area of price rather than a competitive bid to meet a detailed set of final specifications. Four companies, (Concord Control Inc., Gerber Scientific Instrument Company, Instronics Limited and Keuffel & Esser Company) submitted proposals for a full automatic device in answer to Specification A. Two of the companies (Concord Control Inc., and Instronics Limited) have instrumentation capable of automatically digitizing graphic information which with modification would be suitable for use as a tracing type input unit. The other two companies (Gerber Scientific Instrument Company and Keuffel & Esser Company) who are leaders in the field of this type of instrumentation, have completed research on this type of device and have the facilities to produce it. The area of cost is between $110,000 and $175,000.

A consideration of this equipment must include the inherent human disadvantage of tracing instability as a source of error. In the non-automatic equipment the pointer can be precisely positioned over any point before read in is activated. With the automatic equipment human tracing stability is a major factor for

consideration. The only automatic read in tracing unit on which test results were available was the device pro-duced by the Concord Control Inc. These tests, done under the supervision of the U.S. Army Map Service, initially indicated the tracing reliability at 700 measured points on one map to $\pm$ 1/200". Further investigation revealed that the source of data was in fact an etched zinc plate and that the pointer used was a steel needle which lightly rested in the groove formed by the etched lines. Thus a measure of lateral stability was imparted to the tracing operation which would not be found when using regular map input. While this test showed that the system is capable of picking up, storing and producing data to the required degree of accuracy, it could not be related to human tracing stability over a paper map. As the factor was considered to be critical to all tracing of input systems, a preliminary investigation of human tracing error was conducted in Ottawa. In absence of a suitable input unit, the tracing was done with normal drafting instruments. The tracing sample was taken from a typical land resources map with normal density data format and was produced on stable material. A variety of draftsmen and operators was selected, each tracing the master on to translucent stable material, firstly, at normal speed, secondly, at high speed and finally, at slow speed with maximum accuracy. After all tracing had been completed a small area of the master was selected and enlarged on a high precision enlarger. The same area was taken from each of the samples and similarly enlarged. Deviations from the original in excess of line width were measured. As the ultimate output of the computer mapping system would probably be in terms of area or comparison of area, the area of deviation produced by tracing error was measured. The resulting data was plotted stotistically and found to have excessive scatter for any absolute estimate to be established. Based on the data recorded, however, it was found that the error was closely related to boundary length and can be expressed as .6 x 10^{-3} square inches of error per linear inch of boundary measured. To give a more common-place expression of this error it can be calculated that on 1" to 1 mile scale map the error incurred on the overage would be $\pm$ 3 acres in 1300 acres.

While it must be stressed that further tests of operator error are necessary and should preferably be made with a pointer having the same inertia as that to be encountered on the tracing type input device, it appears that tracing error is not great enough to eliminate consideration of the automatic tracing type input unit.

The second inherent disadvantage to the tracing type input unit is the time of tracing the complete map sheet. Tests were run in Ottawa with selected operators tracing a typical map sheet with normal density data format onto translucent material with normal drafting instruments. The working time in hours involved was between seven and ten hours per sheet. The operators selected were used to doing this type of work in a commercial company and it is not thought that much greater speeds could be expected by an operator even with continuous operation of

a tracing type input unit. It can, however, be estimated that low density data format maps would be input at approximately 1/10th of the above time, or, for practical purposes, within one hour. With the tracing type input system, however, it would be essential for the tracing to be verified. Probably the most effective way of doing this is to have the same map traced twice, and use the computer facility to compare the two inputs, perhaps optimize the lines to offset the tracing instability still further, and most important, to indicate an error condition where any boundary does not close. This, however, implies that the same map is traced twice. This would give a minimum input time for a normal density data format map of approximately two days. The input time for low density format maps would be approximately two hours. When it is considered that the system requires an input of one thousand maps in one year, or approximately five maps per day, the area of cost of one automatic tracing type input device makes this disadvantage a limiting factor in the adoption of input devices utilizing the tracing method.

Optical Type

A development of the automatic tracing type transcription device is known as the "optical" type. The concept of this device is to optically enlarge the section of the map being transcribed so that the pointer movement would at all times be within the optically enlarged boundary being traced. The basic design of such a unit would incorporate a base plate on which the original map would be placed, and, attached to it, a pointer coupled to the computer system. This could be pointed to the co-ordinates of the map on the base plate and so numerically record its position. Suspended over the base plate on the rack would be an optical scanning device. This would pick up the image of a small section of the map and transmit an enlarged image onto a small 'tracing' screen. The operator could then use a second pointer electrically coupled to co-ordinates on the edges of the tracing screen to transcribe selected data from the enlarged portion of the map. Under the control of the operator the scanning head would next be moved to the adjacent section of the map. The movement of the scanning head itself would represent an increment in terms of co-ordinates which could automatically index the second portion of the map to be enlarged. The operator could then continue his tracing of the appropriate boundaries with reference to the co-ordinates on either side of the scanning screen.

