Volume 4

DSSAT version 3

Editors

Gerrit Hoogenboom
Paul W. Wilkens
Gordon Y. Tsuji

International Benchmark Sites Network for Agrotechnology Transfer
International Consortium for Agricultural Systems Application
University of Hawaii, Honolulu, Hawaii

IBSNAT, The International Benchmark Sites Network for Agrotechnology Transfer, was a network consisting of the contractor (University of Hawaii), its subcontractors and many global collaborators. Together they created a network of national, regional, and international agricultural research for the transfer of agrotechnology among global partners in both developed and lesser developed countries.

From 1982 to 1987, IBSNAT was a program of the U.S. Agency for International Development under a cost-reimbursement Contract, No. DAN-4054-C-00-2071-00, with the University of Hawaii. From 1987 to 1993, the contract was replaced with a Cooperative Agreement, No. DAN-4054-A-00-7081-00, between the University of Hawaii and USAID.

In 1994, the International Consortium for Agricultural Systems Application or ICASA was established as a non-profit corporation in Honolulu, Hawaii. The ICASA is governed by a Board of Directors consisting of former members of the IBSNAT network, and systems scientists from the Wageningen Agricultural University in the Netherlands and from the Agricultural Production Systems Research Unit, APSRU, of CSIRO in Australia. The ICASA network will oversee the continued development of systems tools, including DSSAT.

Correct citation: G. Hoogenboom, P.W. Wilkens, and G.Y. Tsuji (eds.). 1999. DSSAT v3, volume 4. University of Hawaii, Honolulu, Hawaii.

LIBRARY OF CONGRESS 94-19296
ISBN 1-886684-04-9 (Volume 4)

DSSAT v3

VOLUME 4

Volume 4-1
Decision Support System for Agrotechnology Transfer v3.5

Volume 4-2
Spatial Analysis

Volume 4-3
Agricultural and Environmental Geographic Information System for Windows

Volume 4-4
Displaying Soil Water and Nitrogen Graphs

Volume 4-5
Creating Experimental Performance Files

TABLE OF CONTENTS

Volume 4-1. Decision Support System for Agrotechnology Transfer, v3.5

Tables

Volume 4-3. Agricultural and Environmental Geographic Information System for Windows

Volume 4-6. Concepts for Calibrating Crop Growth Models

Figures

Volume 4-7. Data Requirements for Model Evaluation and Techniques for Sampling Crop Growth and Development

Volume 4-8. Field and Laboratory Methods for the Collection of the Minimum Data Set

Volume 4-1

Decision Support System for Agrotechnology Transfer v3.5

G. Hoogenboom
P.W. Wilkens
P.K. Thornton
J.W. Jones
L.A. Hunt
D.T. Imamura

The University of Georgia, International Fertilizer Development Center,
International Livestock Research Institute, University of Florida,
University of Guelph, and University of Hawaii
International Consortium for Agricultural Systems Applications

CHAPTER ONE. INTRODUCTION

The Decision Support System for Agrotechnology Transfer v3 was released in 1994 with three volumes of users' guides that described the background and operation of the utility programs, crop simulation models, and application programs (Tsuji et al., 1994). Since the release of DSSAT v3, further development of DSSAT continued with (1) improvements to existing crop models, (2) new crop simulation models for additional crops and (3) new application programs and data utility tools. These improvements were made in response to requests from users and participants in training courses and workshops. The updated version of DSSAT is being released as v3.5. We will refer to v3 as the suite of versions from the initial release version in 1994, which will be referred to as v3.0, through to the updated version, v3.5.

Volume 4 of the DSSAT v3 users' guides presents the changes, improvements, and additions included in DSSAT v3.5. The first chapter of Volume 4 describes the changes made to the models as well as the input file structure. The second chapter describes a spatial application program that includes spatial interpolation of simulated data using kriging and other techniques. The third chapter describes an interface between DSSAT and the Geographic Information System (GIS) ArcView. The result is AEGIS/WIN which was developed for operation in a Windows environment. The fourth chapter presents a graphics program for display of simulated soil nitrogen and moisture over time, called NSHOW. This program is especially useful as an education tool to present simulation results to people who are not familiar with crop simulation models. The fifth chapter describes the utility TACREATE for entering summary and time series experimental data. The sixth chapter provides guidelines for model evaluation. The final chapter discusses procedures for experimental data collection for model evaluation. The appendix is a revised version of IBSNAT Technical Report 2, initially printed in 1990, titled "Field and Laboratory Methods for the Collection of the IBSNAT Minimum Data Set for DSSAT."

CHAPTER TWO.
FILE FORMAT

INPUT FILES

The original documentation of the format and structure of the input and output files for the crop models in DSSAT was presented by Jones et al. (1994). As the models were improved and new features added, the input and output files were modified to accommodate these changes. In addition, new applications of the models required modification of the input and output files. For example, spatial application of the models require input of specific coordinate (longitude and latitude) information of a particular experiment. These new spatial input variables, however, are not required to run the models with minimum inputs. The input module of the crop simulation models is backwards compatible with the DSSAT v3.0 file format and can handle files that conform to either DSSAT v3.0 or DSSAT v3.5 file format.

The most significant format change is in the *FIELDS section of the experimental details file, referred to as FILEX. A second tier has been added to handle the spatial aspect for the entire experiment or for each individual treatment of a particular experiment. In the latter case multiple levels can be defined in the *FIELDS section. New variables added include the X- and Y-coordinates (XCRD, YCRD), using any reference system, elevation above sea level (ELEV), area of the field (AREA), slope length of the field (SLEN), field length to width ratio (FLWR) and slope aspect (SLAS). The slope aspect was removed from the first tier in order for slope to be defined completely.

The *FIELDS section is organized by tiers and not by level number. The following presents the changes in format:

	Variable[1]	Header[2]	Format[3]
***FIELDS**			
(*first tier*)			
Slope from horizontal (%)	SLOPE	FLSA	1 R 5 1
(*second tier*)			
Field level	none	none	0 I 3
X-coordinate (any reference system)	XCRD	XCRD	0 R 15 5
Y-coordinate (any reference system)	YCRD	YCRD	1 R 15 5
Elevation, m above sea level	ELEV	ELEV	1 R 9 2
Area of field, m^2 (0 if point/raster)	AREA	AREA	1 R 17 1
Length of slope, m	SLEN	SLEN	1 R 5 0

Field length–width ratio	FLWR	FLWR	1 R 5 1
Slope aspect, degrees clockwise from North	SLAS	SLAS	1 R 5 1

[1] Abbreviations used as variable names in the DSSAT models.
[2] Abbreviations suggested for use in header lines (those designated with "@") within the file.
[3] Formats are written as follows: number of leading spaces, variable type (C=Character, R=Real, I=Integer), variable width, and (if real) number of decimal places.

In the *INITIAL CONDITIONS section several new variables were added. In the DSSAT v3 models, the initial conditions for initial organic residue in the profile at planting or start of simulation were defined in the *RESIDUES AND OTHER ORGANIC MATERIALS section. Although the initialization in this section can be used for sensitivity analysis and single season runs, there are potential errors in initialization when the models are run in a crop rotation or sequence analysis mode. Therefore, the initial conditions that define initial residue at the start of simulation have been moved to the *INITIAL CONDITIONS section. New variables added include initial crop residue (ICRES), initial residue N (ICREN) and P (ICREP) content, initial residue incorporation (ICRIP), and initial residue incorporation depth (ICRID). Another new variable is initial water table depth ICWD)

The following describes the changes in the initial conditions section:

	Variable[1]	Header[2]	Format[3]
***INITIAL CONDITIONS**			
Initial water table depth, cm	ICWD	ICWD	1 R 5 1
Initial crop residue applied, kg/ha	ICRES	ICRES	1 R 5 0
Initial residue N content, %	ICREN	ICREN	1 R 5 2
Initial residue P content, %	ICREP	ICREP	1 R 5 2
Initial residue incorporation, %	ICRIP	ICRIP	1 R 5 0
Initial residue incorporation depth, cm	ICRID	ICRID	1 R 5 0

In the *PLANTING DETAILS section one new variable was added to define initial sprout length for tubers (SPRL). This variable was previously stored in transplant age (PAGE), which caused input conflicts.

***PLANTING DETAILS**			
Initial sprout length	SPRLAP	SPRL	1 R 5 1

In the *RESIDUES AND ORGANIC MATERIALS section one new variable was added to define method of residue incorporation (RMET).

***RESIDUES AND OTHER ORGANIC MATERIALS**			
Residue method of incorporation code	RMET	RMET	1 C 5

In the *HARVEST DETAILS section one new variable was added to define takeoff percentage of harvest by-product (HBPC). This is especially applicable when, in addition to the grains, the straw is also harvested.

```
*HARVEST DETAILS
Harvest by-product takeoff percentage, %     HBPC    HBPC   1 R 5 0
```

In the *SIMULATION CONTROLS section several new options and methods were added. Chemical (CHEM) and tillage (TILL) simulations are new options for future applications of the models. A new method was added for the water balance simulations (MEHYD). The default method is the "Ritchie" soil water balance simulation. In outputs, two new switches were added; CHOUT controls the chemical time series output file and OPOUT controls an operations output file.

	Variable[1]	Header[2]	Format[3]
***SIMULATION CONTROLS**			
Line 2			
@N OPTIONS			
Chemical applications	ISWCHE	CHEM	5 C 1
Tillage applications	ISWTIL	TILL	5 C 1
Line 3			
@N METHODS			
Hydrology	MEHYD	MEHYD	5 C 1
R = Ritchie method (Default)			
Line 5			
@N OUTPUTS			
Details—individual aspects			
Chemical (Y = yes; N = no)	IDETH	CHOUT	5 C 1
Operational (Y = yes; N = no)	IDETR	OPOUT	5 C 1

[1] Abbreviations used as variable names in the DSSAT models.
[2] Abbreviations suggested for use in header lines (those designated with "@") within the file.
[3] Formats are written as follows: number of leading spaces, variable type (C=Character, R=Real, I=Integer), variable width, and (if real) number of decimal places

No changes were made to the other input files, including the weather and soil data files, and the summary and time series experimental data files.

Output Files

Minor changes were made to the output files generated by the models, including the files for simulated growth, water, nitrogen and carbon balances. These changes were

implemented internally to the crop models and do not affect the user. The definitions of the headers in the output files can be found in the data code file, DATA.CDE, located in the C:\DSSAT35 directory. This file can be accessed from the DSSAT shell under data utilities.

TABLE 1. EXAMPLE FILEX IN DSSAT v3.5 FORMAT

```
*EXP.DETAILS: UFGA9801SB NEW IBSNAT 3.5 INPUT EXAMPLE

*GENERAL
@PEOPLE
 BOOTE,K.J. JONES,J.W. HAMMOND,L.C.
@ADDRESS
 UNIVERSITY OF FLORIDA, GAINESVILLE, FL, USA
@SITE
 IRR.PARK,UF.CAMPUS 29.63;-82.37;40.;FLA

*TREATMENTS                        -------------FACTOR LEVELS------------
@N R O C TNAME..................... CU FL SA IC MP MI MF MR MC MT ME MH SM
 1 1 0 0 IRRIGATED                  1  1  0  1  1  1  0  1  1  1  1  0  1
 2 1 0 0 NON-IRRIGATED              1  2  0  1  1  0  0  1  0  0  0  0  1

*CULTIVARS
@C CR INGENO CNAME
 1 SB IB0001 BRAGG

*FIELDS
@L ID_FIELD WSTA .....FLSA .FLOB .FLDT .FLDD .FLDS .FLST SLTX .SLDP ID_SOIL...
 1 UFGA0004 UFGA       5.0     0 DR000     0     0 00000 SA     180 IBSB960001
 2 UFGA0004 UFGA       7.5     0 DR000     0     0 00000 SA     180 IBSB960001
@L ...........XCRD ...........YCRD .....ELEV .............AREA.SLEN .FLWR .SLAS
 1        29.63000       -82.37000     20.00            8026.8 110.   1.0   0.0
 1        29.63000       -82.37000     20.00            8026.8 110.   1.0   0.0

*INITIAL CONDITIONS
@C   PCR ICDAT  ICRT  ICND  ICRN  ICRE  ICWD ICRES ICREN ICREP ICRIP ICRID
 1    SB 78166   100   -99  1.00  1.00 150.0 1000.  2.00  1.00  100.    5.
@C  ICBL  SH2O  SNH4  SNO3
 1     5 0.086   0.6   1.5
 1    15 0.086   0.6   1.5
 1    30 0.086   0.6   1.5
 1    45 0.086   0.6   1.5
 1    60 0.086   0.6   1.5
 1    90 0.076   0.6   0.6
 1   120 0.076   0.6   0.5
 1   150 0.130   0.6   0.5
 1   180 0.258   0.6   0.5

*PLANTING DETAILS
@P PDATE EDATE PPOP PPOE PLME PLDS PLRS PLRD PLDP PLWT PAGE PENV  PLPH  SPRL
 1 78166  -99  29.9 29.9  S    R    91   0    4.0 -99  -99 -99.0 -99.0 -99.0

*IRRIGATION AND WATER MANAGEMENT
@I  EFIR  IDEP  ITHR  IEPT  IOFF  IAME  IAMT
 1  0.75   -99   -99   -99   -99   -99   -99
@I IDATE  IROP  IRVAL
 1 78181 IR001    13
 1 78230 IR001    13
 1 78235 IR001    13
 1 78237 IR001    13
 1 78240 IR001    11
 1 78242 IR001    11
 1 78244 IR001    11
 1 78246 IR001    11
 1 78250 IR001    11
```

```
 1 78253 IR001    11
 1 78256 IR001     8
 1 78259 IR001     8
 1 78262 IR001     8
 1 78265 IR001     8
 1 78269 IR001     7
 1 78272 IR001     8
 1 78279 IR001     7
 1 78283 IR001     8
 1 78294 IR001    10

*RESIDUES AND OTHER ORGANIC MATERIALS
@R RDATE  RCOD  RAMT  RESN  RESP  RESK  RINP  RDEP  RMET
 1 78166 RE001  1000  0.80 -9.00 -9.00   100    15 AP002

*TILLAGE
@T TDATE TIMPL  TDEP
 1 78166 TI987  10.0
 1 78200 TI001   5.0

*CHEMICAL
@T CDATE CHCOD CHAMT CHMET CHDEP CHT..
 1 78166 CH002  1.00 AP001   5.0 WEEDS
 1 78199 CH003  0.50 AP005  10.0 WEEDS

*SIMULATION CONTROLS
@N GENERAL    NYERS NREPS START SDATE RSEED SNAME.....................
 1 GE             1     1     S 78166  2150 BRAGG, IRRIGATED & NON-IR
@N OPTIONS    WATER NITRO SYMBI PHOSP POTAS DISES  CHEM  TILL
 1 OP             Y     Y     Y     N     N     N     Y     N
@N METHODS    WTHER INCON LIGHT EVAPO INFIL PHOTO HYDRO
 1 ME             M     M     E     R     S     C     R
@N MANAGEMENT PLANT IRRIG FERTI RESID HARVS
 1 MA             R     R     R     R     M
@N OUTPUTS    FNAME OVVEW SUMRY FROPT GROUT CAOUT WAOUT NIOUT MIOUT DIOUT LONG CHOUT OPOUT
 1 OU             N     Y     Y     3     Y     Y     Y     Y     N     N    Y     Y     Y

@  AUTOMATIC MANAGEMENT
@N PLANTING   PFRST PLAST PH2OL PH2OU PH2OD PSTMX PSTMN
 1 PL           155   200    40   100    30    40    10
@N IRRIGATION IMDEP ITHRL ITHRU IROFF IMETH IRAMT IREFF
 1 IR            30    50   100 GS000 IR001    10  0.75
@N NITROGEN   NMDEP NMTHR NAMNT NCODE NAOFF
 1 NI            30    50    25 FE001 GS000
@N RESIDUES   RIPCN RTIME RIDEP
 1 RE           100     1    20
@N HARVEST    HFRST HLAST HPCNP HPCNR
 1 HA             0   365   100     0
```

[1] *New variables and headers are shown in bold*

CHAPTER THREE.
NEW CROP SIMULATION MODELS

The original DSSAT v3.0 included the grain cereal model CERES for simulation of growth and development for maize, wheat, barley, sorghum, and millet; the CERES-Rice model for simulation of rice growth and development; the grain legume model CROPGRO for simulation of soybean, peanut and dry bean growth and development and the CROPSIM model for simulation of cassava growth and development. The operation of these models has been documented by Hoogenboom et al. (1994) in Volume 2-2 of the DSSAT v3 Users' guides. A more detailed description of the CERES models is presented in Ritchie et al. (1998), where the growth and development simulation procedures are described for the various cereal crops. A detailed description of the CROPGRO models can be found in Boote et al. (1998), where the growth and development simulation procedures for the grain legume models are presented.

DSSAT v3.5 includes six new crop simulation models. They are potato, sugarcane, sunflower, chickpea, tomato and pasture. The first three models have independent source code for each crop, while the last three models are based on the generic CROPGRO model.

POTATO

The potato model uses the SUBSTOR source code for simulation of growth, development and tuber yield. The code was developed by T.S. Griffin, B.S. Johnson, and J.T. Ritchie. The SUBSTOR model included in DSSAT v3.5 is an improved version of the original model included in DSSAT v2.1. A detailed description of the actual model can be found in Singh et al. (1998). SUBSTOR simulates potato growth and development, as well as the plant and soil water and nitrogen balance. The soil water and nitrogen balance simulation routines are identical to the soil water and nitrogen routines of the CROPGRO and CERES simulation models (Godwin et al., 1998; Ritchie, 1998).

The user has to define a harvest date in the experimental details input file, as the model does not predict physiological and harvest maturity. The potato model uses six coefficients to define the characteristics of a cultivar. Table 2 shows an example of the current cultivars defined for potato. Required genetic coefficients include:

VAR#	Identification code or number for a specific cultivar
VAR-NAME	Name of cultivar
ECO#	Ecotype code or this cultivar, points to the Ecotype in the ECO file (currently not used).

G2	Leaf area expansion rate in degree days
G3	Potential tuber growth rate
G4	Currently not used in the model
PD	Index that suppresses tuber growth during the period that immediately follows tuber induction
P2	Index that relates photoperiod response to tuber initiation
TC	Upper critical temperature for tuber initiation

TABLE 2. GENETIC COEFFICIENTS FOR SUBSTOR-POTATO (PTSUB980.CUL)

```
*POTATO GENOTYPE COEFFICIENTS—PTSUB980 MODEL

@VAR#  VAR-NAME........   ECO#     G2    G3    G4    PD    P2    TC
IB0001 MAJESTIC         IB0001  2000.  22.5  0.20   0.8   0.6  17.0
IB0002 SEBAGO           IB0001  2000.  22.5  0.20   0.7   0.8  15.0
IB0003 RUSSET BURBANK   IB0001  2000.  22.5  0.20   0.6   0.6  17.0
IB0004 KATHADIN         IB0001  2000.  25.0  0.20   0.7   0.6  19.0
IB0005 ATLANTIC         IB0001  2000.  25.0  0.20   0.9   0.6  17.0
IB0006 MARIS PIPER      IB0001  2000.  25.0  0.20   0.8   0.4  17.0
IB0007 KING EDWARD      IB0001  2000.  22.5  0.20   1.0   0.6  17.0
IB0008 DESIREE          IB0001  2000.  25.0  0.20   0.9   0.6  17.0
IB0009 LT-1             IB0001  2000.  25.0  0.20   0.9   0.8  21.0
IB0010 C14-343          IB0001  2000.  25.0  0.20   0.9   0.4  21.0
IB0011 NORCHIP          IB0001  2000.  25.0  0.20   1.0   0.4  17.0
IB0012 SHEPODY          IB0001  2000.  25.0  0.20   0.7   0.6  19.0
```

SUGARCANE

The sugarcane model is based on the CANEGRO model, developed by G.A. Kiker and G. Inman-Bamber of the South African Sugar Association Experiment Station . The model is a hybrid between CERES and CROPGRO models. CANEGRO simulates sugarcane physiology, population dynamics, and water relations. The model has three options for calculating potential evapotranspiration. In addition to the Priestley-Taylor and FAO-Penman methods, the model can also simulate the Penman-Monteith potential evapotranspiration. This method requires daily wind speed and dew point temperature, as well as the other DSSAT weather variables as input. CANEGRO currently does not simulate the soil nitrogen balance and organic carbon. The model predicts only vegetative growth stages, as reproductive stages are not economically critical for farmers. The current model, for instance, does not predict flowering date or maturity date. The user therefore needs to define a harvest date for the crop in the experimental details file. If the harvest date is not defined, the model will use the cultivar coefficient P1, degree days from emergence to harvest maturity, to end the simulation.