There are three immediate advantages of this device. Firstly, the human tracing error is negligible, it could be assumed that the transcribed boundary lines would be well within the limits of error acceptable to the system. Secondly, the human operator could transcribe only those lines necessary and ignore the background data on the map. Finally, the optical scanning unit would increment its own unit areas, and this could be utilized as a numerical indexing base. One disadvantage of this particular system is that the identification of

character of specific types of land separated by the boundaries has to be done on each unit area incremented. This would duplicate many times over the required input of coding data. The second disadvantage is that while the indexing of the incremented areas would be valid for one map, it would be out-of-line, i.e. non coincident, with an adjacent map. Also, as the device described would not rely on any existing instrumentation but would have to be built strictly as a research contract at an estimated cost of $200,000, and as the time of input would be essentially the same as that of the automatic tracing type of transcription device, further investigation along this line is not considered worthwhile at this time.

Line Following Device

The concept of a tracing device that would automatically follow the required boundary lines was put forth early in the investigation. Preliminary research indicated that line tracing sensing heads that could follow lines on paper are available and had an accuracy of .01 of an inch. Two applications of such a device were suggested. The first was that the line following head guided by an operator could pick up the appropriate boundary line with considerable accuracy and so eliminate operator tracing instability. The second suggestion was that a tracing could be made on the original map sheet, taking off only such data as was required, and the line tracing head would then follow this pattern automatically and transcribe it into the computer storage. Further investigation of this device, particularly one produced by the Canadian Westinghouse Company Limited and named 'Linatrol', indicated that when such a tracing head was guided by the operator over a normal map which had background data as well as the lines to be transcribed, the device picked up background data which could not be eliminated without considerable pre-tracing map preparation.

The simplest form of pre-tracing map preparation was the tracing of the appropriate boundary lines, and this led to the second suggested application. It was considered, however, that if such pre-tracing map preparation was envisaged and considered acceptable in an input system, then a scanning unit could be devised which would pick up the data in much the same way as a line following device but which would not need human operator guidance. On this basis the incorporation of the line following device in a tracing type input unit was not followed further at this time.

Scanning Type

The concept of automatically scanning the original map sheet and abstracting those lines pertinent to the survey was also considered early in the investigation. The conclusion was reached that a scanning

type device using a TV type scanner could not resolve data accurately enough, particularly from the 48" x 48" original format. Secondly, it was thought that a device capable of extracting only the boundary data and ignoring the background data would be exceedingly complex and costly.

In the last month of the investigation, however, the inherent disadvantages of the tracing type of device were better documented, and to achieve an input level to accommodate a thousand maps per annum and avoid the committment of a number of tracing type units, it was decided to examine all possible approaches to data input that obviated manual tracing with a transcription device. The main problem of the scanning approach was the inability of a scanning head to differentiate between required and background data. To bring the scanning system within the area of feasability it was decided to divide the input task into two parts. If the boundary data could be extracted from the map and scanned independently, the areas and their characters could be swiftly identified by a non-automatic tracing type of transcription device. The separate boundary data and area identifications could then be merged in computer storage. This concept had the following advantages:

1. The existing, and relatively inexpensive, non-automatic tracing type of input device could be purchased for the input of character data.

2. The tracing of boundary data could be done from either printed or manuscript maps onto a stable overlay by non specialist personnel using only light-tables and pens and utilising neither expensive equipment nor computer time.

3. The scanning problem could be reduced to the task of scanning black lines on white material with no differentiation of data required.

The second major difficulty in adapting a scanning unit to graphic data input was the inherent inaccuracy of the TV type of scanning head. It was thought to be impossible, with existing equipment, to resolve 0.01 inch over a 48" x 48" original. Developments in the automatic mapping field, however, saw the production by the Federal System Division of IBM under research contract to the U.S. Army (GIMRADA) of a scanning device capable of resolving 0.002 inch over a 9 inch x 9 inch format. Permission was gained from GIMRADA to examine this device and enquire into the possibility of the technique being applied to scanning a larger format.

These discussions took place in the last days of July, immediately prior to the production of this report.
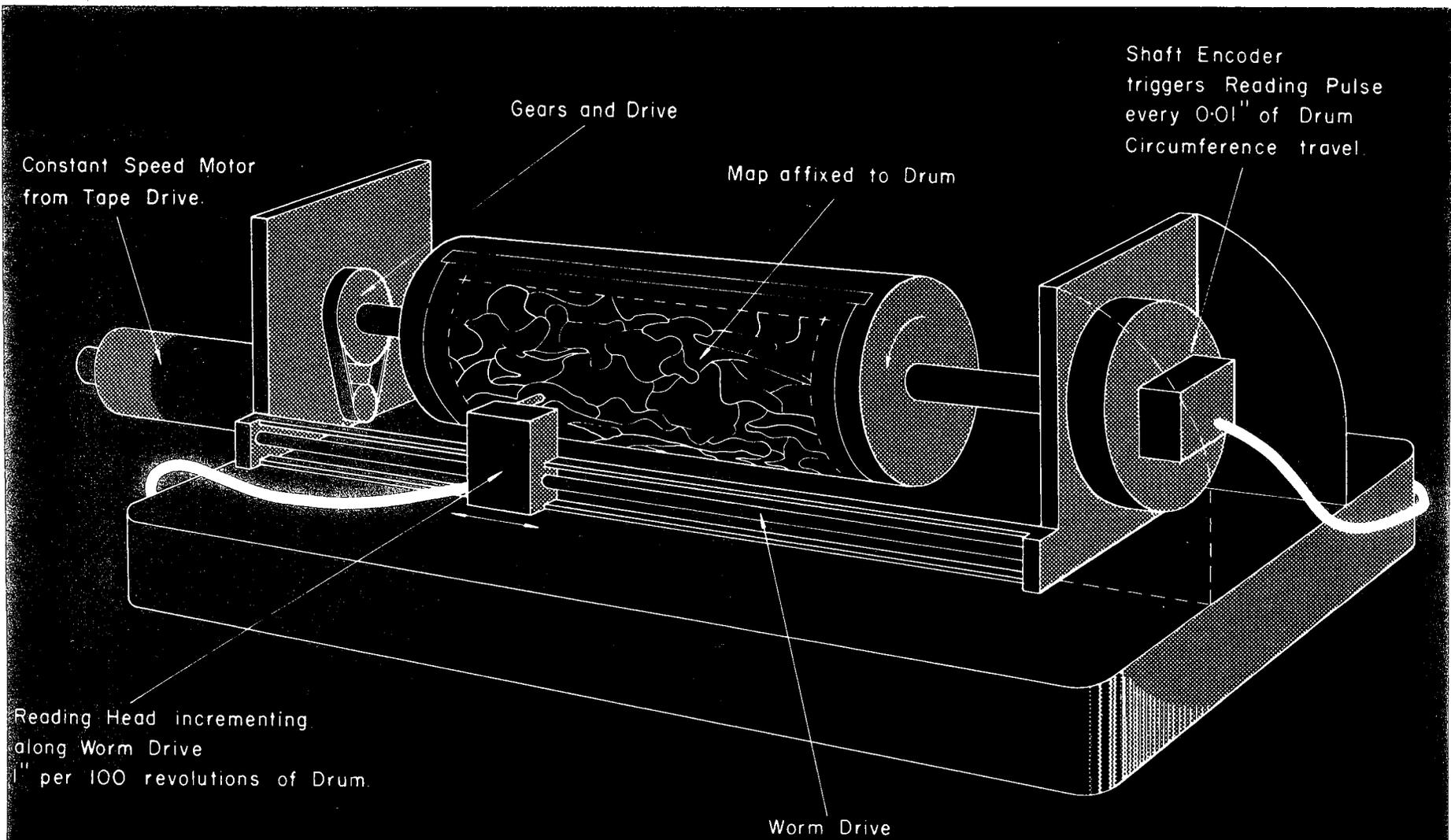
Basically, the device is a TV type of scanning head which scans a narrow band approximately 0.002 inches thick and 4.5 inches long, and is moved horizontally over the scanned area by means of a precision

frame. The frame presently in use is from a first order photogrammetric plotter, and its accuracy of movement controls the precision of the scanning position. In a 48 inch x 48 inch device, the frame would have to be correspondingly larger, and it is doubtful if any existing equipment could be utilized, as it has been in the case of the prototype. It may be possible to scan a larger area by positioning the scanning head further away from the object, though this would need to be researched. It would be possible to reduce a map sheet photographically to 9 inch x 9 inch and use this instrumentation without alteration. This has been offered as a test procedure and would be of value in developing the programming task, but it is considered that the incorporation of pre-scanning map preparation task of photographic reduction should be avoided if possible. The engineers in the discussion gave a verbal estimate of $250,000 to develop instrumentation for a 48 inch x 48 inch scan. This would have to be confirmed by a study if this approach were to be followed.

The cost, the research involved and the need to procure a stable frame with precision movement over large areas are main considerations in this system. Although the instrument is in present operation, transferring data to magnetic tape in a form that allows manipulation and processing, it is considered that a scanning device more aligned to the specific requirements should be the aim of the investigation.

A detailed discussion of our requirements with engineers of IBM Federal Systems Division led to the examination of a Drum Type Scanner which had been built in July, and was about to be tested. This particular device was constructed to test the possibility of scanning larger formats. The material to be scanned is wrapped around the drum, which is then revolved past the reading head. The accuracy of the system is not as great as the TV type of scanning device developed by IBM, but resolution was to .001 inches of the original format. The singular advantage of this system is the fact that it consists entirely of existing instrumentation, and the whole device is within the present state of the art.

Basically, the drum is a hollow stainless steel cylinder. To accommodate a 48 inch x 48 inch map sheet, a drum approximately 16 inches in diameter and 50 inches in length would be desirable. In the prototype a smaller drum was employed — approximately 25 inches long and 10 inches in diameter. The drum was mounted on a standard lathe bed with a heavy, rigid mount. The drive was by gear and belt from a constant speed motor of the type used on a magnetic tape device. The rotational movement of the shaft was measured by a shaft encoder, which could register 1000 units for every inch of rotation at the circumference. In front of the drum, mounted on the screw bar which formerly held the lathe cutting tool, was the reading head. As the drum rotated, so the reading head was incremented along the bar. This occurred at a rate of one inch per 100 revolutions. The reading head

Shaft Encoder
triggers Reading Pulse
every 0·01" of Drum
Circumference travel.