The genetic coefficients have been divided into two groups following the same format as the CROPGRO and other models of DSSAT v3.5. Coefficients included in the ecotype file include a root centimeter to gram ratio to convert root length into root weigh (RTCMPG), a sensitivity factor of yield to water table height (SENSWT); and a critical water table height (CRTWT).

The sugarcane model uses seven coefficients to define the characteristics of a cultivar. Table 3 shows an example of the current cultivars defined for sugarcane. Required genetic coefficients include:

VAR#	Identification code or number for a specific cultivar
VAR-NAME	Name of cultivar
ECO#	Ecotype code or this cultivar, points to the Ecotype in the ECO file (currently not used).
P1	Degree days from emergence to harvest maturity
RATPOT	Maximum # of ratoon crops before reseeding
LFMAX	Maximum # of green leaves on a shoot
G1	General Leaf shape to be used to calculate the maximum area, leaf width, and total leaf populations. Users can choose either 1.0, 2.0, or 3.0, depending on the cultivar characteristics listed below: G1 = 1.0 corresponds to the NCO376 and N14 leaf type—high population (greater than 13 plants/m^2) and narrow leaf (less than 30 mm in width) G1 = 2.0 corresponds to the N12 leaf type—medium population (10 to 13 plants/m^2) and medium leaf width (30 to 50 mm in width) G1 = 3.0 corresponds to the R570 leaf type—low population (less than 10 plants/m^2) and broad leaf (greater than 50 mm in width).
PI1	Phyllochron interval #1. When 0 < heat units < DDTPI
PI2	Phyllochron interval #2. When heat units > DTTPI
DTTPI	Degree day threshold between Phyllochron interval 1 and 2

SUNFLOWER

The sunflower model is based on the OILCROP model, developed by A.J. Hall from the University of Buenos Aires, Argentina and F.J. Villalobos of Cordoba University, Spain in collaboration with J.T. Ritchie of Michigan State University. OILCROP-SUN is a process-oriented, management level model that simulates development, growth and yield of the sunflower crop, as well as the plant and soil water and nitrogen balances. The soil water and nitrogen balance are identical to the CERES and CROPGRO simulation models (Godwin and Singh, 1998; Ritchie, 1998). An earlier version of the

model, including a detailed explanation of the computer code, was described in Villalobos et al. (1992, 1993). Villalobos et al. (1996) presented a description of the most recent version of the model.

TABLE 3. GENETIC COEFFICIENTS FOR CANEGRO-SUGARCANE (SCCAN980.CUL)

```
*SUGARCANE GENOTYPE COEFFICIENTS—SCCAN980 MODEL

@VAR#  VAR-NAME........  ECO#    P1  RATPT  LFMAX   G1   PI1    PI2   DTPI
!                              1     2      3     4     5      6     7
IB0001 NCo376          SC0001 8500.    5.0   11.0  1.0  109.0  169.0 1526.
IB0002 N14             SC0001 8500.    5.0   11.0  2.0  109.0  169.0 1526.
IB0003 GEOFF'S FAV     SC0001 8500.    5.0   11.0  1.0   71.0  133.0  800.
```

To simulate sunflower growth, development, and yield, the model takes into account the following processes:

- phenological development, especially as it is affected by genotype, temperature and photoperiod
- extension growth of leaves and roots
- biomass accumulation and partitioning, especially as affected by phonological development radiation, temperature, and water and nitrogen stresses
- soil water balance and water use by the crop
- soil nitrogen transformations, and crop nitrogen uptake and partitioning among plant parts.

The sunflower model uses six genetic coefficients to define the characteristics of a cultivar. Table 4 shows an example of the current cultivars defined for sunflower. Required genetic coefficients include:

P1 Duration of juvenile phase (in degree days, with a base temperature of 4 °C)
P2 Amount (in days/hour) that development is slowed when crop is grown in a photoperiod shorter than the optimum (which is considered to be 15 hours)
P5 Duration of the first anthesis-physiological maturity stage (in degree days above a base of 4 °C)
G2 Maximum possible number of grains per head (measured in plants grown under optimum conditions and low plant population density)
G3 Potential kernel growth rate during the linear kernel filling phase (in mg/day, measured in plants grown under optimum conditions and low plant population density).
O1 Maximum kernel oil content (%)

A rough estimate of G3 may be obtained by estimating mature kernel weight (KW) from cultivar mature grain weight (GW). To do this, assume KW/GW = 0.75 (a reasonable approximation for many high oil content hybrids) and divide KW by the duration of the kernel growth phase (to a reasonable approximation, grain filling phase—5 days).

Algorithms for estimation of grain oil content, as implemented and tested in OILCROP-SUN can only capture some response of grain oil content to environmental conditions. Estimates of O1 can be obtained from the highest reported grain oil contents for a cultivar, and adjusted for dry weight differences between whole grain and kernel, assuming the oil is stored in the latter.

TABLE 4. GENETIC COEFFICIENTS FOR OILCROP-SUNFLOWER (SUOIL980.CUL)

```
*SUNFLOWER GENOTYPE COEFFICIENTS—SUOIL980 MODEL
@VAR#  VRNAME.........   ECO#    P1    P2    P5    G2    G3    O1
!                              1     2     3     4     5     6
IB0003 894              IB0001 274.0  9.33 600.0 2800.  1.70    65
IB0005 SUNGRO-380       IB0001 295.0 12.32 600.0 2800.  1.70    65
IB0008 SH-3000          IB0001 250.0  2.45 600.0 2800.  1.70    65
IB0009 E-353            IB0001 245.0  3.74 700.0 1500.  2.90    85
IB0011 FLORASOL         IB0001 360.0  1.60 600.0 2800.  1.70    65
IB0013 SUNGRO-385       IB0001 325.0 12.32 700.0 2800.  1.90    65
IB0014 SW-101           IB0001 210.0  3.74 800.0 1000.  2.50    65
IB0015 S-530            IB0001 335.0 12.32 700.0 3700.  1.90    65
IB0016 G-100            IB0001 309.0  2.99 732.0 2800.  1.81    65
IB0017 CONTIFLOR-3      IB0001 325.5  0.14 712.0 2900.  1.67    65
```

CHICKPEA

Chickpea is an important legume crop in the semi-arid tropics and an important source of protein. The crop seems to grow well under rainfed conditions, using residual soil water from a previous crop grown during the rainy season. Singh et al. (1996) developed the CHIKPGRO model based on a modified version of the PNUTGRO model that predicts hedgerow photosynthesis. CHIKPGRO predicts potential yield and water use for semi-arid climates. However, CHIKPGRO is not compatible with the current generation of crop models of DSSAT v3 and higher. The generic grain legume model CROPGRO was therefore converted to include chickpea as a new crop.

The simulation of growth, development, and yield of chickpea is identical to the simulation of the other grain legume crops, e.g., soybean, peanut and dry bean (Boote et al., 1998). One difference is that chickpea is a long day plant, while the other crops

are short day plants. The model requires three genotype coefficient files, including the species file (CHGRO980.SPE), the ecotype file (CHGRO980.ECO) and the cultivar file (CHGRO980.CUL).

The chickpea model uses 15 genetic coefficients to define the characteristics of a cultivar. Table 5 shows an example of the current cultivars defined for chickpea. Required genetic coefficients include:

ECO#	Code for the ecotype to which this cultivar belongs (see CHGRO980.ECO file)
CSDL	Critical Short Day Length below which reproductive development progresses WITH daylength effect (for long day plants) (hour)
PPSEN	Slope of the relative response of development to photoperiod with time (negative for long day plants) (1/hour)
EM-FL	Time between plant emergence and flower appearance (R1) (photothermal days)
FL-SH	Time between first flower and first pod (R3) (photothermal days)
FL-SD	Time between first flower and first seed (R5) (photothermal days)
SD-PM	Time between first seed (R5) and physiological maturity (R7) (photothermal days)
FL-LF	Time between first flower (R1) and end of leaf expansion (photothermal days)
LFMAX	Maximum leaf photosynthesis rate at 30 °C, 350 vpm CO_2, and high light (mg CO_2/m^2 s)
SLAVR	Specific leaf area of cultivar under standard growth conditions (cm^2/g)
SIZLF	Maximum size of full leaf (three leaflets) (cm^2)
XFRT	Maximum fraction of daily growth that is partitioned to seed + shell
WTPSD	Maximum weight per seed (g)
SFDUR	Seed filling duration for pod cohort at standard growth conditions (photothermal days)
SDPDV	Average seed per pod under standard growing conditions (#/pod)
PODUR	Time required for cultivar to reach final pod load under optimal conditions (photothermal days)

TOMATO

The tomato model is based on the generic grain legume model CROPGRO and was developed by Scholberg et al. (1996, 1997). The main difference between the grain legume model and the tomato model is that the nitrogen fixation process is not activated. The main nitrogen processes therefore include nitrogen uptake by the root system as well as nitrogen mobilization from older leaves and stems to new tissue. The original "pod" in the model is considered to be a fruit, with an extremely small shell or

TABLE 5. GENETIC COEFFICIENTS FOR CROPGRO-CHICKPEA (CHGRO980.CUL)

```
*CHICKPEA GENOTYPE COEFFICIENTS—CHGRO980 MODEL
@VAR#  VRNAME.....ECO#  CSDL PPSEN EM-FL FL-SH FL-SD SD-PM FL-LF LFMAX SLAVR SIZLF  XFRT WTPSD SFDUR SDPDV PODUR
!                          1     2     3     4     5     6     7     8     9    10    11    12    13    14    15
990001 Kabuli Type   KABULE 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
990002 Desi Type     DESI   11.00 -.143  37.0   8.0  15.0 38.00 42.00 1.700  150.  10.0  1.00 0.283  29.0  1.00  18.0
IB0001 ANNIGERI      ANNIGE 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
IB0002 K 850         DFAULT 11.00 -.143  37.0   8.0  15.0 38.00 42.00 1.700  150.  10.0  1.00 0.283  29.0  1.00  18.0
IB0003 G 130         DFAULT 11.00 -.143  38.0   8.0  15.0 38.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
IB0004 JG 74         JG 74  11.00 -.143  41.0  11.0  14.0 26.00 34.00 1.700  150.  10.0  1.00 0.190  22.0  1.60  18.0
IB0005 ICCV 88202    DFAULT 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
IB0006 ICCC 42       DFAULT 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
IB0007 ICCV 10       DFAULT 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
IB0008 ICCV 2        DFAULT 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
IB0009 ICCC 32       DFAULT 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
IB0011 ICCC 37       DFAULT 11.00 -.143  30.0   8.0  15.0 35.00 42.00 1.700  150.  10.0  1.00 0.181  29.0  1.20  18.0
```

fruit wall weight and a very high shelling percentage. The model currently does not predict fresh fruit weight.

The simulation of growth, development and yield of tomato is identical to the simulation of the grain legume crops, e.g., soybean, peanut and dry bean, except for nitrogen fixation as mentioned in the previous paragraph (Boote et al., 1998). The model requires three genotype coefficient files, including the species file (TMGRO980.SPE), the ecotype file (TMGRO980.ECO) and the cultivar file (TMGRO980.CUL).

The tomato model uses 15 genetic coefficients to define the characteristics of a cultivar. Table 6 shows an example of the current cultivars defined for tomato. Required genetic coefficients include:

ECO#	Code for the ecotype to which this cultivar belongs (see TMGRO980.ECO file)
CSDL	Critical Short Day Length below which reproductive development progresses with no daylength effect (for short day plants) (hour)
PPSEN	Slope of the relative response of development to photoperiod with time (positive for short day plants) ($hour^{-1}$)
EM-FL	Time between plant emergence and flower appearance (R1) (photothermal days)
FL-SH	Time between first flower and first pod (R3) (photothermal days)
FL-SD	Time between first flower and first seed (R5) (photothermal days)
SD-PM	Time between first seed (R5) and physiological maturity (R7) (photothermal days)
FL-LF	Time between first flower (R1) and end of leaf expansion (photothermal days)
LFMAX	Maximum leaf photosynthesis rate at 30 °C, 350 vpm CO_2, and high light ($mg\ CO_2/m^2\ s$)

SLAVR	Specific leaf area of cultivar under standard growth conditions (cm^2/g)
SIZLF	Maximum size of full leaf (three leaflets) (cm^2)
XFRT	Maximum fraction of daily growth that is partitioned to seed + shell
WTPSD	Maximum weight per seed (g)
SFDUR	Seed filling duration for pod cohort at standard growth conditions (photothermal days)
SDPDV	Average seed per pod under standard growing conditions (#/pod)
PODUR	Time required for cultivar to reach final pod load under optimal conditions (photothermal days)

PASTURE

The main objective of the pasture model is to allow for biomass cover crop simulation in crop rotation and sequence applications. *It should not be used as a stand-alone model.* The pasture model is a very simple model that predicts vegetative growth and development only. It is based on the CROPGRO model, in which all reproductive development phases have been switched off. The model does not predict any increases in biomass of either shell, seed, or pod weight. In its current format it also does not predict flowering. In addition nitrogen fixation is not simulated, although simulation of a nitrogen fixing cover crop can be included in future versions of the model. The experimental data for model development are based on several bahia grass experiments.

The simulation of vegetative growth and development of the pasture model is identical to the simulation of vegetative growth of the grain legume crops (Boote et al., 1998). The model requires three genotype coefficient files, including the species file (G0GRO980.SPE), the ecotype file (G0GRO980.ECO) and the cultivar file (G0GRO980.CUL).

The pasture model does not use any of the genetic coefficients of the cultivar file, except for the number of days to flowering and sensitivity to photoperiod. Both are set to extreme high values to avoid flowering. Table 7 shows an example of the current cultivars defined for the pasture model. Required genetic coefficients include:

ECO#	Code for the ecotype to which this cultivar belongs (see G0GRO980.ECO file)
CSDL	Critical Short Day Length below which reproductive development progresses with no daylength effect (for short day plants) (hour)
PPSEN	Slope of the relative response of development to photoperiod with time (positive for short day plants) ($hour^{-1}$)
EM-FL	Time between plant emergence and flower appearance (R1) (photothermal days)

TABLE 6. GENETIC COEFFICIENTS FOR CROPGRO-TOMATO (TMGRO980.CUL)

@VAR#	VRNAME.....	ECO#	CSDL	PPSEN	EM-FL	FL-SH	FL-SD	SD-PM	FL-LF	LFMAX	SLAVR	SIZLF	XFRT	WTPSD	SFDUR	SDPDV	PODUR
!			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
990001	Tomato Example	TM0001	12.33	0.00	23.0	8.0	17.0	50.00	50.00	1.36	350.	300.0	0.75	.0040	25.0	300.0	42.0
TM0001	SUNNY S-D.	TM0001	12.33	0.00	23.0	8.0	17.0	50.00	50.00	1.36	350.	300.0	0.75	.0040	25.0	300.0	42.0
TM0004	BHN 66 .	TM0001	12.33	0.00	23.0	8.0	17.0	50.00	50.00	1.36	350.	300.0	0.75	.0040	25.0	300.0	42.0
TM0005	FIT. S-D. TOM.	TM0005	12.33	0.00	23.0	8.0	17.0	50.00	50.00	1.36	350.	300.0	0.75	.0040	25.0	300.0	42.0

TABLE 7. GENETIC COEFFICIENTS FOR CROPGRO-PASTURE (G0GRO980.CUL)

@VAR#	VRNAME.....	ECO#	CSDL	PPSEN	EM-FL	FL-SH	FL-SD	SD-PM	FL-LF	LFMAX	SLAVR	SIZLF	XFRT	WTPSD	SFDUR	SDPDV	PODUR
!			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
UF0001	BAHIA	G00001	12.00	0.200	99.0	10.0	18.0	33.00	25.00	1.00	180.	180.0	0.01	0.180	18.0	2.05	20.0

OTHER CROP MODELS

Several models are being developed for crops that are currently not part of DSSAT v3.5. These include the vegetable crops pepper and cabbage, the grain legumes cowpea, faba bean and velvet bean, the oil crop canola or rape seed, the fruit crops pineapple and banana, the root crops taro and tanier, and the fiber crop cotton. Contact the authors about the status and availability of these models.

CHAPTER FOUR.
MODEL CHANGES AND IMPROVEMENTS

General

The crop simulation models and associated files released as DSSAT v3.0 included the '940' identifier. All models released in DSSAT v3.5 include the '980' identifier to correspond to the release year 1998. Hoogenboom et al. (1994) provide additional information about the file naming conventions of the models and species, ecotype and cultivar files. The CROPGRO module in DSSAT v3.5 is called CRGRO980.EXE; the model driver is called MDRIV980.EXE and the input module is called MINPT980.EXE. All executables are located in the C:\DSSAT35 directory. For soybean the species file is called SBGRO980.SPE, the ecotype file is called SBGRO980.ECO and the cultivar file is called SBGRO980.CUL. Similar to the CROPGRO module, these files are located in the directory C:\DSSAT35-\GENOTYPE. The CERES crop module is called GECER980.EXE and located in the same directory as the CROPGRO module. For maize the species file is called MZCER980.SPE and the cultivar is called MZCER980.CUL. These files are also located in the C:\DSSAT35-\GENOTYPE directory. The naming convention for the other models and crops is identical to the convention shown here for CROPGRO and CERES.

Soil Nitrogen Parameter File

A new input file for all crop models has been implemented for the DSSAT v3.5 release, SOILN980.SOL (Table 8). This file is a model parameter file which externalizes many of the coefficients needed for simulating the decomposition of soil organic matter (1 pool) and organic matter added as residue or manure (3 pools). These coefficients were externalized so that researchers, especially those working in long term sequence simulations, might have easier access to formerly hardwired coefficients in the code. If the soil nitrogen parameter file SOILN980.SOL does not exist in the data directory, it is created upon the first run of the model using the values displayed in Table 8. This file can be edited with any text editor.

The file SOILN980.SOL is self-documented. Note that nine residue codes (RE001 to RE009) are specified in the file. For example, if you are applying manure as an organic amendment, it is likely that the decomposition rates and the composition of the manure is very different from corn stalks. It is a trivial matter to assign the decomposition characteristics of whatever the source of organic matter you are applying to a residue code, which is specified in the *RESIDUE section of the FileX. The default (and formerly hardwired coefficients) is represented in values for RE001. The following coefficients are defined in SOILN980.SOL

DMINR Potential decomposition rate of SOM pool. Default value is .8300E-04 per day.
RTCNR C/N ratio of initial root residue. Default = 40.0.
DSNCV Depth to which soil C (SCDD) and total N (SNDD) values are integrated for output to CARBON.OUT. Default value is 20.0 cm.
RE001 First 3 values are the potential decomposition rates of the carbohydrate, cellulose, and lignin pools; next 3 values are the relative proportions of carbohydrate, cellulose, and lignin in the residue or manure dry matter. Defaults=.2000 .0500 .0095 .2000 .7000 .1000. Up to 9 different residue or manure types can be defined.

TABLE 8. SOIL NITROGEN PARAMETER FILE (SOILN980.SOL)

```
*SOIL NITROGEN PARAMETER FILE
!
! Model parameter file which externalizes many of the
! coefficients needed for simulating the decomposition
! of soil organic matter (1 pool) and organic matter
! added as residue or manure (3 pools). If SOILN980.SOL
! does not exist in the data directory, it is created
! upon the first run of the model. Definitions follow:
!
*CHARACTERISTICS
@C VARIABLE      VALUE
DS DMINR         .8300E-04
DS RTCNR         40.0
DS DSNCV         20.0
DS RE001         .2000   .0500   .0095   .2000   .7000   .1000
DS RE002         .2000   .0500   .0095   .2000   .7000   .1000
DS RE003         .2000   .0500   .0095   .2000   .7000   .1000
DS RE004         .2000   .0500   .0095   .2000   .7000   .1000
DS RE005         .2000   .0500   .0095   .2000   .7000   .1000
DS RE006         .2000   .0500   .0095   .2000   .7000   .1000
DS RE007         .2000   .0500   .0095   .2000   .7000   .1000
DS RE008         .2000   .0500   .0095   .2000   .7000   .1000
DS RE009         .2000   .0500   .0095   .2000   .7000   .1000
```

CALCULATIONS IN GENERIC CERES MODEL

The thermal time calculation in the generic CERES model has been modified since the release of DSSAT v3. The CERES v3.5 models incorporate changes implemented by J.T. Ritchie and L.A. Hunt during a maize modeling workshop held at the Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT) in May, 1998. A series of modifications were implemented to simplify and improve thermal time calculation. These steps include

- Base temperature and optimum temperature are now read in from the species file for each cereal crop. The species file is a new feature of the CERES v3.5 model.