Gears and Drive

Map affixed to Drum

Constant Speed Motor
from Tape Drive.

Reading Head incrementing
along Worm Drive
1" per 100 revolutions of Drum.

Worm Drive

SKETCH SHOWING MAIN COMPONENTS OF DRUM-TYPE SCANNING INPUT UNIT

was impulsed 1000 times per inch of revolution of the circumference. This, in effect gave a network of points, with 1,000,000 per square inch. Such resolution would not be required for ARDA purposes; 100 points per inch (10,000 per square inch) would be adequate. The reading head operated on a reflected light principle: the intensity of a spot of light projected from the reading head onto the drum was measured. On the prototype, nine shades of graduation from black to white were being accurately recorded. The present requirement is simply black or white, to provide ease in conversion from analogue to digital form. The prototype could also operate in the reverse mode, putting out a beam of varying intensity which could "draw" a photograph on light sensitive material. This is, in fact, the facsimile principle. It would not be necessary on instrumentation to fulfill the present requirements.

The speed of rotation was governed by the speed of pick-up of the shaft encoder. The speeds possible with a .01 resolution would be higher than those on the prototype, and read in time could be further decreased by using additional read heads. It is estimated that using six read heads it would be possible to read in a 48 inch x 48 inch format in 16 minutes.

It is emphasized that the device is in a preliminary stage of development, and that tests to determine its capabilities and output problems have just started and are continuing. The design of this instrument does, however, answer the requirements for a transcription device, and it is based on existing instrumentation. Discussions with the engineers responsible for developing this equipment estimated that the cost of developing similar instrumentation to meet the ARDA requirements would be in the area of $85-$125,000. This would need to be coupled with a magnetic tape drive, the cost of the total system being estimated at $135-$170,000.

While this system would appear to be closely aligned with ARDA requirements, it would be impossible to write detailed specifications for its costing or manufacture without considerable study and development. It is also considered that, as in all the instruments outlined, the parameter of size radically alters the problems of instrumentation. Thus the production of this instrument, while relying entirely on existing equipment, could perhaps be expedited as a research and development project by a manufacturer with a background of experience in the field, and the capability to produce the required instrument.

In order to be as fully cognizant with the possibilities of this approach, and also of the problems that may arise, it is recommended that close contact be maintained with all tests and studies that are carried out in the next months that may be pertinent to this enquiry.

CHAPTER 5

INPUT PROCESSING

Data File Organization

The data file organization is a critical factor in the system. The file may be considered to consist of three portions, an index file, a characteristics file and a boundary file. The index file would provide the means of access to data in the other files, and the characteristics file would contain all descriptive material related to specific areas. The boundary file would be the largest and the most critically effected by various data representations.

The size of the data file naturally depends on the amount of data, both graphic and statistic, that is read into the system. It must be stressed that the following estimates are most tentative and are based on assumptions that could be in error by several degrees of magnitude.

A review of the data available to the computer mapping system and an analysis of that data allow the assumption that the Canada-wide coverage of any one type of map has an average of 254 sheets, and that the average number of square inches on each sheet that has to be transcribed is 900. This gives a total of 228,600 square inches to be transcribed in any one set of maps. If the data is to be read in with a frequency of 100 points per inch, i.e. 100^2 per square inch, and if local coding as estimated requires 24 bits (4 characters), then based on a magnetic tape capacity of 20,000,000 characters per reel, the tape requirement for five sets of maps would be 2,286 reels, i.e. $\dfrac{(5 \times 228,600 \times 4 \times (100)^2)}{(20 \times 10^6)}$.

To reduce this large number of reels, coding of the location tags and identifiers was considered. If this is done on a sheet basis, local coding can be reduced to 3 bits, which equals 1/2 character. On the same basis as above, the tape requirement for five coverages would be 290 reels, i.e. $\dfrac{(5 \times 228,600 \times \frac{1}{2} \times (100)^2)}{(20 \times 10^6)}$.

A further reduction in the number of reels may be achieved if only the data at boundaries is coded. From an analysis of the available data, it can be deduced that the average area of 2 square inches has a boundary of 6.5 linear inches associated with it. As this is the boundary of two parcels of land, the boundary effectively connected with area equals 3.2 inches for 2 square inches of map. There are thus 1.6 inches of boundary associated with each square inch of map sheet. It is estimated that if less than 8 distinct land types occur within any one square inch the boundaries within that square inch can be referenced by a directory of 11 characters.

On a basis that one boundary crossing is recorded for every 1/100th of an inch of boundary, and that each crossing would require 2 characters, the tape requirement would be $(1.6 \times 2 \times 100 + 11)$ per square inch of map. The tape requirements for five coverages using this method would then be in the order of 21 reels.

The above figures are based on the analysis of a normal density data format map. As 2 coverages will be at a lower density data format, which will require approximately 1/10th of the above capacity, it can be suggested that the above number of reels is probably high for the coverages involved.

Because a very compact file requires special notation, it would require increased processing time over other, less compact file representation. However, for a large file, magnetic tape handling would contribute greatly to the total time for the job: thus, between the two extremes of file organization, an optimum exists for a particular system. Further study would be needed to substantiate these figures and give an accurate assessment of tape requirements. This will only be possible when actual input data is processed. It may, however, be suggested that the most economical tape file in regard to ease of retrieval of data is between 20 and 200 reels of tape for five map coverages of Canada.

Co-ordinate System

A co-ordinate system is necessary to reference all common data in a system. Several co-ordinate systems exist which may be applied to a national inventory area with greater or lesser degrees of accuracy. Among the Terrestrial systems directly related to the surface of the earth are geodetic co-ordinates, astronomic co-ordinates and the township and range system (the latter not being a co-ordinate system but rather a partitioning of areas with useful metrical properties). The Spherical systems are those which consist of transforming geodetic latitude and longitude to a sphere in a manner which may preserve clearly defined ellipsoidal properties, the advantage being that computations involving these can be treated as a sphere. The Map Projection systems consist of transforming either geodetic or spherical co-ordinates to a plane in a manner which may preserve certain clearly defined ellipsoidal or spherical properties. The advantage is that computations involving these properties may be treated as though the earth were a plain disk.

The question of which co-ordinate system to use as a base for the computer system is to some extent independent of the form of the input data, in that mathematical correlations can be simply and easily made between the systems within the computer storage. The main factor concerning this particular study, however, is the ability to compute areas from the base co-ordinate system after the data has been recorded. This can be done from any co-ordinate system, though the task is much simpler if rectangular co-ordinates can be used. With

this in mind, an investigation was made of a co-ordinate reference system related to the geographic system but in all areas having a local rectangular co-ordinate base. A co-ordinate frame of this type combined with data grouped into blocks provides a scheme which can rapidly access groups of points in the file without excessive unproductive file motion.

Obviously a rectangular frame co-ordinate system will have an error when related to spherical co-ordinates. An overlap would occur at the top corner of each block. However, given ten mile blocks aligned to a meridian at latitude 45° a ten mile edge of the block would in fact be represented by true distance by 10 (1- .0025) miles on the surface. Thus the block co-ordinate system gives only a two count error based on one hundred points per mile. Even if no correction is made during input, the position error at the corner of the blocks is limited to a maximum of 100 feet on the surface. It is felt that the percentage area error in measurements taken from such a co-ordinate system would be well within the total area measurement error that could be tolerated within the system and the adoption of such a rectangular co-ordinate system would measurably reduce the task of machine computation of areas.

Machine Capability

The general machine capability which would be required to process data is partly dependant upon the magnitude of the problem and the period of time specified for the processing of data. It would appear, from the assessment of available data and the expected digital output of the input units, that the capability of an IBM 7000 series system is desirable for input processing and establishment of the data file. In the early stages of the programme development, the IBM 1401 would be an adequate instrument to prove trial programs and to process test areas. While it is impossible to estimate the processing time required, as the amount of input data is unknown, the form of digitizing from the input unit is unknown and all factors of file coding and co-ordinate system have received only preliminary study, it can be suggested that the area of processing time that might be envisaged would be in the order of one hundred hours of IBM 7000 series time for the total number of available map sheet coverages, i.e. sixteen hundred maps. Further research and definite parameters are necessary to substantiate the processing time actually involved. The processing time here involves the major processing of the raw data into a tape file from which assessments can be made with a maximum of retrieval economy.

CHAPTER 6

ASSESSMENTS

Assessment Requirements

The Federal Agricultural Rehabilitation and Development Administration jointly with the related agencies in the Provincial Governments have the task of understanding the present day agronomic situation in Canada and of guiding and encouraging the most rewarding use of these resources. Basic to the understanding of the present situation is the correction of sufficient data and its assembly in understandable form. This requirement has been clearly recognized by ARDA. The collection of data to give a first approximation of agricultural conditions across the whole of Canada is proceeding as the Canadian Inventory of Land Capability. Within this framework, the closer studies of the rural development areas give rise to data pertinent to conditions in those areas. Irrespective of the method used, the accumulated data will have to be assembled and correlated into a form that can be handled and understood. As the exact type and amount of data to be collected from the various surveys and studies is not known at this time, it is obviously impossible to give a detailed analysis of the form in which the data will be of most value. A consideration of the data which is available for compilation and detailed discussions with ARDA staff who are presently planning the collection of the data has led to the following outline of data compilation on presentations that will have to be made.

(a) Statistical compilation for Canadian Inventory.

1. Acreage measurements of different classes of land (capability for agriculture, forestry, recreation, wildlife) by townships, counties, provinces and regions.

2. Comparison of present land use with assessed land capability by townships, counties, provinces and regions.

3. Comparison of some of the socio-economic factors with assessed land capability and present land use. The socio-economic factors envisaged are those outlined in the preceding chapter regarding available statistical information. This latter step of data reduction would involve a generalization of the physical data to apply to the census subdivisions.

(b) Map and Statistical compilations and correlations for detailed analysis in project areas. This will involve similar types of analysis to those described in (a) but on a more detailed basis and with the added requirement of map representation of some types of analysis.