- The base temperature was fixed for each species and not allowed to vary by growth stage
- Crown temperatures during early development are initially set to soil temperature for thermal time calculation for development (DTT). During later phases air temperature instead of soil temperature is used for calculation of thermal time. For winter cereals, such as wheat and barley, the model also checks if snow is present. In this case the model uses air temperature instead of soil temperature.

This new modification improves model performance by more accurately estimating temperature at the growing point of the cereal crops during early development. For users of DSSAT v3.0 and v3.1, this will likely entail recalibration of crop varieties they have added to the DSSAT cultivar database.

Soil Water and Nitrogen Balance

In the previous version of DSSAT v3.0, the soil water and nitrogen balance in the CERES and CROPGRO models were similar but not identical. Since 1994, a concerted effort between the International Fertilizer Development Center (IFDC), University of Florida, and the University of Georgia resulted in the modularization of the CROPGRO and CERES models, as well as the other new models, including CROPSIM, SUBSTOR, and OILCROP. In DSSAT v3.5 the soil water balance and soil nitrogen subroutines are identical. This should result in easier maintenance of all crop simulation models, as well as an improvement of the applications programs, especially crop rotation and sequencing simulations. In the latter case, there is true continuity when crops are changed in a rotation planting scheme.

A detailed description of the soil water balance simulation routines can be found in Ritchie (1998), while the soil nitrogen routines are described by Godwin and Singh (1998). No major modifications were made to the soil water and nitrogen routines in DSSAT v3.5. Soil water evaporation and the soil water content calculation for the top layer were slightly modified to account for both soil evaporation and root water uptake by any roots present in that layer. A new water stress factor was added to the CROPGRO model to account for extreme wet conditions. This water stress factor will be added to the other models in the near future. The models now have two new output soil files, WATBAL.OUT and NBAL.OUT, related to the conditions of the soil profile. The file WATBAL.OUT presents a summary of the soil water status at the start and end of a simulation, while the file NBAL.OUT presents a summary of the soil and plant nitrogen status at the start and end of a simulation. Both files can used as an error checking and debugging tool in case there are errors in the soil water, nitrogen or carbon balance simulation.

Species Files for Non-Leguminous Crops

With the release of DSSAT v3.5, the models other than the CROPGRO family of models now have a species file associated with them. These files are named *crtyp980.SPE*, where *crtyp* represents the crop and model type, i.e., MZCER980.SPE for the CERES-Maize model. An example of the species file for maize is shown in Table 9.

Species files were constructed to externalize some coefficients that were embedded in the code and could only be changed by editing the code and recompiling. Note that the default coefficients for all crops are appropriate for most conditions. Like the new SOILN980.SOL file, the species files are automatically generated by the model if the file does not exist prior to model execution. Default values, that were previously embedded in the code, are used by the model when the file is automatically generated.

The species files are, necessarily, different from crop to crop, although there are common elements. In the maize species file (Table 9), there are both array variables and constants. A common element to all species files are the CO2X and CO2Y arrays. They describe the default response to elevated levels of CO_2. In this case, individuals working on climate change who need to modify a crop species response, can easily accomplish it by editing this array.

Additional variables are also described (and documented) within the species file and should only be changed with care. Two of the more important variables are BASE and TOPT, the base temperature and optimum temperature for growth and development. In the example shown in Table 9 for maize these are 8 and 34° C, respectively. If you believe, for example, the base temperature to be different from 8° C, you can change it here without affecting or modifying the code. In most cases, you should not vary or change these coefficients.

TABLE 9. SPECIES COEFFICIENTS FOR CERES-MAIZE (MZCER980.SPE)

```
*MAIZE SPECIES COEFFICIENTS—GECER980 MODEL

*CHARACTERISTICS
@C X(EN) Y(CH) YSTD
MZ OPT   CO2X     0  220  330  440  550  660  770  880  990 9999 ! CO2 X axis
MZ OPT   CO2Y  0.00 0.81 1.00 1.03 1.06 1.10 1.13 1.16 1.18 1.25 ! CO2 Y axis
MZ OPT   RWEP  1.50
MZ OPT   PORM  0.02                  ! Minimum pore space
MZ OPT   RWMX  0.03                  ! Max root water uptake
MZ OPT   RLWR  0.98                  ! Root length weight ratio
MZ OPT   SDSZ .2750                  ! Seed size (g/seed)
MZ OPT   BASE   8.0                  ! Base temperature
MZ OPT   TOPT  34.0                  ! Temperature optimum
MZ OPT   TYPE     1                  ! Cereal type
```

CHAPTER FIVE.
GRAPHING SIMULATED AND EXPERIMENTAL DATA

SPECIFYING OBSERVED SOIL LAYER DATA FOR WINGRAF

The capability to graph simulated data using the same soil layers as those used for observed data was not documented in the WINGRAF chapter of DSSAT v3 (Chan et al., 1994). The following is a description of the naming convention to be used for headers when setting up a file for measured time series data FILET. The first letter of the header specifies the variable measured. W is used for soil water content, T is soil temperature, N is soil nitrate and A is soil ammonium. The second space is reserved for values that specify the layer number in ascending order from the surface. A maximum of 15 layers can be defined, using the hexadecimal numbering system. The surface layer should be given a value of 1, the ninth layer should be given a value of 9, the tenth layer a value of A, etc. The third and fourth spaces together specify the thickness of the layer in centimeter. Below are several examples for all four variables.

VARIABLE	HEADER EXAMPLE	EXPLANATION
Soil Water (W)	W110	Surface layer, 10 cm thick
Soil Water (W)	WA30	Tenth layer, 30 cm thick
Soil Water (W)	WF25	Fifteenth layer, 25 cm thick
Soil Temperature (T)	T220	Second layer, 20 cm thick
Soil Nitrate (N)	N515	Fifth layer, 15 cm thick
Soil Ammonium (A)	A115	Surface layer, 15 cm thick

When observed data are input using this naming convention for headers, WINGRAF will display the variable name preceded by an infinity symbol (∞). The standard soil layers will also be displayed. To provide an appropriate comparison of simulated and observed data, WINGRAF recalculates the simulated output values provided by the model, which are for the standard soil layers, and graphs the resultant new values for the same soil layer as the observed data.

IMPROVEMENTS

A number of functional changes have been made to WINGRAF. These are outlined in the following section, along with the limitations and caveats.

- The user can save user-selected variables and recall them each time WINGRAF is started.
- New printer drivers have been added to the system. The new QC drivers permit high quality output for many drivers (no more screen dumps). A good choice for lasers is the HPGL mode with A- size paper. The dialogs for setting this up are all

new, and can be modified at the time of printing. The new drivers depend on FREE DOS memory to construct images for printing. Hence, it might be difficult to create and print a hardcopy with network drivers already loaded. The heap is limited to a maximum of 200K, which means there should be enough DOS memory left (assuming that you started with 580K of available DOS memory) to print at a 300 DPI LaserJet resolution, without causing overflow memory problems in the computer.

- The default font is now LITT with a font size of 5. This is acceptable for the screen and for most printer settings. If you print out a hardcopy, and the font is too big or too small, the font size can be interactively changed just before your next print. To use the bitmapped font, type in 'defaultfont' in the dialog box (size = 1). To use another stroked font, just type in the name of the font, assuming it is available, e.g. 'goth', 'scri', etc., in the dialog box, You will need to experiment with font sizes.
- The data arrays are constructed on the heap now, rather than in the data segment. With the heap set at a maximum of 200K, the arrays can handle about 1400 data points, rather than 365 data points on the previous version.
- User files can now be loaded. You can now plot (1) long durations of output, for say tree crops, or (2) crop rotation and sequential outputs. For example, for the water section, multiple years can be plotted by reading data as YRDOY (85123) rather than DOY (123). You may also plot any other OUT file that follows the general OUT file format. If you have a section in DATA.CDE labeled USER, it will pick up header definitions here for plotting.
- If you are plotting summary output, make sure there are no blank lines in the experimental summary file, FILEA, between the header and the data.
- Be sure there is a CR/LF or a blank line at the end of the experimental time series file FILET to have all lines read.
- Be sure there is enough memory in your DOS window session.

The following section shows the new format for the GRAPH.INI file.

```
[Wingraf]
color0=1
color1=15
color2=14
color3=12
color4=15
color5=10
color6=11
color7=13
interval=4
tics=5
days=Y
symbols=C
```

```
thickness=T
plot=P
exp=A
screenmode=3
sound=1
driver=28
port=0
orientation=1
file=OUTPUT00.GRA
fontsize=5
font=litt
[WMgraf]
color0=1
color1=15
color2=14
color3=12
color4=15
color5=10
color6=11
color7=13
interval=-9
tics=-9
driver=28
port=0
orientation=0
fontsize=5
usetics=Y
usekeyword=1
useabv=1
font=litt
file=OUTPUT
```

Application Graphics Features

The application graphics program WMGraf has several features that have not been documented. WMGraf is used in the application programs for seasonal, sequence and spatial analysis, as well as in WeatherMan. After plotting a graph with WMGraf, a number of options are available. Pressing the letter 'P' or 'p' will send the graph to the current printer and output destination. Pressing the letter 'G' or 'g' toggles the graph gridlines on and off. The default value is to construct a graph with grid lines. Striking any other key terminates WMGraf and returns the user to the calling program.

Additional options are available to users with a mouse. An arrow shaped cursor will appear in the upper left corner of the plot. Clicking on the left mouse button will display the coordinates of any point within the graph, with the X-axis value in the upper left corner of the screen and the Y-axis value in the upper right corner. This way, the user may click on any point and ascertain the X- and Y-data values for this point. In addition, clicking on the right mouse button will bring up a menu on the

right side of the displayed graph. You may choose to toggle the grid, print the graph, or exit WMGraf by clicking on the appropriate menu item with the left mouse button.

A new feature has been added to WMGraf that allows you to load any font (*.chr) file from disk. The default font, e.g. LITT, is still included with WMGraf. However, there is inadequate data space to bind all the fonts into WMGraf. You, therefore, have to load them from disk. From the section on WMGRaf of GRAPH.INI, add or modify the line font=XXXX, where XXXX is the name of the font such as TRIP for triplex, EURO for European, etc. You can also experiment with different fonts by setting the size at 1, rather than the pre-set value of 6. You can also use the DEFAULTFONT, an 8*8 bitmap font available from the ROM-BIOS of all computers.

CHAPTER SIX.
TROUBLE SHOOTING AND SUGGESTIONS

Converting Data Sets From DSSAT v2.1 to DSSAT v3.0

Please compare your new converted data set(s) with your old DSSAT v2.1 data set(s) before you run the models. The convert program does not convert automatic irrigation and related options. If you created new cultivars and cultivar parameters in your DSSAT v2.1 genetics file, you need to add these cultivars to the new DSSAT v3.0 genetics file first, before you run the convert program. For some crops several of the cultivar specific parameters have been changed.

Converting Data Sets From DSSAT v3.0 to DSSAT v3.5

DSSAT v3.5 does not include a utility to convert files from DSSAT v3.0 to DSSAT v3.5 format. The crop models and all applications program, however, are able to read both formats. You can import your old file using the experimental details file create utility XCREATE and save your file. XCREATE will allow you to add the values for the new input parameters, such as latitude and longitude, and save it in the new format. XCREATE, however, will not move your initial organic residue from the *RESIDUES AND OTHER ORGANIC MATERIALS section to the *INITIAL CONDITIONS SECTION.

File Managers

If you installed DSSAT v3.5 on a drive and directory different from C:\DSSAT35, please manually edit the file DSSATLST.FLE to point to the correct data directories. Additional data directories can also be added to the file DSSATLST.FLE.

If you have problems with “Out of memory” error messages when using the LIST/EDIT functions in DSSAT v3.5, the cause is a limitation in the TVED.EXE editor program that is included and installed as the default editor of DSSAT v3.5. This editor can only edit files smaller than 64K in size. Editing files bigger than this will result the “Out of memory” error. To solve this problem, go to DSSAT v3.5 SETUP menu and under TOOLS and change to another editor of your own choosing. One common editor is EDIT.COM, which is distributed with the MS-DOS, Windows 95 and Windows 98 operating systems.

System Errors

DSSAT v3.5 is a DOS-based system. Many of the potential problems may be associ-

ated with the operation of your DOS window or your Windows 95 or Windows 98 operating environment. Below is a simple checklist that you can use in case you have problems with one or more components of DSSAT v3.5.

1. Have you checked the README.V35 document, included with your copy of the software, for a possible solution to your DSSAT v3.5 problem?

2. Does your system fulfill the minimum system requirements for running DSSAT v3.5?

Processor:	80486 or better
RAM:	640K (minimum **590K** free DOS RAM required)
Display:	VGA color or better
Hard Disk:	Required, approximately 25 MB for a complete installation of DSSAT v3.5
Operating System:	DOS 3.3 or higher

3. After DSSAT v3.5 has been installed, check to see if there is enough free hard disk space to run DSSAT v3.5. DSSAT v3.5 requires approximately 1 MB of free hard disk space to write temporary and other files while it is running.

4. Did you add the statement C:\DSSAT35 or equivalent and SET DSSAT3=C:\DSSAT35 or equivalent to your AUTOEXEC.BAT? At the prompt, type SET to check what the current setting are.

5. Check your input data. Especially check for valid format and valid values, and make sure that there are no strange ASCII characters in the input files. Please make sure that there are also no hidden "TAB" characters in your file.

6. Use a program called CHECKIT or something similar to do a system diagnostics and test your computer.

7. For MS-DOS v6 or higher users:
   If you are having problems running any of the DSSAT v3.5 components, check to see if you are using EMM386.EXE in your CONFIG.SYS. If so, you may need to add the NOEMS and NOVCPI options to get DSSAT v3.5 to run. For example,

   DEVICE=C:\DOS\EMM386.EXE NOEMS NOVCPI

NOVCPI and NOEMS are BOTH required, since NOVCPI without NOEMS does not disable VCPI support by EMM386.

8. For Pentium II computers:
   The analysis programs under Seasonal and Sequence use a utility called AUTO.EXE. This utility is used to spool part of the programs to the hard disk and to save memory. It has been reported that this can cause a system error or occasional "hanging" on Pentium II computers, especially those that are faster than 200 Mhz. When this occurs, check your COMSPEC statement by typing SET in a MS-DOS window. Add the following statement to AUTOEXEC.BAT

   SET COMSPEC=C:\COMMAND.COM

   if the SET statement shows something different. Reboot your computer to make this statement active. This will solve most of the analysis problems, but not all.

9. Memory problems
   Certain programs within DSSAT v3.5 will run with less than 590 Kbyte memory. However, both the seasonal, sequence and spatial analysis programs require 590 KByte memory, especially for displaying results graphically. If these programs do not seem to work, please try to free up more memory by deleting some TSR programs or network drivers from either your AUTOEXEC.BAT or CONFIG.SYS files. Do not forget to reboot your computer after you make changes to these files.

10. Other problems
    When you have completed all of the above and still have not been able to determine the cause of your DSSAT v3.5 problems, contact us at "ibsnat@hawaii.edu" or by fax at 808-956-3421.

    Please be sure to include a clear description of your problem, including MS-DOS or Windows 95 and 95 version, memory status, contents of your CONFIG.SYS and contents of your AUTOEXEC.BAT, and other specifics that might help us solve your problem.

REFERENCES

Boote, K. J., J. W. Jones, G. Hoogenboom, and N. B. Pickering. 1998. The CROPGRO model for grain legumes. p. 99-128. In: [G. Y. Tsuji, G. Hoogenboom, and P. K. Thornton, editors]. *Understanding Options for Agricultural Production.* Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, the Netherlands (ISBN 0-7923-4833-8).

Chan, H., P. W. Wilkens, G. Hoogenboom, J. W. Jones, and D. C. Godwin. 1994. Graphing simulated and experiment data. p. 245-281. *In:* [G. Y. Tsuji, G. Uehara, and S. Balas, editors] *DSSAT version 3, Volume 2.* University of Hawaii, Honolulu, HI (ISBN 1-886684-02-0).

Godwin, D.C. and U. Singh. 1998. Nitrogen balance and crop response to nitrogen in upland and lowland cropping systems. p. 55-77. In: [G. Y. Tsuji, G. Hoogenboom, and P. K. Thornton, editors]. *Understanding Options for Agricultural Production.* Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, the Netherlands (ISBN 0-7923-4833-8).

Hoogenboom, G., J. W. Jones, P. W. Wilkens, W. D. Batchelor, W. T. Bowen, L. A. Hunt, N. B. Pickering, U. Singh, D. C. Godwin, B. Baer, K. J. Boote, J. T. Ritchie, and J. W. White. 1994. Crop models. p. 95-244. *In:* [G. Y. Tsuji, G. Uehara, and S. Balas, editors] *DSSAT version 3, Volume 2.* University of Hawaii, Honolulu, HI (ISBN 1-886684-02-0).

Jones, J. W., L. A. Hunt, G. Hoogenboom, D. C. Godwin, U. Singh, G. Y. Tsuji, N. B. Pickering, P. K. Thornton, W. T. Bowen, K. J. Boote, and J. T. Ritchie. 1994. Input and output files. p. 1-93. *In:* [G. Y. Tsuji, G. Uehara, and S. Balas, editors] *DSSAT version 3, Volume 2.* University of Hawaii, Honolulu, HI (ISBN 1-886684-02-0).

Ritchie, J.T. 1998. Soil water balance and plant water stress. p. 41-54. *In:* [G. Y. Tsuji, G. Hoogenboom, and P. K. Thornton, editors]. *Understanding Options for Agricultural Production.* Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, the Netherlands (ISBN 0-7923-4833-8).

Ritchie, J.T., U. Singh, D.C. Godwin, and W.T. Bowen. 1998. Cereal growth, development and yield. p. 79-98. *In:* [G. Y. Tsuji, G. Hoogenboom, and P. K. Thornton, editors]. *Understanding Options for Agricultural Production.* Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, the Netherlands (ISBN 0-7923-4833-8).

Scholberg, J.M.S., K.J. Boote, J.W. Jones, and B.L. McNeal. 1997. Adaptation of the CROPGRO model to simulate field-grown tomato. p. 135-151. *In:* [M.J. Kropff, P.S. Teng, P.K. Aggarwal, J. Bouma, B.A.M. Bouman, J.W. Jones, and H.H. van Laar, editors]. *Application of Systems Approaches at the Field Level.* Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, the Netherlands (ISBN 0-7923-4286-0).

Scholberg, J.M.S., B.L. McNeal, G. Hoogenboom, J.W. Jones, K.J. Boote, and S.R. Olson. 1996. Application of the CROPGRO model to predict fruit yield of fresh-market tomato. *Agronomy Abstracts* 14.

Singh, P., and S.M. Virmani. 1996. Modeling growth and yield of chickpea (Cicer arietinum L.). *Field Crops Research* 46:41-59.

Singh, U., R.B. Matthews, T.S. Griffin, J.T. Ritchie, L.A. Hunt, and R. Goenage. 1998. Modeling growth and development of root and tuber crops. p. 129-156. *In:* [G.Y. Tsuji, G. Hoogenboom, and P. K. Thornton, editors]. *Understanding Options for Agricultural Production.* Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, the Netherlands (ISBN 0-7923-4833-8).

Villalobos, F.J., A.J. Hall, and J.T. Ritchie. 1992. OILCROP-SUN: A crop growth and development model of the sunflower (*Helianthus annuus L.*) crop. *Proceedings 13th International Sunflower Conference,* International Sunflower Association: Pisa, Italy. pp.453-458.

Villalobos, F.J., A.J. Hall, J.T. Ritchie, and F. Orgaz. 1996. OILCROP-SUN: A development, growth and yield model of the sunflower crop. *Agronomy Journal* 88(3):403-415.

Villalobos, F.J., V.O. Sadras, A. Soriano, and E. Fereres. 1992. Effects of plant population and genotype on sunflower dry matter partitioning and yield. *In* (E. Scaife, editor). *Proceedings 2nd. Congress European Society of Agronomy.* European Society of Agronomy, Wellesbourne. pp. 144-145.

Volume 4-2

Spatial Analysis

P.K. Thornton
H.W.G. Booltink
J.J. Stoorvogel

International Livestock Research Institute,
Wageningen Agricultural University
International Consortium for Agricultural Systems Applications

ACKNOWLEDGMENTS

The development of this software was partially supported through a grant from the School of Production Ecology of the Wageningen Agricultural University, Wageningen, The Netherlands.

CHAPTER ONE. INTRODUCTION

The most recent version of the DSSAT, version 3.5, allows the user to perform spatial simulation experiments and to analyze the results. Normally, the crop models are run on a "point" basis, i.e., input data in terms of soil and weather conditions are assumed to relate to one location, e.g., a field or plot where an experiment was carried out. With the increasing power and speed of personal computers, and the increasing availability of powerful Geographic Information System (GIS) software, a number of research groups have used the DSSAT crop models to investigate various spatial applications (see, for instance, Singh *et al.*, 1993; Negabhan *et al.*, 1994; Hoogenboom and Gresham, 1995; Engel *et al.*, 1997; Booltink and Verhagen, 1997; Thornton *et al.*, 1997b). With the advent of International Consortium for Agricultural Systems Applications (ICASA) as an umbrella organization for expanded collaboration between the former IBSNAT group and the modeling groups in the Netherlands and elsewhere, it was decided in early 1996 to incorporate some rudimentary spatial analysis capabilities into future revisions of the DSSAT software.

The programs developed are open-ended, in that it is straightforward to import or export relevant files to a commercial GIS package. It is expected that most users will have their own copies of a GIS, such as IDRISI (Eastman, 1994) or Arc/Info (ESRI, 1992), or some other type of spatially-oriented program such as Surfer (Surfer, 1989).

As in seasonal and sequence analysis in the DSSAT, there are three basic steps involved in spatial analysis:

1. The creation of an appropriate model input file;

2. Running the crop model(s) using a special controller program or driver (the crop models themselves need no modification);

3. Analyzing the results of the simulation using one of two basic tools: a geostatistics program that allows the user to create a "surface" of model response for an area, based on a finite number of sample points in the area (described in Chapter 5 of this document), and a separate program that allows model results to be mapped (described in Chapter 4 of this document). Both programs use a shareware mapping utility developed by one of us (JJS); this module is described in Chapter 6 of this document. Model results can also be investigated using the seasonal analysis program in the DSSAT (documented in Tsuji *et al.*, 1994, Volume 3, Section 1).

It should be noted that the software is fully consistent with AEGIS/WIN, and users can run the models within that system using the same spatial driver, MAPDRV.EXE, that is described below. AEGIS/WIN is different in that it calls the DSSAT v3.5 from a Windows interface; the programs outlined here are called from within the DSSAT v3.5 shell. Otherwise, the various spatial analysis systems are compatible. The highlights of the spatial analysis program are documented in Thornton *et al.* (1997a). AEGIS/WIN is documented in the same journal by Engel *et al.* (1997) and in section 3 of this volume (Engel *et al.*, 1999).

The links between the three steps are shown in Figure 1. Spatial analyses are run from the DSSAT shell under the menu "Analyses—Spatial". The three steps are explained in more detail below.

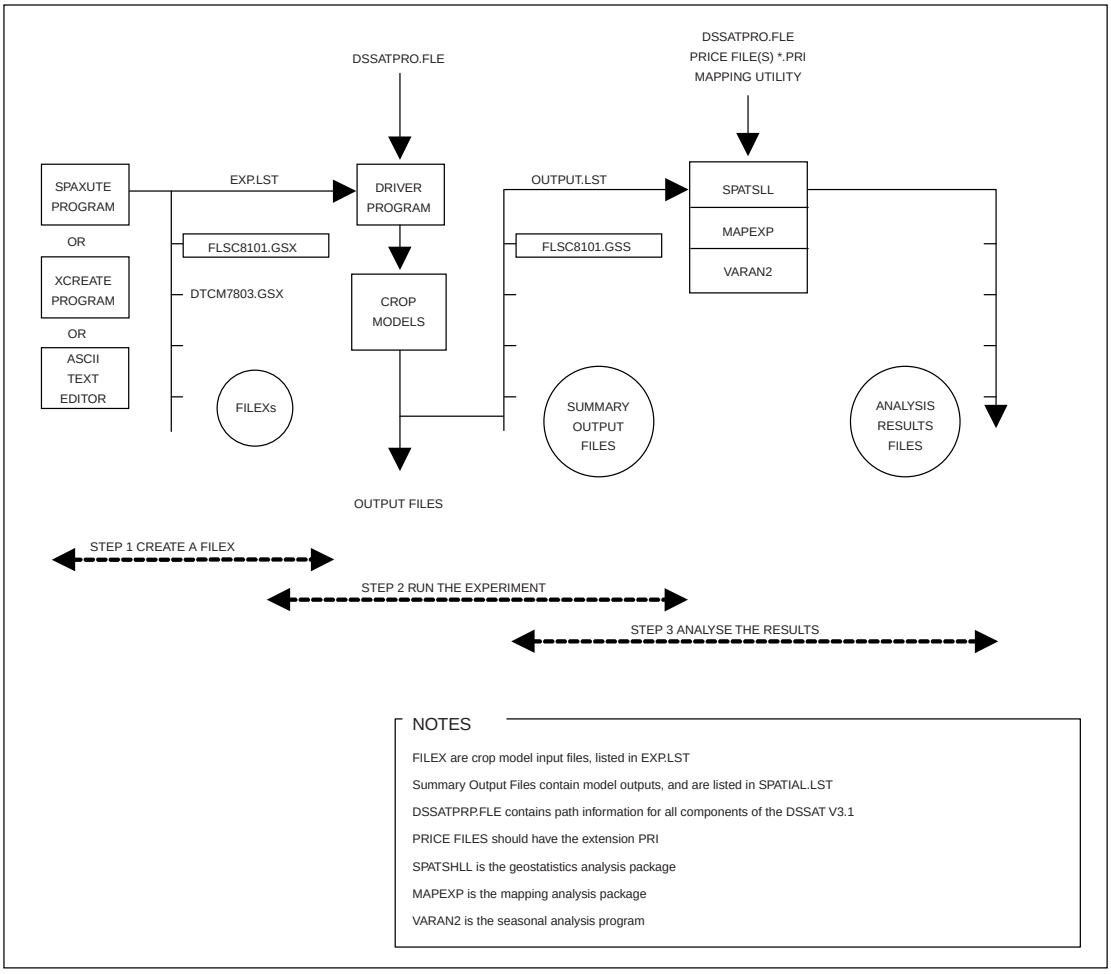


Figure 1. Steps in Spatial Analysis

CHAPTER TWO.
CREATING A SPATIAL FILEX

General Considerations

The basic model input file, FILEX, can be adapted to represent spatial experiments that can then be run using the DSSAT crop models. The essential elements of a spatial FILEX are the same as for FILEXs constructed for model validation, sensitivity analysis, and seasonal and sequence analyses, except that they may be much larger. Consider an area made up of four fields of slightly different soil type; a FILEX to represent this situation could be built up be specifying four treatments in the FILEX, each differing in the "field" factor level, FL, that specifies a different field or location. Part of such a FILEX is shown below:

```
*TREATMENTS                        -------------FACTOR LEVELS------------
@N R O C TNAME..................... CU FL SA IC MP MI MF MR MC MT ME MH SM
 1 1 0 0 Field 1                    1  1  0  1  1  0  1  1  0  0  0  0  1
 2 1 0 0 Field 2                    1  2  0  1  1  0  1  1  0  0  0  0  1
 3 1 0 0 Field 3                    1  3  0  1  1  0  1  1  0  0  0  0  1
 4 1 0 0 Field 4                    1  4  0  1  1  0  1  1  0  0  0  0  1

*FIELDS
@L ID_FIELD WSTA....  FLSA FLOB  FLDT  FLDD  FLDS  FLST SLTX SLDP    ID_SOIL
 1 FLSC0001 FLSC8101   -99    0 DR000     0     0 00000  -99  151 IB00000002
 2 FLSC0002 FLSC8101   -99    0 DR000     0     0 00000  -99  151 IB00000021
 3 FLSC0003 FLSC8101   -99    0 DR000     0     0 00000  -99  151 IB00000031
 4 FLSC0004 FLSC8101   -99    0 DR000     0     0 00000  -99  151 IB00000004
```

This example shows a different soil type for each field, and this sequence of "treatments" (in reality, a set of self-contained management scenarios imposed on a number of fields) can then be run using the appropriate crop models and the results analyzed. For spatial analyses, a number of changes to the FILEX have been made. One of these is to introduce a second tier to the *FIELDS section, consisting of x- and y-coordinates, elevation, polygon area, and other such information (the contents of the new *FIELDS section are shown in Table 1).

It is very important to note that there are no provisions in the DSSAT to take account of interactions between the various spatial elements as specified in the FILEX. As stated above, the list of treatments in the file contains completely self-contained scenarios, and the fact that they may represent contiguous polygons in a coverage or map of an area is, as far as the models are concerned, irrelevant. For this reason, the "spatial" capabilities of the DSSAT are, in a sense, very limited. This does not mean, however, that sophisticated spatial (in the true sense of the word) analyses cannot be done. For example, Arc/Info can be used

to create a map of runoff and infiltration from a digital elevation model; this information can be turned into soil profiles of differing physical characteristics, and then the models can be run using the DSSAT for these different profiles and the results mapped.

TABLE 1. VARIABLES IN THE REVISED *FIELDS SECTION OF FILEX

	Variable[1]	Header[2]	Format[3]
*FIELDS			
Tier 1			
Level number	L	L	0 I 3
Field identifier	FLDNAM	ID_FIELD	0 C 8
Weather station identifier	WSTA	WSTA	1 C 8
Slope, redefined from 1 C 5	FLSA	FLSA	1 R 5 2
Degrees to obstruction	FLOB	FLOB	1 R 5 0
Drainage code	FLDT	FLDT	1 C 5
Drainage depth	FLDD	FLDD	1 R 5 0
Drain separation	FLDS	FLDS	1 R 5 0
Surface stoniness, %	FLST	FLST	1 C 5
Soil texture code	SLTX	SLTX	1 C 5
Soil depth	SLDP	SLDP	1 R 5 0
Soil identifier	SOILN	ID_SOIL	1 C 10
Tier 2			
Level number	L	L	0 I 3
X-coordinate, any reference system	XCRD	XCRD	0 R 15 5
Y-coordinate, any reference system	YCRD	YCRD	1 R 15 5
Elevation, m above sea level	ELEV	ELEV	1 R 9 2
Area of the polygon, m_2 (0 is point/raster)	AREA	AREA	1 R 17 1
Length of slope, m	SLEN	SLEN	1 R 5 0
Polygon length-width ratio	FLWR	FLWR	1 R 5 1
Slope aspect, degrees clockwise from north	SLAS	SLAS	1 R 5 1

[1] Abbreviations used as variable names in the DSSAT models.
[2] Abbreviations suggested for use in header lines (those designated with "@") within the file.
[3] Formats are written as follows: number of leading spaces, variable type (C=Character, R=Real, I=Integer), variable width, and (if real) number of decimal places.

The creation of a FILEX for use in spatial analysis will usually make use of the program XCREATE or a text editor and a FILEX template. A utility exists in the DSSAT v3.5, however, to help create parts of a FILEX, as spatial FILEXs may be very long. There are two parts to this process:

- Creating a values file that contains the attributes for each distinct polygon under study;
- Taking this values file and creating all or part of a FILEX.

Users with their own GIS software may create the values file for step 2 in any way they like. The utility in the DSSAT for doing this assumes the user has images in IDRISI format. The process is described below.

Step 1

Assume we have a soils map in the form of an IDRISI image. It consists simply of a rectangular grid with numbers in each grid box:

```
Soil          1   2   2
              1   1   1
              1   1   1
```

and a treatment coverage for the same raster points,

```
Treatment     1   1   1
              1   1   2
              1   1   2
```

What these numbers actually refer to may be specified in a values file, which is a file that relates the numbers in the image to what they represent. Thus the soils values file might consist of the following:

```
0000
0001 IB00000001
0002 IB00000002
```

The identifier on the left is the number that appears in the coverage, the code on the right is the DSSAT soil identifier. For the treatment coverage, we might assume that the numbers refer to the factor level for fertilizer, i.e., rasters with a value of 1 will receive 20 kg N ha^{-1}, those with 2 will receive 40 kg N ha^{-1}. A spatial experiment may be made up of many of these layers or coverages, each relating to one particular variable or factor.

The next step is to overlay these layers to produce what may be termed an "experiment" coverage, where all the unique combinations of the experiment are defined. This step can be done using the utility SPAXUTE in the DSSAT (note that individual coverages are assumed to exist already). In the example above, if we overlay the soils and treatment layers, we get an experiment coverage like this:

```
Experiment     11   12   12
               11   11   21
               11   11   21
```

where the first number is simply the treatment number, and the second the soil number. This gives three unique combinations, 11, 12, and 21. These are put into a "values" file with the appropriate DSSAT headers:

```
@PID  MF    ID_SOIL
0001   1 IB00000001
0002   1 IB00000002
0003   2 IB00000001
```

The header PID is shorthand for the Polygon Identifier (the number that is actually stored in the experiment coverage, rather than the values "11" or "12", for instance), MF is the fertilizer factor level, and ID_SOIL is the soil identifier.

This is the end of step 1. At this stage, we have an experiment coverage that contains a unique number for each unique combination of the factors that we are interested in (in the example, MF and ID_SOIL), and an associated values file that contains details of the combinations.

As noted above, there are a number of commercial GISs that can do this manipulation simply. Because IDRISI for DOS and for Windows is a relatively inexpensive and readily available package, routines have been written to allow this manipulation of IDRISI images to be carried out from the DSSAT directly.

Step 2

The second step is to take the experiment coverage and values file and make a FILEX that the models can use. From the DSSAT, the program SPAXUTE can be used to create a "skeleton" FILEX that contains the *TREATMENTS and *FIELDS sections as specified in the values file. Note that if you choose to update an existing FILEX with new treatment and field sections, then all the other sections in the existing FILEX will simply be copied out to the updated file. If the FILEX does not exist, then the program will insert default sections corresponding to the levels found in the *TREATMENTS section created (see below for default factor levels and thus default section creation in a new FILEX). In either case, you will need to edit the resultant FILEX, as there will not necessarily be correspondence between all the factor levels specified and appropriate blocks in the file. There will be a correspondence for the *FIELDS section, however.

To illustrate, assume you have a treatment coverage with 20 unique treatments. A FILEX will be created with 20 treatments with, among other things, 20 field levels, and the *FIELDS section of the file will contain 20 entries. However, you may have specified four cultivar levels. The resultant FILEX will contain a *CULTIVARS section, as this is the default if no cultivar level header (CU) is found in the treatment values file, but it will contain only one level (level number 1). The other three cultivar levels will need filling in before the FILEX can be run with the driver.

Using the simple example above, the program SPAXUTE would produce the following in the FILEX:

```
  :
  :
*TREATMENTS                        -------------FACTOR LEVELS------------
@N R O C TNAME..................... CU FL SA IC MP MI MF MR MC MT ME MH SM
 1 1 0 0                            1  1  0  1  1  0  1  0  0  0  0  0  1
 2 1 0 0                            1  2  0  1  1  0  1  0  0  0  0  0  1
 3 1 0 0                            1  3  0  1  1  0  2  0  0  0  0  0  1
  :
  :
  :
  :
*FIELDS
@L ID_FIELD WSTA....  FLSA  FLOB  FLDT  FLDD  FLDS  FLST SLTX SLDP ID_SOIL
 1 FLSC0001 FLSC8101 -99       0 DR000     0     0 00000 -99   151 IB00000001
 2 FLSC0002 FLSC8101 -99       0 DR000     0     0 00000 -99   151 IB00000002
 3 FLSC0003 FLSC8101 -99       0 DR000     0     0 00000 -99   151 IB00000001
@L            XCRD            YCRD     ELEV           AREA  SLEN  FLWR  SLAS
 1             -99             -99      -99          0.0000     0   1.0  90.0
 2             -99             -99      -99          0.0000     0   1.0  90.0
 3             -99             -99      -99          0.0000     0   1.0  90.0
  :
  :
```

There are three field levels rather than two, since there is a unique field identifier (ID_FIELD) for each of the three combinations of soil and fertilizer level. To get this FILEX to run with the models, it would be necessary simply to edit the FILEX produced to ensure that there were appropriate entries in the cultivar section, the initial conditions section, the planting section, the fertilizer section with two levels, and the simulation controls block.

The SPAXUTE Utility

If you have IDRISI images to combine to form an experiment coverage and values file, then you can use SPAXUTE to do this for you. The first menu option in this program is to overlay IDRISI coverages (Menu 1, Table 2). This operates on the basis of a list. You add the details you want to the list, then execute the list (option 1.5). For each layer or coverage, an IDRISI image file must be selected (with an IMG extension, option 1.1.1). The file type may be ASCII or binary, and the data type may be byte, integer, or real. Secondly (option 1.1.2), a values file associated with the coverage just chosen may be specified. If no values file is associated, then

the values in the image itself will be used for subsequent processing. Thirdly, a header is specified (option 1.1.3). The header chosen may be one of the following:

- Any factor level code from the *TREATMENTS section of the FILEX (e.g. CU for cultivar, MF for fertilizer, etc);
- Any of the variables from the *FIELDS section (Table 1), such as soil ID or slope;
- The MASK header; if this is chosen, then the coverage is treated as a mask for the resultant overlay. Values in the mask image of 0 will be ignored or skipped over; non-zero values will be processed normally. This option is useful if you specify a land-use coverage, for example (see below).

TABLE 2. SPAXUTE MENU STRUCTURE

MAIN MENU	SECONDARY MENU	TERTIARY MENU
1 Overlay IDRISI Coverages	1.1 Add an entry to list	1.1.1 Select mage file
		1.1.2 Select values file
		1.1.3 Select a header
	1.2 Delete entry from list	1.2.1 Entry number to delete:
	1.3 Coverage < >[1]	1.3.1 Enter coverage file name:
	1.4 Val file < >	1.4.1 Enter values file name:
	1.5 Execute list (hot key)	
	1.6 Quit	
2 Convert values file to a FILEX	2.1 Create FILEX (hot key)	
	2.2 Select a FILEX name	2.2.1 Enter 8 character name:
	2.3 Select a coverage < >	2.3.1 Select coverage name:
	2.4 Select values file < >	2.4.1 Select values file name:
	2.5 INCLUDE option toggle[2]	
3 Utilities	3.1 Edit a file	3.1.1 Select file name:
	3.2 Rename a file	3.2.1 Select file name:
		3.2.2 Rename to:
	3.3 Copy a file	3.3.1 Select file name:
		3.3.2 Copy to:
	3.4 Call up the mapping module	
	3.5 Exit	
4 About SPAXUTE		
5 Exit		

[1] The < > indicates that the appropriate file name is echoed to the menu screen (either a default name or the name selected by the user).
*[2] A toggle between producing a skeleton FILEX (the default) and writing out *TREATMENTS and *FIELDS sections as separate files (with extensions TR and FL, respectively) that can be INCLUDED in another FILEX.*

Once these three items have been chosen for an entry, the list on the screen reflects the options chosen. Another overlay may be added (up to 10 in total). If a mistake is made, an overlay can be deleted from the list (option 1.2).

Before processing the list, the resultant experiment coverage can be given a name (the default is TREAT.IMG), option 1.3, and the resultant values file can be named (the default is TREAT.VAL), option 1.4. Once all is ready, the "Execute list" option can be pressed (1.5) to process the list.

A number of checks are made:

- All IMG files must exist, together with their associated DOC files.
- All specified values files must exist.
- All IMG image files selected must have the same number of rows and columns.

Note that you may specify the same values file for more than one coverage entry in the list, and you may also specify the same image file for more than one entry, if you want. Coverages may be combined in any order; the fields created in the values file will have variable widths, depending on the contents of the field, and will be right-justified with the header in the field.