(c)   Analysis of data by areas or regions to show in map and statistical form areas of like characteristics for two

or more elements e.g. soil capability and farm income: soil capability, present land use, size of farms and

farm income, etc.

Precision of Area Evaluation

The above assessments often require a definition of land parcels which have a selected group of

properties. These land areas are defined by the intersection of all areas which have any of the selected properties.

Usually the boundaries of the contributing areas do not coincide, so that the resultant area has a composite boundary.

A second type of assessment requires distribution of all land in a region into categories. The categories would be

defined by rules operating on a selected set of properties. These assessment techniques require the measurement of

area from maps. They also require the recognition of intersection of overlaid areas.

If a computer mapping system were to be used, and the general principles of input data processing

co-ordinate system and coding described above had proved satisfactory, the retrieval method would essentially be

the scanning of the grid point representation of boundaries. Study was completed to estimate the error involved in

the estimate of areas of the irregular shapes represented by the grid point method. The error from this source appears

to depend directly on the area associated with the grid unit and on the square root of the number of points needed

to record the boundary. If the boundary is described by a simply connected line then the error due to boundary

length appears to be independent of shape. Preliminary study indicates that the error from this source is considerably

less than ± 1 acre per mile of boundary. This is considered to be within the limits of error incurred when the same

task is done manually.

System Capability

The computer mapping system would be capable of four basic modes of data retrieval.

Mode 1 - Two Map Set - Area Compilation.

This assessment would provide a listing on one map of all the types that occur within specific

boundaries indicated on the second map, and a readout of the acres in each division. This assessment has immediate

application to the acreage measurements of different classes of land by townships, counties, provinces and regions.

A particular land capability map would be compared with the appropriate census subdivision map and all land types

within one census subdivision would be tabulated and their areas measured in acres. Some thought, however, should

be given to the statistics required from this assessment as the classes of land may be subdivided and differentiated

by numerous factors. This could result in an output of a large volume of statistics that were not required. The

grouping of some classes and subdivisions of land capability data could be done within the computer system and only the output required actually printed. If, however, at any time a particular type of detailed land description, e.g. eroded land, needed to be isolated, this could easily be done.

Mode 2 - Three Map Set - Area Compilation.

In this method of assessment, the types of overlap that occur are tabulated and the areas of each specific type of overlap that fall within areas defined on the third map set are printed. In this category lies the comparison of present land use with assessed land capability by townships, counties, provinces and regions. Again, note should be taken that the total statistical record would be a large volume, and selection of the required output is necessary.

Mode 3 - Two Map Set - Area Compilation with Related Factors.

In this type of assessment, a number of variables can apply to any one area denoted on one map set. A desired profile of these variables can be located and area tabulation of the second set for the appropriate area can be printed. A simpler form of Mode 3 is essentially a mode 1, two-map compilation with the variable factors printed out at the head of each appropriate division. Into this category would fall the comparison of some of the socio-economic factors with the assessed land capability and present land use. If desired the physical data can be averaged and the character of the census division thus calculated. This character can then be compared with the appropriate socio-economic factors.

Mode 4 - Three Map Set Area Compilation with Multiple Factors applying to one map set.

This type of assessment is essentially the same as Mode 3, but with the variable factors related to a compilation made from the first two map sets.

The type of data used in the above assessments does not alter the nature of the assessment. The same functions can be performed whether the data is complex and refers to a small area or is coarse and refers to a larger area. Complex assessments may be made by combination of the basic modes.

Depending on the output of statistics required, the fine or coarse scan may be used for assessments. In the initial acreage assessments it may be necessary to reference every recorded point on the tape reels and analyze them. This fine scan would provide detailed data across Canada or for that part of Canada selected for this type of assessment. If, however, small subdivisions of the data were not required it would be most desirable to use a coarser scan, with only every other, or every third point, referenced. The economy of this can be emphasized by saying that by referencing every point instead of every other point the time involved is squared.

It may be that the information put into the computer system has varying degrees of reliability. If required, the degrees of reliability can be noted as the information is put in the system and the printed output of compiled statistics can indicate the degrees of reliability of the information compiled. If the reliability varies from area to area, instructions can be given to the computer to complete assessments only in those areas where the information is of a certain reliability, and to indicate those areas where it is not.

If required, data regarding an area can be simulated and compared with existing data. In this manner the effect of proposed changes in any one area could be examined.

If data changed in any one area at any time the change of data could be input into the system in the normal manner and the previous assessment re-run and the effect of the change determined.

The computer mapping system would also have the capability of data distribution around a point. If data can be supplied to the system which can be related to a point of known co-ordinates, and a network of these points can be established, the boundaries between various factors can be optimized in the computer system and compared with other data as if they were map data. This facility may be of particular value if the steps are taken to identify the headquarters location of farms. It is estimated that this could be done for over 75% of Canada if reduction of raw material presently in the hands of the Dominion Bureau of Statistics can be arranged. Given this information, a computer could record the pertinent factors regarding each farm from the census return from that farm. This would give a point-by-point analysis of the agricultural area from which comparisons could be made to the relatively detailed land capability, forest capability and present land use data.