If processing is successful, an experiment coverage or image will be created (the file type will be ASCII and the data type INTEGER), containing unique identifiers for each combination of factors and variables as specified in the list, with an associated values file containing the attributes of each combination.

The next menu item (2, main menu, Table 2) will take the experiment values file created (using option 1 or using some other method) and produce appropriate *TREATMENT and *FIELDS sections for a FILEX. To use this option, first select a FILEX name (option 2.2); it is best to use the standard DSSAT v3 convention of an institute ID, site ID, year, and field ID, e.g., UFGA8101. Do not enter an extension. The FILEX may exist or it may not; as noted above, if it does exist, then only the fields and treatments sections will be modified. If it does not exist, a skeleton FILEX will be written out. In either case, the FILEX itself, UFGA8101.GSX (note the GS extension, for "spatial"), will need to be edited in a separate operation, to ensure that the file is complete and that all relevant variables have been initialized.

Before creating or modifying the FILEX, the names of the experiment coverage and values file must be specified (again, by default, these are named TREAT.IMG and TREAT.VAL). You can use options 2.3 and 2.4 to change these. These three slots must have entries; attempting to create the FILEX sections with any of these blank will result in a warning message.

This part of the program will process coverages in the following formats:

- IDRISI for DOS and IDRISI for WINDOWS image files with a file type of binary or ASCII and a data type of byte, integer, or real (the associated DOC file must exist);
- Surfer GRD grid files in ASCII format;
- Arc/Info ARC grid files in ASCII format.

Note that most values in the coverages are dealt with as integers; the exceptions are coverages from which x- and y-coordinates and elevations are to be extracted. The program operates as follows: if headers exist in the values file, then these values will be used in the appropriate place in the *FIELDS and *TREATMENTS sections of the FILEX. If there is no header for a particular variable, and thus no values, then defaults are used. The default values for the factor levels are as follows:

CU	FL	SA	IC	MP	MI	MF	MR	MC	MT	ME	MH	SM
1	-	0	1	1	0	0	0	0	0	0	0	1

There is no default value for the fields factor level; normally this is set to the treatment number, unless a FL header exists in the values files. Thus cultivar, initial conditions, planting details, and simulation control block levels are set to 1, the others to 0. For the values in the *FIELDS section, default values are as follows:

Variable	Default Value
ID_FIELD	institute ID + site ID + field ID
WSTA	institute ID + site ID
FLSA	-99
FLOB	0
FLDT	DR000
FLDD	0
FLDS	0
FLST	00000
SLTX	-99
SLDP	-99
ID_SOIL	IB00000000
XCRD	-99
YCRD	-99
ELEV	-99
AREA	-99
SLEN	0
FLWR	1.0
SLAS	90.0

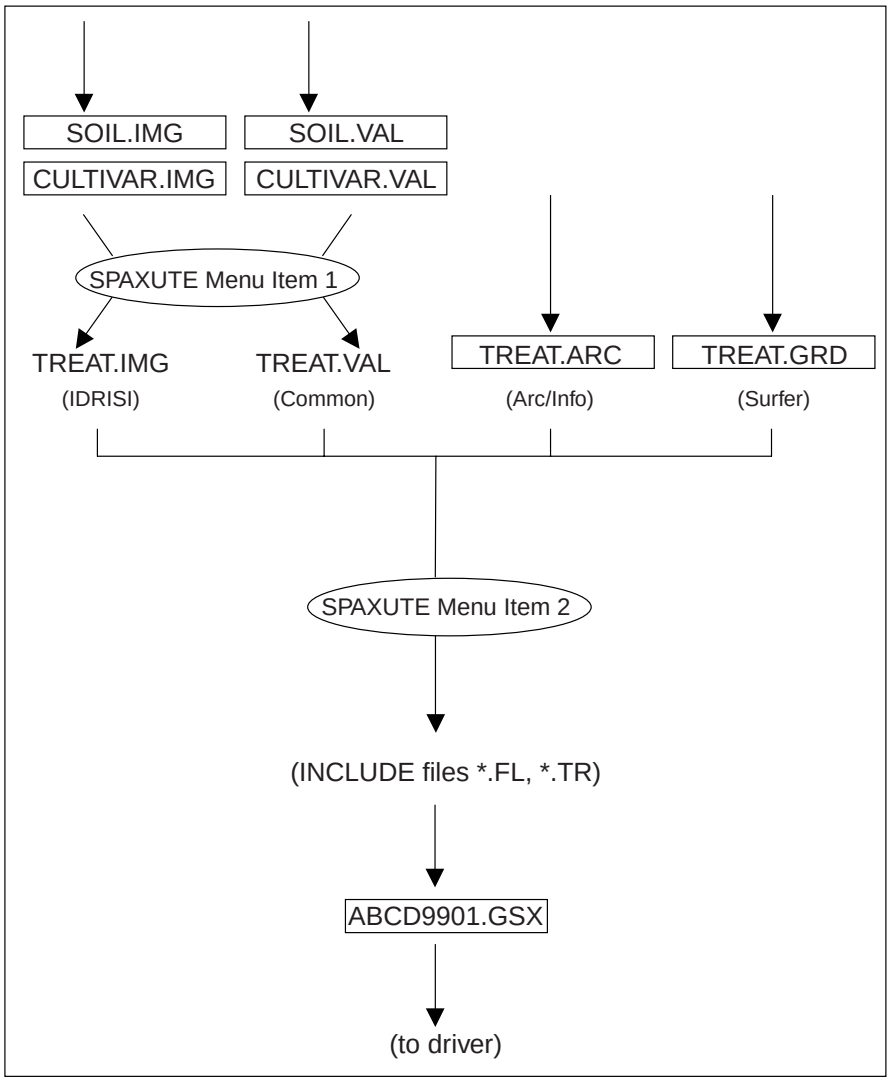


Figure 2. Spatial FILEX Creation. Coverage Files are Required Inputs.

Some of these defaults will need to be edited before running the experiment, such as the soil ID, for example. For all of these variables, if the appropriate header exists in the values file, then those values will be used instead. Default values are used only when there is no information available concerning the variable or factor level.

After processing is complete, a FILEX will have been created containing the *FIELDS and *TREATMENTS sections. Note that the program operates by reading the experiment image itself and then matching the value in each grid cell with an entry in the values file. It may be that there are more entries in the values file than there are in the image; this

causes no problems. It may alternatively be that there are values in the image for which there are no corresponding entries in the values file. A warning message is issued about these "orphan" values in the image file when found, and of course their corresponding entries in the FILEX will have default values for all variables.

The scheme of files in SPAXUTE is shown in Figure 2, with sample file names. The files shown in boxes must exist before SPAXUTE is run. Of course, TREAT.IMG may already exist (you might create this using other software); but while part of the point of SPAXUTE is to make this easier to do from within the DSSAT, the experiment image and values files do not necessarily have to be created using SPAXUTE.

There are provisions in SPAXUTE to edit, copy and rename files (menu options 3.1, 3.2, and 3.3), and to enter the mapping module (3.4) and plot IDRISI images, Surfer grid files or Arc/Info ARC files. The editor accessed will be the one defined by the user in the setup file DSSATPRO.FLE, accessible through the SETUP menu in the DSSAT v3.5 shell. The mapping module that is accessible from the shell can also be defined by the user, but this will normally be the default program SGIS9801.EXE.

There are some limits to the program. A maximum of 10 coverages may be combined to form the experiment coverage and values files. The maximum size of the IDRISI images, Surfer grid files, and Arc/Info ARC files currently supported is 900 rows by 900 columns. FILEXs in DSSAT v3.5 can now support treatment numbers and factor levels up to 999, i.e., 999 unique polygon IDs can be simulated in one file.

Another important limit to the program is the number of items that can be handled by the scrolling routine when selecting a file name. The limit is 60; if there are more than 60 files in the directory (if the search is for all files, *.*) or more than 60 matching files (if the search is for *.VAL, for example), then only the first 60 will be picked up by the program. Users may occasionally find this annoying (it may also explain why you cannot apparently find a file that you know exists), but it does provide a very good incentive to keep your spatial data directories as uncluttered as possible.

For SPAXUTE, as for the other parts of the spatial analysis software, there are fairly rigid conventions as regards file extensions. These are listed below. Some of these relate only to the mapping of grid files using the utility option in the program.

Users should note that one of the headers that can be applied to a treatment values file relates to economics. The header, available from Menu Item 1.1.3 (Table 2), allows price and cost files to be specified spatially. The effect of this header is similar to a true factor level in the treatments section of the FILEX. This is discussed further in Section 4 below.

```
Extension      File Type
```

IDRISI Files

```
IMG      IDRISI for DOS/WINDOWS image file, any file type and data type
DOC      IDRISI for DOS/WINDOWS document file for an image file
VEC      IDRISI for DOS/WINDOWS vector file (ASCII file type)
DVC      IDRISI for DOS/WINDOWS vector documentation file
```

Surfer Files

```
GRD      Surfer grid file (only ASCII file type supported)
```

Arc/Info Files

```
ARC      Arc/Info raster file (only ASCII file type supported)
LIN      Arc/Info line file (ASCII file type)
CRD      Arc/Info point file (ASCII file type)
```

Model Input Files

```
GSX      DSSAT v3.5 spatial FILEX experiment file
FL       DSSAT v3.5 FILEX *FIELDS section file
TR       DSSAT v3.5 FILEX *TREATMENTS section file
```

Common Files

```
VAL      Values file associated with a grid file or coverage
DAT      Database file for the mapping module SGIS
LUT      Lookup table file for the mapping module SGIS
```

An Example

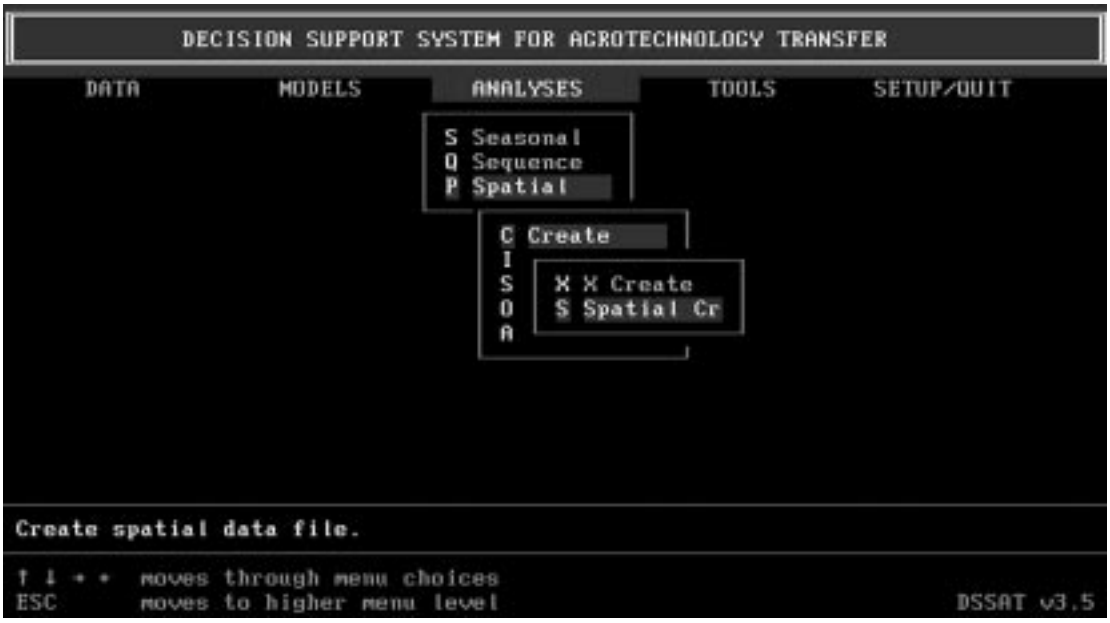
Some coverages are distributed with the DSSAT software. They can be used to illustrate the use of SPAXUTE. The region is an (essentially imaginary) area about 6 km by 7 km, with 14 categories of soil type and parts of four districts or administrative units. The soils coverage is named SOIL.IMG with SOIL.DOC its documentation file, and the district coverage is called DISTRICT.IMG with DISTRICT.DOC. It should be remembered that every IDRISI coverage, raster or vector, has a documentation file associated with it. You can look at these coverages with IDRISI if you have a copy, or by using the mapping program in the DSSAT through the "Utilities" menu item of SPAXUTE. The soil values file, SOIL.VAL, consists of the following:

```
000
001   IB00000001
002   IB00000002
003   IB00000003
004   IB00000004
005   IB00000005
006   IB00000006
```

```
007   IB00000007
008   IB00000008
009   Unknown
010   IB00000009
011   IB00000010
012   Water
013   IB00000011
014   IB00000012
```

The code by each identifier is the 10-character DSSAT soil identified, apart from two categories ("Unknown" and "Water"). The district image is made up of four categories, numbered 1 to 4.

Go to the DSSAT shell, and choose Analyses—Spatial—Create and select the spatial FILEX utility program SPAXUTE. Choose main menu option 1, "Overlay IDRISI coverages" (Screen 1). Now add an entry to the list. Select the file DISTRICT.IMG; do not select a values file for it (press the <ESC> key), and choose the "Cultivar" header. Note that entry number 1 is now displayed in the list as specified. What this means is that the program will use the district coverage to assign the cultivar level in the FILEX that we are going to create. Because there are four district identifiers, the FILEX will have four cultivar levels, with level 1 in those grid cells corresponding to district number 1, level 2 in district number 2, and so on.



Screen 1.

Add a second entry to the list — SOIL.IMG, and this time select the values file SOIL.VAL and the header "Soil Type" (you will need to scroll down the headers list to find this). Once you have two entries in the list that look like this,

```
+—  Image File   Values File  Head  -+
|  1 DISTRICT.IMG              CU     |
|  2 SOIL.IMG     SOIL.VAL     ID_S   |
|                                     |
```

execute the list, using the default coverage and values file names, TREAT.IMG and TREAT.VAL (Screen 2). This may take a few seconds, depending on the speed of your computer. Once completed, go back to the main menu, and choose "Utilities", and "Edit a file". Pick the file just created, TREAT.VAL, and you should see three columns, headed "@PID CU ID_SOIL", which are the polygon identifier, the cultivar level, and the soil identifier, respectively. There should be 30 entries in this file, that is, 30 distinct combinations of soil type and cultivar (which is really district in this example). Notice that entries PIDs 21 and 22 have a soil ID of "Water" and "Unknown" respectively. Exit the editor. You can also have a look at the coverage just created, with the 30 distinct polygons.

Now, in the main menu of SPAXUTE, choose the second option, "Convert values file to a FILEX". You will need to enter a FILEX name — select this, and enter "DTCM7001",

Screen 2.

then press <ENTER>. Note that the three slots in the file menu are now filled in with something. Ensure that the final menu item is set to "Output a skeleton FILEX". Select "Create FILEX" (Screen 3). This will take a few seconds to run. Once done, again go to "Utilities" and select to edit the file just created, DTCM7001.GSX. Note that for "treatments" 21 and 22 (corresponding to PIDs 21 and 22), the actual values of "Water" and "Unknown" from the values file SOIL.VAL have been placed in the slots for the soil ID. This illustrates an important point: the program does not check that what is written to FILEX is a valid model input. You should always check the FILEX manually. This is especially important if you have entries in a coverage or in a values file for one of the factor levels that exceed 999, the current maximum number of levels allowed in DSSAT v3.5 (you will get FORTRAN "format overflow" asterisks in the relevant fields).

One way to deal with this problem is to delete entries 21 and 22 from the FILEX. You could go ahead and do this from DTCM7001.GSX from the treatments section (you could erase the relevant entries in the fields section also, although this is not required; if you do, remember to edit both tiers).

To illustrate another way in which areas can be masked out of images, we will make another FILEX, also called DTCM7001. First, overlay IDRISI coverages as before, but this time add a third entry, using the LANDUSE.IMG coverage, with no values file, and selecting the header **Mask**. If you look at the land-use coverage, it is a coverage showing

Screen 3.

areas in cereals coded with a 1 and other landuses with 0. Your list should look like this (Screen 4):

```
+—  Image File   Values File  Head  -+
|  1 DISTRICT.IMG              CU     |
|  2 SOIL.IMG     SOIL.VAL     ID_S   |
|  3 LANDUSE.IMG               Mask   |
|                                     |
|                                     |
```

Rename the treatment values and coverage files to TREAT2.VAL and TREAT2.IMG. Now process this list, and the program will ignore all grid cells that have a "mask" value of 0 (including the "Water" and "Unknown" soil types), and produce a coverage and values file for only those unique combinations of soil type and district that have a land-use value greater than 0. Then create the FILEX, giving the name DTCM7001. Notice that with the land-use mask, there are now 23 unique combinations of soil type and district.

In general, some care should be taken in combining coverages using SPAXUTE. If, for example, no values file is specified for a coverage, then the actual values that occur in the coverage itself will be used in determining unique combinations of factors. For the standard FILEX factor levels (such as cultivar, CU, initial conditions, IC, and

Screen 4.

simulation controls, SM), this presents few problems. Values from 1 to 999 are allowable, and these values are read from the coverages (or the appropriate values files) and manipulated as integers. For some of the variables in the *FIELDS section, however, problems may arise. Thus if you specify a soils coverage with no associated values file, then the identifiers in the coverage will not only be used to define unique combinations, but will also be written out to the FILEX itself, instead of, say, the 10-character soil identifier code that is required by the crop models.

Some variables in the *FIELDS section relate to polygons, and some to points. It should be noted that SPAXUTE handles the variables XCRD, YXRD, and ELEV in a different fashion from the other variables. These location and elevation variables make more sense related to points rather than polygons, while AREA (polygon area) has no real meaning for a point. In SPAXUTE, coordinates and elevations are written to the file as real variables. If you want to create a spatial FILEX from a digital elevation model (DEM) and a set of points using SPAXUTE (such as for carrying out a geostatistics experiment and analysis, see Section 5), then you need to ensure that you have a DEM image in IDRISI format and a points image file that will be used as a mask (each point should have a unique identifier in the image). In producing the treatment coverage, you would combine the following coverages:

- The points image with the mask header;
- The DEM image with the XCRD header (for the X-coordinates of the points);
- The DEM image with the YCRD header (the Y-coordinates); and
- The DEM image with the ELEV header (the elevation).

You may select coordinate and elevation headers for a "polygonal" image (as opposed to a "points" image), but you should be aware that the values of these variables that are extracted from the relevant images will pertain to the last occurrence of a gridcell with the appropriate identifier in the image file (the image is read from the top left-hand corner to its bottom right-hand corner). An example of creating a DEM is shown in Thornton *et al.* (1997a).

It should also be noted that if you create the treatment values file using SPAXUTE, then this is created as a direct access file with a record length of 100; direct access greatly speeds program performance when it comes to writing out the appropriate parts of the FILEX. If you create the treatment values file using some other method, then there is no need to worry about the format of the file; the program writes out a copy of the treatment values file called XX.YY with the required access and record length, and this is the copy that is used in SPAXUTE (from Menu Item 2 onwards, Figure 2). The temporary copy is deleted at the end of processing. If you find a file

named XX.YY in your spatial data directory, it may be left over from a less-than-graceful exit from SPAXUTE, in which case you may delete it with impunity.

Using INCLUDE Blocks in a FILEX

Another change to the FILEX that has been implemented in the spatial analysis modules of DSSAT v3.5 is to permit the use of INCLUDE statements. This allows parts of a FILEX to be stored in separate, distinct files, so that very large and unwieldy FILEXs can be constructed and maintained more easily. Assume we have an experiment file named ABCD8101.GSX. With the new INCLUDE statement, fragments could be written as follows:

```
(in file ABCD8101.GSX ...)

*TREATMENTS
#INCLUDE 'ABCD8101.TR '

*FIELDS
#INCLUDE 'ABCD8101.FL '
```

Then the file ABCD8101.TR would contain the following:

```
*TREATMENTS                       -------------FACTOR LEVELS------------
@N R O C TNAME.................CU FL SA IC MP MI MF MR MC MT ME MH SM
 1 1 0 0 Field 1                1  1  0  1  1  0  1  1  0  0  0  0  1
 2 1 0 0 Field 2                1  2  0  1  1  0  1  1  0  0  0  0  1
 3 1 0 0 Field 3                1  3  0  1  1  0  1  1  0  0  0  0  1
 4 1 0 0 Field 4                1  4  0  1  1  0  1  1  0  0  0  0  1
```

and the file ABCD8101.FL the following:

```
*FIELDS
@L ID_FIELD WSTA.... FLSA  FLOB  FLDT  FLDD  FLDS  FLST SLTX SLDP ID_SOIL
 1 FLSC0001 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000002
 2 FLSC0002 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000021
 3 FLSC0003 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000031
 4 FLSC0004 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000004
```

The names of the included files may be anything the user likes. Currently, they must be located in the same directory (no path may be specified), in other words the spatial data directory as specified in DSSATPRO.FLE (C:\DSSAT35\SPATIAL, by default). However, it makes sense to give INCLUDE files the same name as the experiment to which they refer, and the extension might reasonably be the factor level code (FL for fields, for example, with TR for treatments, although TR is not a factor level itself).