Machine Requirements

It is estimated that all of the above assessments could be adequately carried out by a medium scale computer with a capacity of the IBM 1401. While it is not possible at this time to determine accurately the computer time involved, unsupported opinion indicates that the order of magnitude of time for a Mode 1, two-map set area compilation of the entire Canadian Inventory area across Canada would be in the order of five to six hours. The more complicated assessments can be suggested pro-rata to this figure.

CHAPTER 7

OUTPUT

Tabular Output

A great majority of the results from the computer mapping system would be in the form of statistical tables. These present no technical problems at all and can be produced with existing equipment. The most convenient form of output would be on the standard machine processed continuous sheets which can be later separated and bound. If required, sheets with a particular form overprinted on them can be arranged at a very small cost. Separate sheets of existing forms can be used on existing equipment, but this method is considerably slower than the printed continuous strip and no reason for its application can be seen at this time.

It is recommended that a study be made regarding the desirability of recording selected types of output in machine readable form. It is most probable that certain of the statistical results could be incorporated in later assessments. It would be most convenient to have a printout of results and simultaneously either punch out to IBM punch cards or storage on magnetic tapes. These could be appropriately indexed and become part of the master file. The selection of which data to keep will become possible when the required assessments are known in more detail.

No cost is given at this point for printing or card punching machines as they are considered to be part of a normal computer installation.

Graphic Output

Graphical output in the terms of the computer mapping system can be thought of as primarily the production of maps illustrating the results of particular data compilations and assessments. Given the facilities to produce maps, however, there is no reason why the same facilities should not be employed in the production of graphs or diagrams illustrating the statistical results.

Three types of instruments are commonly used in conjunction with computers to produce graphic information. These are the printer, the drum type plotter and the table plotter.

Using the normal printer a number of coarse representations of map data can be achieved. Perhaps the simplest way to achieve map representation is to build up lines on the printer paper using a single symbol. The end result is a very coarse line, but if the area represented is small enough and the scale of print out is large enough, the line joining the plotted points will in fact be a very accurate boundary. A development of

this technique is for different symbols to be used on lines of different values though this has more application to problems requiring a contour type output. For the boundaries of characterized areas, the character of each individual area could be printed in its center. Another technique for handling area data in the printer type output is the block printing of one symbol over the area concerned. Immediately adjacent is a similar block print out using a different symbol, the boundary between the two being determined not from an obvious line but by an examination of the change of value of the symbols, which completely cover the sheet. This system has the advantage of a graduation of characters, and the indication of wide boundary zones.

The main disadvantage of the printer output is that if accurate representation of anything but coarse data is required, the scale that has to be used demands a great deal of paper to be used with a subsequent difficulty in handling and the necessity of a reduction step.

The second type of graphical output unit that can be described is the drum type plotter. This has basically the same principle as the normal printer but instead of the printed symbol a line segment is drawn. As the step size is frequently .005 inches, an increment can be made at the speed of 18,000 line segments per minute on paper which is up to $29\frac{1}{2}$ inches wide by 120 feet in length. This type of printer has many applications. It can be put on-line to both small and medium scale computers such as the IBM 1620 and the IBM 1401 respectively. The cost of drum type plotters varies between \$5,000 and \$35,000. Various models are available to answer varying requirements.

The third type of plotter referred to is the table plotter. This consists of a printing or tracing head which is guided on a carriage over the surface of a plotting table. The print out is either in the form of carefully spaced printed points with an overprinted symbol or in the form of a pen, pencil or scribed line. In general they are more accurate and work to a closer tolerance than the drum printer or normal printer. Accuracy is frequently in the range to .002 inches at any point on the plotting surface. The plotting table size is frequently 30 inches x 30 inches but surfaces of up to 78 inches x 150 inches are available. These are essentially off-line analogue devices where the required magnetic tape is taken from the computer and placed on a tape drive alongside the table. In this way main computer time is not used in producing the required maps. Manual control of the pointer is usually possible from a control panel located alongside the table. This facilitates the plotting of co-ordinate points and the establishing of a frame of reference of a map. Visual read out of the x and y co-ordinates can usually be made on the console. This latter feature is most valuable for re-entry of position of co-ordinates during manual setups. Speed of operation is high and is available up to speeds of 500 inches of line traced per

minute. Typical models of this type of plotter are made by Gerber Scientific Instrument Company, Keuffel and Esser Company, Contraves of Switzerland, The Wild Company of Canada Limited and James W. Stevenson and Company (Coradi of Switzerland). The cost of each instrument varies from $35,000 to $120,000 depending on the degree of precision and capability desired.

Summary of Output Units

All of the above devices are stock items and could be purchased as the requirement arose. If graphic output is desired from the computer mapping system the nature of the data to be represented and the precision that is required would determine the type of plotter necessary. It can be noted, however, that the development of plotting instruments is very rapid. Many of the instruments that exist today did not exist two years ago. While the present plotters have the capability of producing technical output to handle all requirements of the computer mapping system it is recommended that the type of plotter to be obtained is not decided until the nature of the output required is known in detail and that a further assessment of the available plotters be made at that time.