From the DSSAT, the program SPAXUTE can be used to create the TREATMENTS and FIELDS sections of the FILEX. These will be stored as separate files that can be included with INCLUDE statements from a base FILEX. To do this, the output option can be toggled (see Main Menu item 2.5, Table 2) from the default setting, "Output a skeleton file", to the alternative setting, "Output INCLUDE files only". If this option is selected, the FILEX name selected may exist or it may not; the filename is used to name the output files, which will be given extensions FL and TR for the fields and treatments sections, respectively (UFGA8101.FL and UFGA8101.TR, in the example discussed above). The main FILEX itself, UFGA8101.GSX, will need to be created or edited in a separate operation, with INCLUDE statements pointing to the two files just created in the appropriate sections (*FIELDS and *TREATMENTS).

It is very important to note that there are only a few DSSAT v3.5 modules that allow INCLUDE statements in a FILEX. These are the following:

- The spatial FILEX utility SPAXUTE.EXE, described in this section;
- MAPDRV.EXE, the spatial model driver, described in the next section. MAPDRV can handle include files in the FILEX whether it is run as part of the DSSAT v3.5 "Spatial Analysis" or as a component of AEGIS/WIN;
- MAPEXP.EXE, the mapping analysis program described in Chapter 4, which needs access to FILEXs to ascertain relevant economic factor levels for price-cost files;
- The geostatistics software SPATSHLL described in Chapter 5, which handles INCLUDE statements in the FILEX for analysis and manipulation.

No other DSSAT version 3.5 modules will recognize INCLUDE statements, including XCREATE, so do not attempt to use XCREATE if you want to use INCLUDE statements in your experiment files.

SPAXUTE Error and Information Messages

The error and information messages are largely self-explanatory. A code of 0 indicates a fatal error (the program will terminate immediately), 1 a non-fatal error, although in this case progress is not possible (and may prove fatal also).

```
0 007 Cannot access file :
Usually this arises when the file DSSATPRO.FLE, the DSSAT set up
file, cannot be found.

1 008 Read error in subroutine GETCOD, file :
In reading the DSSATPRO.FLE setup file, an error occurred; check the
format of the file.
```

0 009 Code could not be found :
A required code could not be found in the DSSATPRO.FLE setup file. Check the file format.

1 055 Problem in finding the appropriate help screen
Help screen not found; check that the file SPAXUTE.HLP exists in the program directory.

1 058 There were no matching files found!
No files matching the mask requested were found in the current directory.

1 060 The requested subdirectory to be searched doesn't exist

1 205 An error occurred trying to edit a file

1 206 An error occurred trying to rename a file

1 207 An error occurred trying to copy a file

1 208 The list is currently empty — choose option above
This occurs if the coverage list is empty and the "execute list" option is chosen. Put something in the list to process.

1 209 Choose a name for FILEX before proceeding

1 211 The file must be in the current data directory .

1 213 Choose a coverage before proceeding

1 214 Choose a values file before proceeding

1 215 A file name must be specified for each image

1 216 File not found

1 217 Image document file not found :

1 218 Selected images have different numbers of rows or columns
This indicates that of the coverages selected, one or more of the images has a different number of rows and columns from the others. Images to be overlaid should be of the same size (and, obviously, relate to the same area).

1 219 Unknown file type,
A file type other than ASCII or BINARY was read from the image document file.

1 220 Unknown data type,
A data type other than BYTE, INTEGER or REAL was read from the image document file.

1 221 Too many rows or columns in selected image files
Image files have either more than 900 rows OR more than 900 columns OR more than 900 rows and 900 columns.

1 223 Maximum is 10 coverages — delete an entry to add another

1 231 Coverage file lacks an extension
You must specify an extension for the coverage—IMG for an IDRISI image, ARC for Arc/Info grid file, GRD for a Surfer grid file.

1 232 Coverage extension not valid (IMG, ARC, GRD)
See error 231.

1 233 One or more file names missing

1 234 File read error

1 235 WARNING—Number of orphan coverage classes found :
Entries found in coverage with no corresponding entries in the values file; defaults will be used for all variables in the *TREATMENTS and *FIELDS sections of the FILEX.

1 236 Output image name must differ from all input image names
The name of the output image cannot be used as the name of an input image; choose another output file name.

1 237 Output values file name must differ from all input values names
The name of the output values file cannot be used as the name of an input values file; choose another output values file name.

1 238 Coverage ID greater than 999,
An identifier was found in the image file specified that exceeded 999, the current limit in the DSSAT v3.5 for the number of treatments and factor levels.

1 239 Number of treatment combinations exceeds 999
In processing the images selected, more than 999 unique treatment combinations have been found. This exceeds the current limit in the DSSAT version 3.5 for number of treatments and factor levels. Simplify your coverages or reduce the number overlaid if this error occurs.

CHAPTER THREE.
RUNNING CROP MODELS WITH A SPATIAL FILEX

General Considerations

The spatial analysis driver program takes the FILEX selected or created by the user, and runs through each treatment in the experiment, calling the appropriate crop model. The spatial driver is a variant of the seasonal analysis driver, and like that program it can be used to generate replicated, spatial experiment output files.

The driver is a FORTRAN program that allows the user to pick an entry from EXP.LST (i.e., a particular FILEX). The FILEX is then read and various controls are set. Each treatment specified in the FILEX is then run in its order of appearance in the *TREATMENTS section of the FILEX. The appropriate model is called for each treatment (the model is run under the command of the driver program), and it will be run for as many replicates as are specified by NYERS in the FILEX. When all the treatments of the selected simulation have been run, the results listing file OUTPUT.LST is updated with the name and description of the summary output file produced. The user can then quit the program or choose another FILEX for running with the spatial driver. The principal files associated with the driver, MAPDRV, are shown in Figure 3.

The spatial driver can be run in a stand-alone mode or through the DSSAT v3.5 shell. It is important to note that in which ever mode the driver is run, the program expects to find all the model program(s) in the same directory (in the case of the DSSAT v3.5 shell, this will generally be the C:\DSSAT35 directory). If model executable files are not found in that directory, then an error message is printed and the user must arrange the executable files so that this condition is met. The file DSSATPRO.FLE is also expected to be in this same directory, with appropriate pointers to the relevant data and program directories. From within the DSSAT v3.5, the current directory for seasonal runs will generally be C:\DSSAT35\SPATIAL, where such files as EXP.LST, OUTPUT.LST and appropriate FILEXs are stored. The file EXP.LST may contain up to 99 experiments, and each FILEX may include up to 999 “treatments” (FILEX treatment entries).

The user interface of the driver program operates in the same way as the seasonal driver; see DSSAT v3 (Tsuji *et al.*, 1994), Volume 3 Chapter 1, pp. 14-17. Once it is activated, if the colors on the screen are difficult to see (such as when the program is being run on a monochrome laptop computer), you can force the program to display screens in monochrome by pressing the ALT-F2 keys. Note that you cannot reverse the color scheme during a session with the program, nor is your preference stored

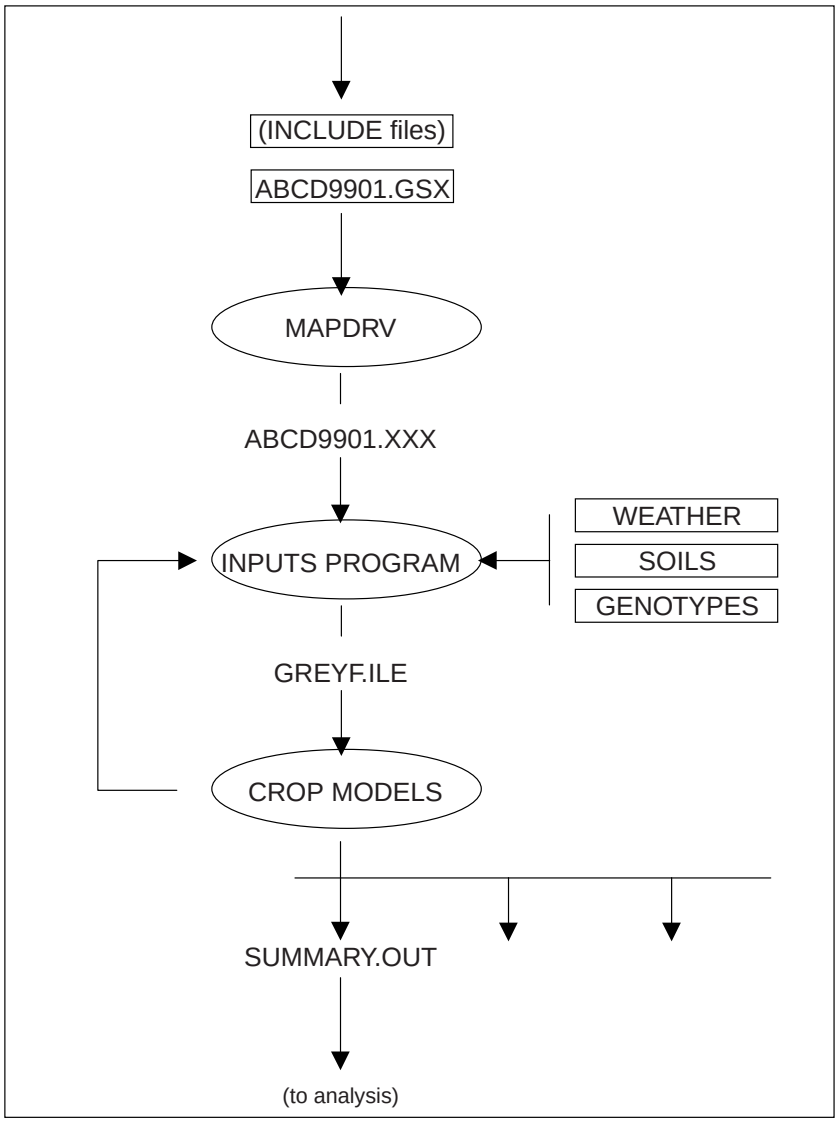


FIGURE 3. RUNNING A SPATIAL FILEX WITH MAPDRV

from one session to another. Having chosen a valid FILEX name, the driver checks to see if there is an existing output file with the same name in the current directory (controlled within FILEX). If there is, then the user is warned that the output file will be overwritten. You can exit the program at this point to rename or move the existing output file. The driver then issues appropriate commands and calls up the crop model(s). Once simulations are under way, each model run is summarized on the screen, so that the user can chart progress of the experiment. As for seasonal and sequence FILEXs, experiments with many treatments and/or many replicates may take

a great deal of time to simulate, especially on older personal computers. No keyboard input should be made, unless an error occurs, until the simulations have finished. To abort the simulation runs before they have finished, press the escape key <ESC> (you may need to keep trying, to interrupt the driver successfully).

For each FILEX simulated, the driver updates or creates a listing file, which stores the names of the summary results files (in similar fashion to EXP.LST, which stores the names of available FILEXs in the current directory). This file, OUTPUT.LST, is of the same format and is used by the analysis programs described below. Simulation results are stored in the summary output file, with extension GSS.

Driver Operation

The spatial driver operates somewhat differently from the seasonal driver. Because of the INCLUDE option now allowed in spatial FILEXs, the driver first reads the selected FILEX and adds in the lines from any INCLUDE files, to create a temporary file named (say) ABCD9901.XXX. This is the file that is read by the models. This means that the driver copies the existing EXP.LST to EXP.OLD, and a new EXP.LST is created with one entry, the temporary file ABCD9901.XXX. It may also be noted that, unlike standard FILEXs, this file is created and read with direct access, greatly facilitating speedy and efficient access. One implication of the temporary FILEX is that, if the driver and models fail for any reason, and the run is aborted, then it may be that the temporary copy of EXP.LST with the single entry is left "stranded" in the data directory. In this case, EXP.LST can be updated from the shell using the file management option (INPUTS). Alternatively, go to DOS, and in the data directory, issue the two following commands:

```
DEL EXP.LST
REN EXP.OLD EXP.LST
```

Another feature of the spatial driver is that the program checks for unique treatments. If a treatment is found that has already been simulated, then that treatment is not re-simulated, but the appropriate output record is copied to the correct place in the summary output file. In other words, the one-to-one (or one-to-many for a replicated experiment) correspondence between treatments in the FILEX and outputs in the summary output file is preserved. ***THE SUMMARY OUTPUT FILE IS THE ONLY OUTPUT FILE FOR WHICH THIS IS TRUE;*** in the other output files, if already-simulated treatments are skipped, their outputs are NOT written to the appropriate places. This feature is true for MAPDRV if it is run from AEGIS/WIN or from the DSSAT shell.

Currently, the definition of an identical treatment is as follows: the factor levels are the same, except for the fields level FL; but in the fields section, if the weather ID, the soil ID and the soil depth is the same as a previous treatment, then the treatment is judged identical. In terms of the models, none of the other variables in the *FIELDS section have any impact on current capabilities. The program scans for identical treatments at the start of the run. If any are found, a message is echoed to the screen, and the user must press a key to continue with the model runs.

It is possible to run the spatial model driver in batch mode (this might be useful if the program is called from another software package, for example; this is also how the driver is run from AEGIS/WIN). The command to do this is as follows:

```
C:\DSSAT35\SPATIAL> c:\dssat3\mapdrv n
```

where the uppercase path is the DOS prompt from the current data directory, the lowercase portion is the command, and <n> is a number that corresponds to the entry number in the file EXP.LST. If <n> is not valid (less than or equal to zero, or greater than the number of entries in EXP.LST), then an error message is echoed to the screen, and batch mode running is not possible. If <n> is valid, then the driver will process the input files and run the models to completion with no keyboard input required.

You can use the seasonal analysis program VARAN2 to look at the summary output files generated from spatial FILEXs. The program VARAN2 accesses output file names from the same listing file, OUTPUT.LST, as is updated by the driver. This is a change from DSSAT version 3, where each listing file had a different name, depending on the type of analysis performed (i.e., SEASONAL.LST and SEQUENCE.LST). Output listing file names are now common, regardless of type of analysis (OUTPUT.LST).

As for seasonal and sequence analysis, there are a number of FILEX run control options that are important to set correctly. One of these is the summary output file option (SUMRY), which should be set to A (for "Append") in spatial FILEXs, if output file headers are to be correctly written out. This will also avoid the writing out of (unprompted and usually unexpected) Y and Z output files. The output filename code FNAME should be set to Y, ensuring that output files are named in accordance with the input FILEX name (otherwise generic output names are used, such as "SUMMARY.GSS", which are successively overwritten). It should also be remembered that if you have a FILEX with multiple simulation control blocks, then the number of replicates specified (NYERS) should be the same in all blocks. Different values will result in unpredictable behavior by the driver, and there will be a good possibility of errors occurring with respect to run counters.

An Example

Before running the FILEXs created at the end of Section 2, you may need to update the EXP.LST experiment listing file. Either use the "ANALYSES—SPATIAL—INPUTS" option from the shell, or you can edit the file directly. Once this is done, start the model driver from the shell with the "SIMULATE" option, and choose the file DTCM7001.GSX. This is an experiment with 5 replicates of 23 or 28 treatments, depending on how you created it in Chapter 2, so it may take a while to run. If the driver gives you an error, check that you deleted treatments 21 and 22 from the DTCM7001.GSX file (the soil ID for these treatments was not valid), if you did not use the "mask" option in section 2.3 to create it. We will look at this experiment below.

MAPDRV Error and Information Messages

Various information and error messages are produced by the driver program. The format was described in Chapter 2.

```
0 001 Cannot find experiment file :
The FILEX experiment listing file could not be found in the current
directory.

0 002 No entries found in file :
The FILEX experimental listing file was found, but it contains no
valid entries.

1 003 Cannot find file :
The specified file cannot be found (non-fatal).

0 004 Error reading :
An undefined read error occurred when attempting to read the file.
The problem is usually caused by an incorrect format of the file.

0 005 No valid treatments found in file :
No valid treatments were found in the FILEX selected.

0 006 Error in *SIMULATION CONTROL factor levels :
The FILEX specified by the user is defective:  FNAME was not set or
not found, or NREPS was not set or not found, or the simulation
control factor level with the lowest number specified by the user
did not have a corresponding *SIMULATION CONTROL section.

0 007 Error in *CULTIVAR factor levels in file :
A cultivar factor level was specified in the selected FILEX that did
not have a corresponding cultivar description in the *CULTIVARS
section of the file.
```

1 008 Model runs completed!

1 009 SUMMARY.OUT renamed to SUMMARY.GSS

1 010 OUTPUT.LST updated, added file :
Specified summary output file added to the listing file.

0 011 A required file was not found .. aborting :
This usually refers to the absence of the file DSSATPRO.FLE, which could not be found.

0 012 Excessive path length error :
You need to reduce the path length as specified in DSSATPRO.FLE for the crop model "cc" (e.g. MZ). Path lengths as specified in the DSSATPRO.FLE must be less than 50 characters in total length, including drive letter, colon, and leading and trailing slashes (\). If these paths are modified using the SETUP menu of the DSSAT v3.5 shell, then there is no problem, as the screens will not allow you to enter paths longer than this. If you modify DSSATPRO.FLE with a text editor, then you need to ensure that this total length is not exceeded.

0 013 No code error :
The crop code specified (e.g. MZ) could not be found in the file DSSATPRO.FLE. This will occur if a cultivar is specified in FILEX that does not have a valid crop code.

0 014 Read error in file :
An undefined read error occurred when reading the specified file; check the format of the file.

0 020 A problem was encountered executing another program
The driver was not able to execute another program; check the free RAM that you have on your computer, and increase it if possible by unloading unnecessary resident programs.

0 021 A problem was encountered erasing a file
The specified file could not be erased.

0 022 A problem was encountered copying a file
The specified file could not be copied successfully.

1 055 Problem in finding the appropriate help screen
The required help screen, or the help screen file MAPDRV.HLP, could not be found.

1 092 Cannot find #INCLUDE file :
An include file referenced in a FILEX could not be found in the data directory.

1 093 Effectively identical treatments found :
One or more treatments were found to be identical (as defined above)—a warning only.

1 094 Direct access read error in file :
A read error in the temporary FILEX, probably due to a format problem in the original file(s).

1 095 Read or treatment error in updating output file :

1 096 Direct access read error (second entry) in file :
A read error in the temporary FILEX, probably due to a format problem in the original file(s).

1 100 Error!
Undetermined error.

CHAPTER FOUR.
ANALYZING SPATIAL EXPERIMENTS

General Considerations

There are three ways in the DSSAT shell to analyze spatial experiments. First, the seasonal analysis program can be used to tabulate and graph results from the summary output file, but note that the limit of 20 treatments in the program will prevent large files being analyzed completely (one way to avoid this limit is to chop up large summary output files into several smaller ones, each with 20 treatments or less). The program is described in detail in DSSAT v3 (Tsuji *et al.*, 1994), volume 3.1, on pages 23-53.

Secondly, the geostatistics analysis package can be accessed from the shell. These programs are described in detail in this document in Chapter 5.

The third option is to use a utility called MAPEXP that reads the summary output file and allows the user to map results. The notes in this Section describe this utility.

The menu options are listed in Table 3, and the principal files for the program, MAPEXP, are shown in Figure 4. The program operates very like the seasonal analysis program. First, an output file is chosen from those available in the output listing file OUTPUT.LST (Menu option 1, Table 3), after which the user can build up a list of variables to map. The list can hold up to five variables. The variables available for mapping are all those that are written to the summary output file, with the addition of gross margin per hectare. These are listed in Table 4. For each variable chosen, if each treatment was replicated, a number of statistics can be mapped: the mean, the standard deviation, the minimum value, the 10th, 25th, 50th, 75th, or 90th percentiles, the maximum value, and the coefficient of variation of the distribution. Alternatively, you can choose to map a particular replicate's value, in which case you are prompted for the replicate number to map. Once the list contains at least one entry (if more than 5 are added, the fifth entry in the list is simply overwritten with the most recent entry), you can map the outputs. Before this is done, you must select a base map file. This will normally be the "experiment" coverage derived earlier (see Chapter 2 above).

TABLE 3. MAPEXP MENU STRUCTURE

MAIN MENU	SECONDARY MENU	TERTIARY MENU
1 Select a spatial output file	1.1 Choose file	
2 Select model outputs to map	2.1 Choose a variable	
	2.2 Choose statistic	2.2.1 Pick replicate
3 Map model outputs in list	3.1 Map (hot key)	
	3.2 Select base map name < >[1]	3.2.1 Select file name
	3.3 Select a LUT file < >	3.3.1 Select file name
	3.4 Select price VAL file < >[2]	3.4.1 Select file name
4 Utilities	4.1 Edit a file	4.1.1 Select file name:
	4.2 Rename a file	4.2.1 Select file name: 4.2.2 Rename to:
	4.3 Copy a file	4.3.1 Select file name: 4.3.2 Copy to:
	4.4 Call up the mapping module	
	4.5 Exit	
5 About MAPEXP		
6 Exit		

[1] The < > indicates that the appropriate file name is echoed to the menu screen (either a default name or the name selected by the user).

[2] This item appears only if gross margin per hectare is one of the output variables to be mapped.

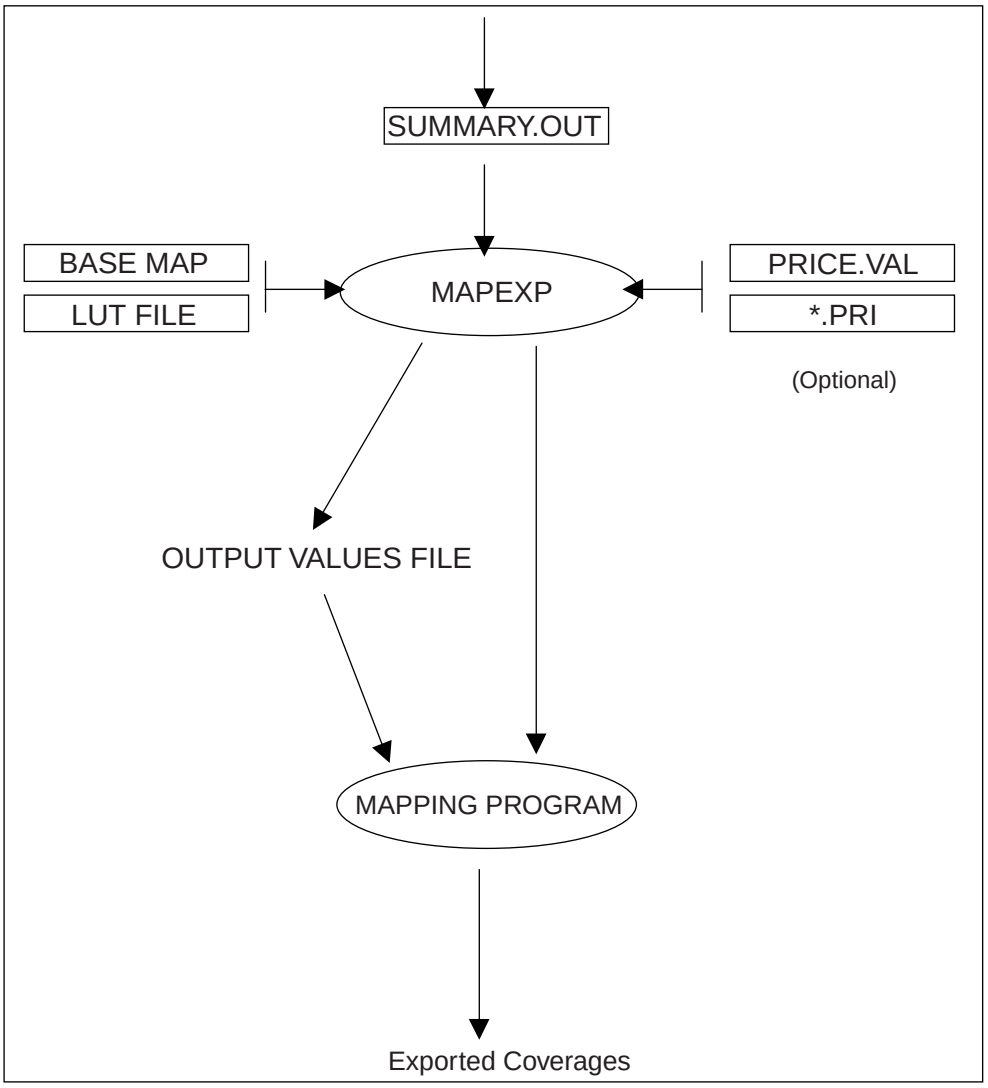


FIGURE 4. RUNNING SPATIAL ANALYSIS WITH MAPEXP

Table 4. Variables Available for Mapping from the Program MAPEXP. For Variables 1 to 5, YRDOY is the Date Expressed as Year Number, Followed by Day of Year.

	Code	Description
1	SDAT	Simulation starting date (YRDOY)
2	PDAT	Planting date (YRDOY)
3	ADAT	Anthesis date (YRDOY)
4	MDAT	Physiological maturity date (YRDOY)
5	HDAT	Harvest date (YRDOY)
6	DWAP	Weight of planting material (kg DM ha^{-1})
7	CWAM	Weight of plant tops at maturity (kg DM ha^{-1})
8	HWAM	Yield at maturity (kg DM ha^{-1})
9	HWAH	Yield at harvest (kg DM ha^{-1})
10	BWAH	Harvested byproduct (e.g. straw) dry weight (kg DM ha^{-1})
11	HWUM	Harvest product individual weight (mg DM)
12	H#AM	Harvest product number at maturity (number m^{-2})
13	H#UM	Harvest product number per unit at maturity (no. $unit^{-1}$)
14	IR#M	Number of irrigation applications during the season
15	IRCM	Total seasonal irrigation applied (mm)
16	PRCM	Total seasonal precipitation (mm)
17	ETCM	Total seasonal evapotranspiration (mm)
18	ROCM	Total seasonal surface runoff (mm)
19	DRCM	Total seasonal water drainage (mm)
20	SWXM	Extractable soil water in the profile at maturity (cm)
21	NI#M	Number of nitrogen applications during the season
22	NICM	Total inorganic nitrogen applied (kg ha^{-1})
23	NFXM	Total nitrogen fixed during the season (kg ha^{-1})
24	NUCM	Nitrogen uptake by the plant during the season (kg ha^{-1})
25	NLCM	Nitrogen leached from profile during the season (kg ha^{-1})
26	NIAM	Soil inorganic nitrogen at maturity (kg ha^{-1})
27	CNAM	Nitrogen content of plant tops at maturity (kg ha^{-1})
28	GNAM	Harvested product nitrogen content at maturity (kg ha^{-1})
29	RECM	Total crop residue applied (kg ha^{-1})
30	ONAM	Organic soil nitrogen at maturity (kg ha^{-1})
31	OCAM	Organic soil carbon at maturity (kg ha^{-1})
32	PO#M	Number of phosphorus applications during the season
33	POCM	Total phosphorus applied (kg ha^{-1})
34	CPAM	Phosphorus in the plant tops at maturity (kg ha^{-1})
35	SPAM	Soil phosphorus content at maturity (kg ha^{-1})
36	GMAR	Gross margin (gross return minus variable costs) ($ ha^{-1})

You can also choose a look-up table (LUT) file with which to map the results. This controls the colors used in the map. To produce the maps, the program calculates the necessary statistics and writes out what is, in essence, a values file for the coverage

chosen. The DSSAT mapping utility SGIS9801 is then called. Operation of this mapping utility is discussed in Chapter 6. MAPEXP also allows the user to edit, rename and copy files, and to enter the mapping utility directly (as for the spatial FILEX utility program SPAXUTE.EXE).

Gross Margin Calculations

The thirty-sixth variable that can be mapped (Table 4) is gross margin per hectare. This has to be derived from the summary output file, and is calculated in the same way as in seasonal analysis (see DSSAT version 3 (Tsuji *et al.*, 1994), Volume 3, pp 34-40). Spatially variable price files can be used for these calculations, based on the header EC and associated values that may be specified in a treatment values file.

The analysis program will attempt to access price files directly from the treatment values file; names of these price files are assembled from the entry in the values file plus the extension ".PRI". All price files are assumed to be located in the economic data directory specified in the DSSAT v3.5 setup program (slot ECD, with default value C:\DSSAT35\ECONOMIC). Thus if you overlay a price file coverage directly with an associated values file using SPAXUTE, for example, you may obtain a treatment values file as follows:

```
@PID    ID_SOIL  CU  EC
0001 IB00000005   1   1
0002 IB00000003   3   3
0003 IB00000010   1   1
0004 IB00000002   1   1
:
:
```

In this case, the program will search for price files named "1.PRI", "3.PRI", etc. It would make more sense to produce the treatment values file using a price-file values file, resulting in something such as:

```
@PID    ID_SOIL  CU       EC
0001 IB00000005   1 ABCD0001
0002 IB00000003   3 ABCD0003
0003 IB00000010   1 ABCD0001
0004 IB00000002   1 ABCD0001
:
:
```

In this case, the program will search for price files named "ABCD0001.PRI" and "ABCD0003.PRI", etc. If you choose to map gross margins, you must enter the name

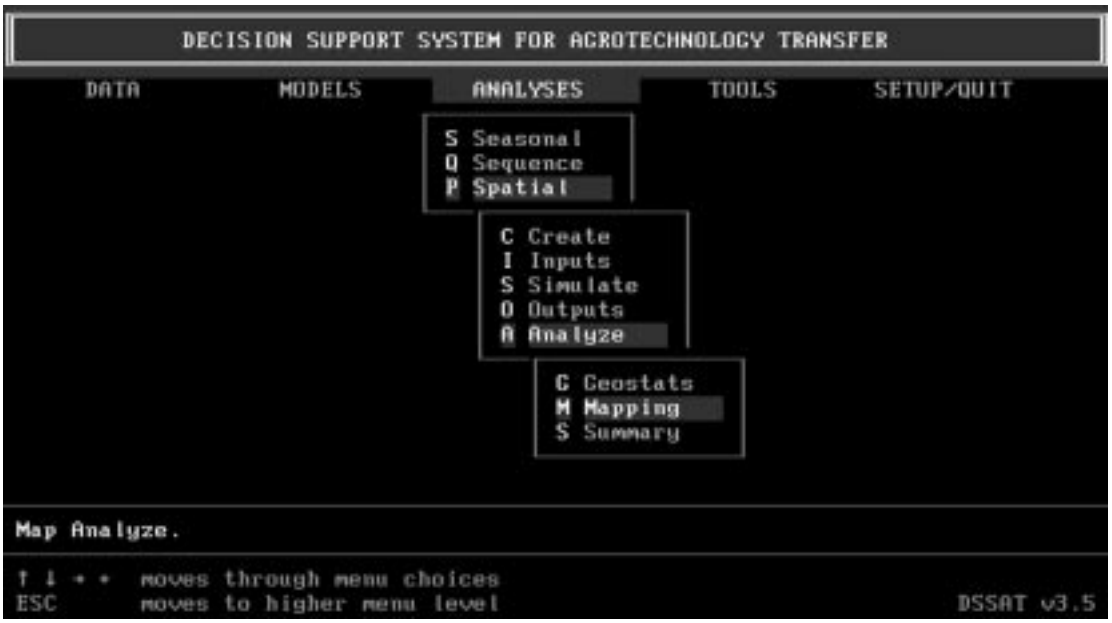
of the treatment values file, assumed to be in the current directory (in distinction to the price files themselves, which, as noted above, are located by default in the DSSAT35\ECONOMIC subdirectory, but controlled in the file DSSATPRO.FLE which is accessible through the SETUP option in the DSSAT shell). In reading the price files, no variability is taken into account (i.e., IDIS = 0, see the file DEFAULT.PRI in the default directory C:\DSSAT35\ECONOMIC), and the prices and costs used will be those that appear in the first block in the file with the appropriate 2-character crop code header (e.g., "MZ" for maize). If no price files can be found for the relevant crop codes, then default costs and prices are used as follows:

	Cost/Price	Default Value	
1	Price of grain	$110.00 t^{-1}	
2	Price of byproduct (e.g., stover)	$ 0.00 t^{-1}	
3	Base production costs	$240.00 ha^{-1}	
4	N fertilizer cost	$ 0.45 kg^{-1}	
5	Cost per N fertilizer application	$ 12.00	
6	Irrigation costs	$ 0.50 mm^{-1}	
7	Cost per irrigation application	$ 12.50	
8	Seed cost	$ 0.46 kg^{-1}	
9	Organic amendments	$ 0.00 t^{-1}	
10	P fertilizer cost	$ 0.00 kg^{-1}	(ignored)
11	Cost per P fertilizer application	$ 0.00	(ignored)

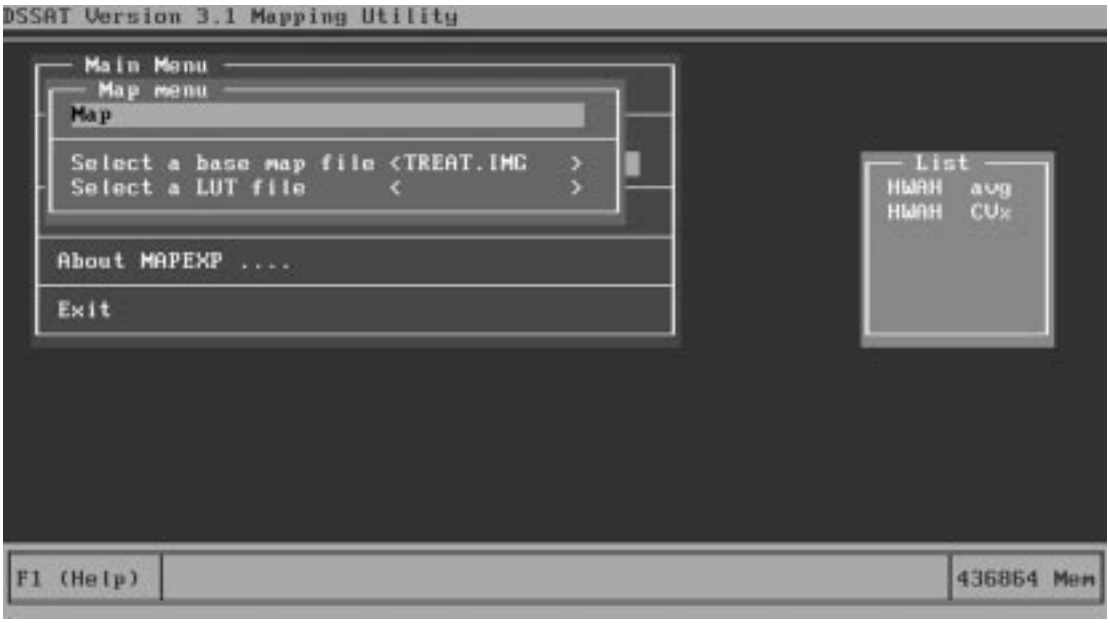
An Example

Select the mapping and GIS utility program from the "SPATIAL—ANALYSES" menu of the DSSAT shell (Screen 5). First, you must select a spatial output file for analysis. Pick the experiment DTCM7001. Note the summary screen that appears after the program has read the file, especially the number of runs and the replications found. Now pick menu option 2, "Select model outputs to map". A scrollable list appears with the 36 outputs (Table 4). Pick harvest yield (number 9); now pick a statistic from the brown menu — choose the Mean. The green list on the right of the screen now shows the first entry, with the header HWAH (for harvest yield) and the statistic ("avg" for the mean). Choose the same variable from the list, but pick the coefficient of variation (CV) this time. The green list should have two entries. To proceed, press the escape key <ESC>, and choose to "Map model outputs in list". Select the base map file, TREAT.IMG, that we created above, and leave the LUT slot empty (Screen 6) [see Section 6 for a discussion of LUT].

Once the slots in the menu have appropriate entries, press the "Map" option. The program will produce a database file named DTCM7001.DAT consisting of three



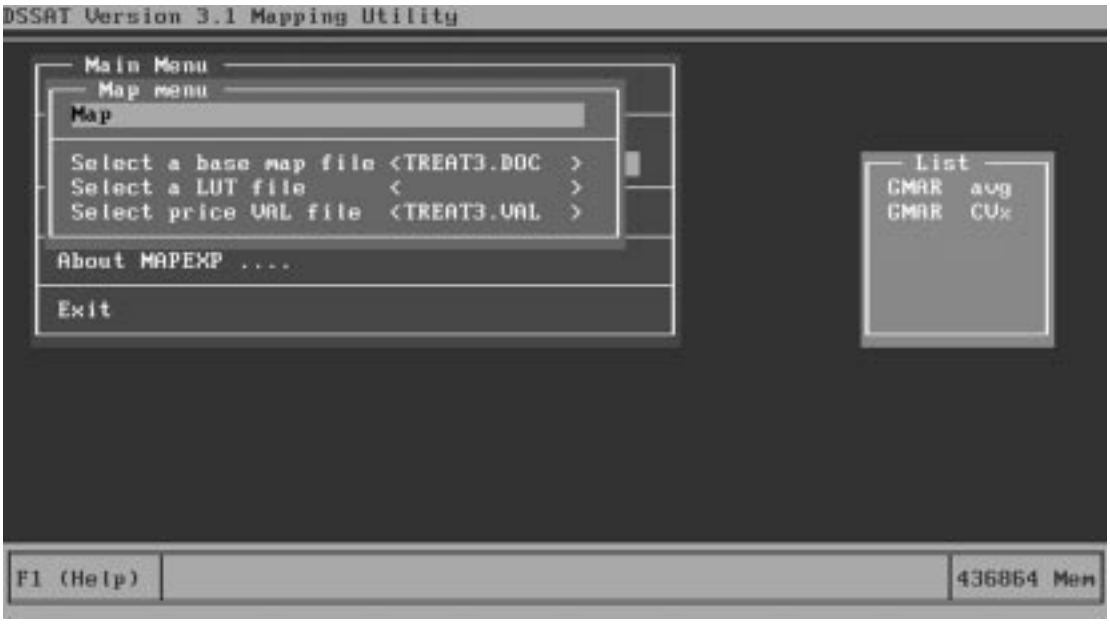
Screen 5.



Screen 6.

columns: the polygon ID, the calculated mean harvest yield, and the calculated coefficient of variation of mean harvest yield. The program then calls the mapping utility, and the first data column is mapped (mean yield). To map the second column, go to view 2 by pressing <Page Down> or "View down". Toggle the legend by pressing "L". You should be able to see clearly the influence of soil type on yields, and the polygons in the image that are not filled in, corresponding to soil categories 9 and 12 ("Unknown" and "Water", respectively). When done, press "Q" to exit, and investigate the masked outputs of DTCM7001 using the TREAT2 image and values files.

When you choose to map gross margin, you can use spatially variable price files by selecting a treatment values file with the "EC" header and variable (without this, the program will simply use the default costs and prices shown above in Chapter 4). To do this for the example, you will need to go back to the SPAXUTE program (Chapter 2) and make a new treatment values file and coverage. Proceed as you did to make TREAT2.IMG and TREAT2.VAL, but as the fourth overlay, select the district coverage DISTRICT.IMG, the price values file PRICEFIL.VAL distributed with the software, and the "prices" variable header, and name the output TREAT3.VAL and TREAT3.IMG (Screen 7). Now, back in the mapping analysis program, you can map gross margin, using the four price files ABCD0001.PRI , ABCD0002.PRI, ABCD0003.PRI and ABCD0004.PRI (located in the default C:\DSSAT35\ECONOMIC subdirectory), so that we are using different costs and prices associated with the four districts in the image.



Screen 7.

There are provisions in MAPEXP that allow you to copy, rename and edit files, and to call up the mapping utility directly. As for the program SPAXUTE, an important limit to the program is the number of items that can be handled by the scrolling routine when selecting a file name. To repeat, if there are more than 60 matching files in the directory, then only the first 60 will be picked up by the program. This may explain why you cannot apparently find a file that you know exists, but it provides a good incentive to keep your spatial data directories as uncluttered as possible.