CHAPTER 8

SUMMARY AND RECOMMENDATION FOR THE COMPUTER MAPPING SYSTEM

Input System

A review of the possible input systems leads to the conclusion that the requirements can best be met by a system comprising a drum type of scanning transcription unit used in conjunction with a simple non-automatic tracing type of transcription unit, the former linked to a buffer unit and a magnetic tape drive, the latter with an output of punch cards to be merged with the tape file. This system has the following advantages:

1. Without further instrumentation it will have the capacity to handle an input of 1000 maps per year.

2. The magnetic tape drive, core buffer unit and the card punch are standard units. The non-automatic tracing transcription device is an existing unit that can be purchased ready made. The drum type of scanning transcription unit can be constructed from existing instrumentation. The research required for the realization of this system is not that directed to the working of new methods but that of the control of existing and reliable instruments. With this type of instrumentation design performance can be gauged in a research phase and system reliability should be high.

Input Processing

Routine reduction of input data to a master tape file will require the capacity of a computer with the capability of the IBM 7000 series. Initial input data processing, and trials and tests that may be carried out for this purpose, can be completed on a medium scale computer facility such as offered by the IBM 1401 series. Further research is needed in this area to determine the detail of tape file compression and the exact use of the co-ordinate system. It is considered that the methods outlined in the text will be basic to these studies and that while the programming task will be substantial, no major problems will be encountered. It is estimated that three to four man-years of programming will be required in this area, though with a suitable staff this can be completed within a period of eighteen months. The first nine months of the work will require the services of senior systems personnel, thereafter regular programming personnel can be satisfactorily employed.

Assessments

The selection and programming of assessments, i.e., calculation and tabulation of output data, must proceed in close co-operation with ARDA and provincial authorities. It is recommended that a team comprised of an agro-economist and a mathematician-statistician, with a suitable backup of programmers, be assembled to

make initial recommendations regarding the assessments that are possible and pertinent, and constantly review the need for assessments that are being made and the form in which they can be best presented.

Output

The output of the data presents no particular technical problems. Statistical data can be output on the printers associated with the computers and if graphical output is desired, one of the existing plotters can be utilized, depending on the requirement.

Recommendation for Implementation

It is recommended that the following outline be followed in implementing the development of the computer mapping system.

1. View and report on the companies capable of producing the drum type scanning transcription unit. Negotiate a research and development contract for the research necessary to the system and the production of the system.

2. Start research on input processing for drum scanning type of input system.

3. Procure non-automatic tracing type of transcription type unit and couple to existing punch card unit.

4. Form a team of agrologist, statisticians and programmers to research required assessments.

5. Convert non-automatic tracing type of transcription unit to semi-automatic, still utilizing punch card output.

6. Start output of map data, preferably from a pilot area. Relay this data to the team working on input processing and to the team working on assessments.

7. Photographically reduce tracing data from sample maps and scan on GIMRADA scanning device to produce experimental tape data. Relay this data to the input processing team and to the assessment team.

8. Arrange for collection of maps at optimum level of data representation across Canada.

9. Investigate statistical information and initiate transformation of raw data onto punch cards if necessary. Simultaneously investigate the possibility of point location of data.

10. Run preliminary assessments on medium scale computer facility (e.g. IBM 1401).

11. Feed back results of assessments to input processing programmers and to research being carried out on drum type of scanning transcription unit.

12. Receive and install drum type of scanning transcription unit with an associated buffer unit and magnetic tape drive.

13. Convert semi-automatic tracing type of transcription unit to normal mode.

14. Commence routine input of map data.

15. Start tape file using IBM 7090 capability for input data processing.

16. Start programming assessments with basis of tape operation.

17. Run preliminary assessments from tape data.

18. Review assessments and revise as necessary.

It is considered that steps one to eleven should be initiated in the first twelve months after commencing the development of the system. Routine assessments from magnetic tape data should be available in a further nine months.

The question of which organizations or agencies should carry out this work is beyond the scope of this report but the following observations made from the technical standpoint may be of value:

1. As detailed specifications cannot be written without further research, the production of the drum type of scanning input system would best be handled as a negotiated research and development contract. The company concerned should be chosen after the limited number of possibilities in Canada and the USA have been reviewed for experience and capability.

2. As suitable equipment exists, specifications can be drawn up for a non-automatic tracing type of transcription unit and it be let to tender.

3. Existing IBM 526 punch card units can be coupled with a non-automatic tracing unit.

4. The computer facility required for handling all parts of the system exist within the Government of Canada and, as the input system is off-line, it would be possible to process magnetic tapes at various locations as computer time became available.

5. The research into input processing and assessment requirement will require experienced senior personnel on a full time basis. This cannot be emphasized too strongly. It is further recommended that project teams be formed of such personnel and every attempt be made to maintain them so that the experience with the system in the initial stages can be cumulative.

6. As the requirement for graphic output is determined, a review should be made of the existing instrumentation and specifications should be drawn up that would allow tenders to be called for the production of any necessary new instruments.

7. The computer mapping system development should be co-ordinated by one agency, either within ARDA, or working in close consultation with ARDA.

8. The existing unit record equipment in the Department of Agriculture, Data Processing Centre would be suitable for the conversion of available raw statistical data to punch card form.

* * * * * * * * *