MAPEXP Information and Error Messages

Fatal errors are indicated by an index code of 0, non-fatal errors by 1.

0 001 Cannot find directory file :
The summary output listing file could not be found in the current directory.

0 002 No entries found in file :
No valid entries were found in the summary output listing file.

1 003 Cannot find file :
The specified file cannot be found in the appropriate directory.

0 004 Error reading :
An error occurred during file reading; this is usually due to incorrect format of the specified file.

0 007 Cannot access file :
The specified file cannot be found in the appropriate directory.

1 008 Read error in subroutine GETCOD, file :
An error occurred in reading the specified file.

0 009 Code could not be found :
The program attempted to find the specified code in the DSSATPRO.FLE, but could not find it. The pathname associated with the code specifies the location of the graphics program for the seasonal analysis graphics.

1 013 Cannot find file :
The specified file cannot be found.

1 021 A problem was encountered copying a file
The specified file could not be copied successfully.

1 022 A problem was encountered erasing a file
The specified file could not be erased successfully.

1 023 Error in TUKEY.FOR
An error occurred when calculating percentiles of a distribution.

1 024 READ Error in file :
A read error occurred in the specified file; check the format of the file.

1 025 Unequal treatment reps in file :
Unequal treatment replicates were found in the selected summary output file; exit and fix the file to avoid this message.

1 026 READ Error in DATA.CDE
An error occurred when reading the file DATA.CDE; check that it is complete and the format is correct.

1 029 Number of treatments exceeds 20
See message 031.

1 030 Number of replicates exceeds 30
See message 031.

1 031 File is too big to analyze:
This error message arises in conjunction with one of three other messages, and refers to the fact that the selected summary output file is too large with respect to total number of runs (error message 028), number of treatments (error message 029), or number of replicates (error message 030).

0 038 Error in filename :
Error in filename specification.

1 039 Crop code not matched in price file :
There was no section in the price file corresponding to the specified crop code; you must edit the file so that this section exists.

1 040 Read error in price file :
Check the format of the price file; a read error occurred.

1 050 Cannot find a file .. choose again :

1 051 You must first select a model output file ...

1 055 Problem in finding the appropriate help screen
The appropriate help screen could not be found.

1 058 There were no matching files found!

1 060 The requested subdirectory to be searched doesn’t exist

1 100 Error!
Undetermined error.

1 201 Warning: any existing, unrenamed map files may be overwritten
Rename existing map files if you want to save them.

1 202 Using defaults; at least one price file does not exist :
Some gross margins will be calculated using the default prices and costs outlined in the text.

1 203 Using defaults; at least one crop code not found in file :
Default prices and costs are also used if not all crop codes can be found.

1 204 Warning: file has only one replicate anyway
Summary output file has only one replicate — no distribution variables can be calculated.

1 205 An error occurred trying to edit a file

1 206 An error occurred trying to rename a file

1 207 An error occurred trying to copy a file

1 208 The list is currently empty — choose option above
The mapping list is empty; there must be at least one entry to proceed.

1 209 Choose a base map file before proceeding
A base map must be chosen; normally this will be the “experiment” coverage, perhaps created using SPAXUTE.

1 210 The map file must be in the current data directory ...
The specified map file must be located in the data directory, as specified in DSSATPRO.FLE and the DSSAT setup.

1 211 The LUT file must be in the current data directory ...
The look-up table file specified must be located in the data directory, as specified in DSSATPRO.FLE and the DSSAT setup.

1 212 An error occurred trying to execute another program
Check your amount of free RAM, it may be a memory problem.

1 215 No price values file found ... defaults used for all polygons
The program could not find the price values file specified, so no external price files will be accessed.

0 216 Fatal format error in values file
Check the format of the values file specified

1 217 No valid price files found ... defaults used for all polygons
Price file names specified in values file could not be located; no external price files will be accessed.

0 219 Header not found in file :
Check the format of the specified file.

0 220 Price file header not found in file, default prices used :
No header “EC” was found in the treatment values file; make an overlay with this header if you want to access a set of price files.

CHAPTER FIVE.
GEOSTATISTICAL ANALYSIS OF SIMULATION RESULTS

General Considerations

Variation in space is defined as the phenomenon that something changes in space. At a certain location something is observed. At a very small distance from this place something different is observed, a deviation. The deviations may increase as the distance increases. The description of variability (how large the deviations are as the distance increases) may be important for process-based interpretation. When cultivating plants or when interpreting soil conditions, spatial interpolation is necessary (Stein, 1993).

Developments in geostatistics allow us to deal quantitatively with spatial variation. The different stages include an analysis of the spatial dependence, i.e., how large is the variation as a function of the distance between observation, production of computerized maps, and determination of the probabilities of exceeding some critical threshold values, such as leaching levels. The main distinction with statistics is that in geostatistics variables are used that are linked to locations. Some authors use terms such as "regionalized variables" or "geovariables".

This brief introduction on geostatistics within DSSAT v3.5 is not meant to be complete. Rather, it will describe the various elements used and guide the user through the software. Information about geostatistics should be obtained from specialized textbooks.

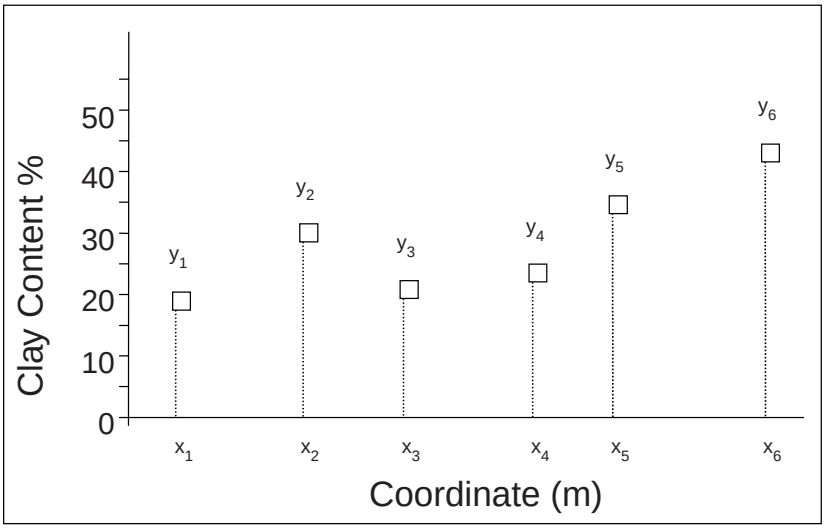


Figure 5. Data Along Transect (Adapted from Stein, 1993)

Semi-Variance

Imagine you have **n** observations (simulations) along a transect, that is a line in a field. We could think of these observations relating to clay content, crop yield, or leaf area index, for instance, Figure 5. Some observations, n=6, are taken at different distances from each other.

Along the horizontal axis the coordinate along the transect is given, measured in meters from an arbitrary origin and denoted x_1 through x_6. The observations are denoted y_1 through y_6. Using simple regression tools one could determine a regression fit which best fits the observations. This line may then be used to predict the value at each location.

In geostatistics a different method is used. Observations in space are linked through their coordinates and each observation has its own specific place in space. The value of coordinate (x) is essentially linked with variable (y). Such spatial variables are, therefore, denoted x(y) and are called "regionalized variables". One important characteristic of regionalized variables is that data close to each other are more likely to be similar than data collected at large distances from each other. Consider two points along the transect x and x+h (the latter point being located at a distance h from the first point x). Equation (1) defines the semi-variance, gamma(h), as a function of h, the lag. In one dimension, gamma(h) can be estimated at regular intervals by sampling along a transect.

$$\text{gamma}(h) = 1/(2(n-h)) \sum_{i=1}^{n-h} [z(x_i) - z(x_i+h)]^2 \qquad (1)$$

Thus given a set of values $z(x_1)$, $z(x_2)$, ..., $z(x_n)$, we can estimate the semi-variance, where h is any integral multiple of the sampling interval (Webster, 1985).

Figure 6 shows the comparisons involved for h=1, 2, and 3. The result is an ordered set of values that constitute the sample semi-variogram. These can be plotted. An example is presented in Figure 7. If the distances between the observations are not equal to each other, distance classes are created: all pairs of points with about the same intermediate distance are grouped into one distance class. An example of such a distance class is the class of all points with an intermediate distance between 0.5 and 1.5 m. To this class a pair of observations with a distance of 0.6 m is assigned, as well as a pair of observations with a distance of 1.4 m, but not a pair of observations with a distance of 1.6 or 0.3 m. The average distance within this class will be approximately equal to 1 m. The next distance class may contain all the points with intermediate distances between 1.5 and 2.5 m, etc. The width of such a distance class is called the lag length. The choice of lag length is generally up to the user, and typically influences

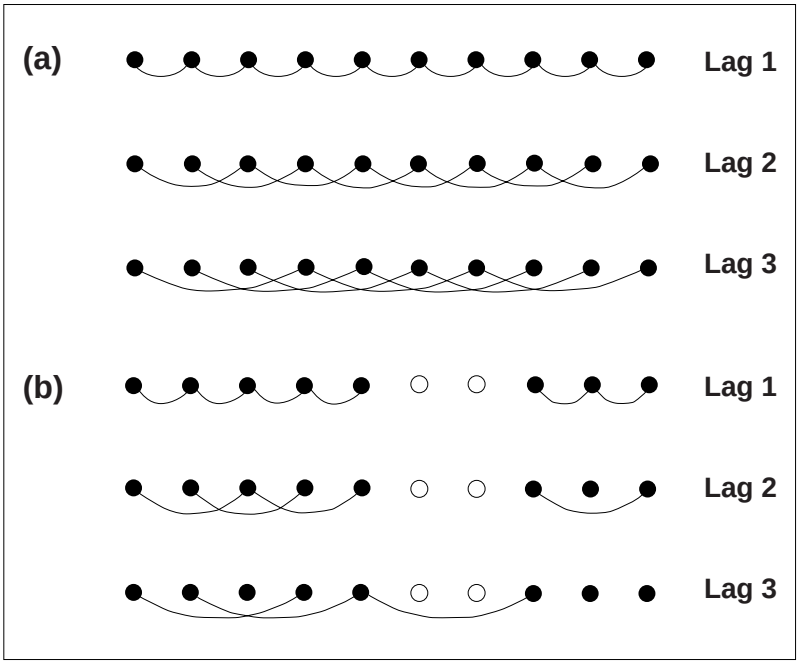


Figure 6. Comparisons for Estimating Semi-Variances on a Linear Transect at Lags of 1, 2, and 3 (Adapted from Webster 1985). A) for Complete Data. B) with Missing Observations (Open Circles).

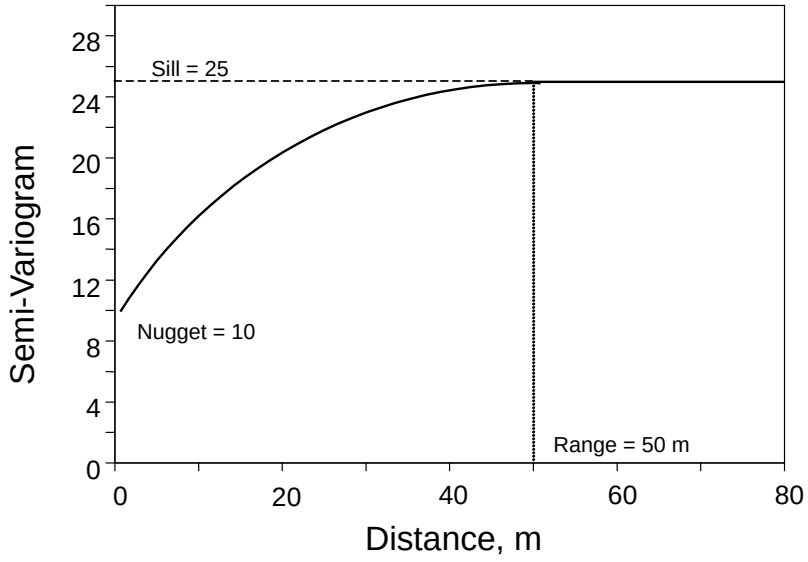


Figure 7. Sample Semi-Variogram with Characteristic Structures: Nugget, Sill and Range (Adapted from Stein, 1993).

the number of distance classes. If the lag length is chosen to be large the number of lags will decrease, and vice versa. For all practical purposes there are some general rules which have to be obeyed in order to obtain reliable semi-variogram estimates:

- The number of pairs of observations points in each class must exceed 30;
- The maximum distance h between observation points (the lag length) for which the semi-variogram may be determined should not exceed half the length of a transect, or the smallest dimension of a field or plot. If, for example, the field dimensions are 200 by 300 m, the number of lags multiplied by the lag length should not exceed 200/2=100 m; 10 lags of 10 m, or 12 lags of 8 m, are valid combinations, while 12 lags of 10 m exceed the 100 m limit and is not, therefore, a valid choice.

There are certain features of the semi-variogram worth mentioning. First, the semi-variogram is independent of the place where the regionalized variables are located. The squared pair differences have the same expectation, regardless of whether they are measured at one part of the transect or at another part of the transect. Similar differences between observations independent of the part of the area are expected. Second, in many practical studies one observes an increase of the semi-variogram with increasing distance between the observation locations. This implies that the dependence decreases with increasing distance h between locations. Loosely speaking, observations close to each other are more likely to be similar than observations at a larger distance from each other.

For spatial interpolation, it is necessary to fit a specific function (model) through the semi-variogram estimates. On account of certain (geo)statistical considerations, not all the functions one could think of are allowed. On the other hand, some functions that are allowed are seldom used in practice. In this software tool the following models have been included: spherical, exponential, gaussian, circular, bessel, linear, and linear with sill. The mathematical descriptions of these models can be found in specialized geostatistical textbooks and will not be discussed any further. It is worth mentioning that the spherical, exponential and gaussian models are the most widely used. The linear model with sill is only allowed for data along a transect; for spatial interpolation, this model is not recommended. The gaussian model is characterized by its horizontal behavior at the origin. This model is often encountered when there is uncertainty with respect to the precise location of the observations. All of these models are characterized by a number of typical structures (see also Figure 7):

- The **sill** variance [C] is the value where the semi-variogram reaches maximum variance (the observations are becoming uncorrelated);
- The **range** [r or a] of the semi-variogram is a measure of the distance up to which

spatial dependence extends—loosely speaking, the distance at which the sill value is reached. In a number of models, such as the exponential, bessel, and gaussian, the sill value is never reached. The range is therefore defined as the value where the variance reaches 95% of the sill value;

- The **nugget** effect [C_0] contains the non-spatial variability. The sources of such variability are, for instance, operator bias, measurement error, and model uncertainty. The nugget effect does not measure systematic deviation in the observations, such as the case where all measurements are shifted with a fixed constant value.

Whenever the sill value is reached, the sill/nugget ratio gives an indication of the variability to be assigned to spatial variability. If the ratio is close to 1 the non-spatial component is dominant; in the case of a value considerably larger than 1, the spatial dependence is dominating.

WARNING: *In cases where a strong trend is evident, such as simulation results on a strong hill slope, it is generally better to "de-trend" the data first, and then apply the geostatistics to the de-trended data.*

Kriging

In addition to analyzing spatial dependence of observed or simulated properties, the main reason for obtaining a semi-variogram is to use it for estimation purposes. The method of spatial interpolation estimation used is generally referred to as "Kriging", named after D. G. Krige, who devised the technique for use in South African gold-fields. It is essentially a means of weighted local averaging in which the weights are chosen so as to give unbiased estimates while at the same time minimizing the estimation variance (Webster, 1985).

In general we may wish to estimate the value of a property at a place we have not visited. By including "neighborhood" points (observations close to the point where we want the prediction) and the spatial dependence through the semi-variogram, one can make an estimate of the prediction and its variance. This procedure can be performed for individual points, series of points, or points densely covering an area (for example, to create maps or contour plots). The predicted value of a property is in that case the best estimate for that point. The variance represents its uncertainty (note that the square root of the variance of the prediction error gives the prediction error in original units, and is thus a measure of prediction accuracy).

A special case of kriging called "disjunctive kriging" (Cressie 1991, Finke and Stein, 1994) can be used to create probability maps of exceeding a specified threshold or cut-off value (such as a map with the probability of exceeding a simulated leaching loss to

the groundwater, or of exceeding a specified yield). Essentially, the observed distribution of the variables is transformed into a normal distribution by means of so-called Hermite polynomials, which are used to predict the distribution in a location, followed by a back-transformation. The probabilities vary between 0 and 1.

The Geostatistical Utility

The software included with the DSSAT v3.5 has been brought together from existing programs that have been widely tested and used. A brief summary of the software and the most relevant references are as follows:

- To calculate semi-variances: SPATANAL, developed by Staritsky *et al.* (1992);
- Semi-variances were modeled using WLSFIT, developed by Heuvelink (1995);
- Ordinary and disjunctive kriging options: DISCO, developed by Finke (1993);
- Mapping tools: USTED, developed by Stoorvogel (1995).

The Shell

When selecting the geostatistics option under the "Analysis" menu, the spatial shell is activated. The first menu prompts you to select an experiment that you want to analyze (e.g. FLSC8101). The experiments which are available are listed in the file OUTPUT.LST. After selecting an experiment, the next menu prompts for a variable that you want to analyze. This list is based on the headers in the summary output file. The last item in the list is the elevation, which is read from the input file (FILEX), and this can be used to create a digital elevation model (DEM). The third menu allows you to select an analysis tool from those available:

- Calculating semi-variances;
- Fitting a semi-variance model;
- Kriging;
- Mapping;
- File rename;
- Exit.

The different tools are discussed below. It should be noted that there is a fixed sequence of analysis which has to be followed. First, the semi-variances have to be calculated. The results from this are used in fitting the semi-variogram, which in turn is used in the kriging option.

Input files required:

```
1 DSSATPRO.FLE
2 OUTPUT.LST
3 FLSC81O1.GSS    (as an example summary output file)
4 FLSC8104.GSX    (as an example FILEX)
5 DATA.CDE        (to read headers for variables)
6 WEATHER.CDE     (to read headers for variables)
7 Include file(s) (optional)
```

Output files produced:

```
1 FLSC8101.CRD   (file containing the treatment number, X- and Y-
                 coordinates, and elevations; again this is an ex-
                 ample experiment)
2 SPATSHLL.DAT   (file containing drive, path, experiment name, and
                 column number of the variable used)
```

Images can only be produced for single-year, single-replication simulations. The NYERS and NREPS options within the simulation control section of FILEX should be set to 1 in both cases. If you want to analyze multiple years or replications, you will need to repeat the simulation procedure.

```
*SIMULATION CONTROLS
@N GENERAL   NYERS NREPS START SDATE RSEED SNAME....................
 1 GE            1     1     S 81089  2150 N SPATIAL ANALYSES TEST
```

Note also that, since the software checks for fixed file extensions, the option FNAME within the simulation control section of FILEX should be set to Y in order to create a summary output file with the experimental name attached to it (such as FLSC8101.GSS).

```
@N OUTPUTS FNAME OVVEW SUMRY FROPT GROUT CAOUT WAOUT NIOUT MIOUT DIOUT  LONG
 1 OU          Y     Y     Y     3     Y     N     Y     Y     N     N     Y
```

Calculating Semi-Variances

The first menu contains the following items:

1 Number of lags
2 Lag length
3 Estimation method
4 Missing data flag

The choice of lag number and lag length depend on the dimensions of the field and the data themselves. There is a wide choice, so long as the restrictions mentioned in the section on Semi-variance are borne in mind.

The estimation method options are "Classical" and "Robust". "Classical" is strongly recommended. If a file contains missing values they will be ignored if the appropriate missing data value is specified. The second menu gives an overview of statistical parameters such as mean, standard deviation, number of observations, etc. To continue you have to press <Enter>. Next a screen with the calculated semi-variances as a function of their distance classes (lag length and number) is presented. It allows a visual inspection of the results. In general one has to see whether or not spatial dependence is present (the sill should be higher than the nugget of the semi-variogram) and if the data points show independence at a certain distance (in other words, the sill should be reached). By pressing <Enter> the continuation menu appears. If the result are satisfactory one can leave the semi-variance tool. The results will automatically be written to an output file.

Input files required:

```
1  DSSATPRO.FLE
2  SPATSHLL.DAT
3  FLSC81O1.GSS      (as an example summary output file)
4  FLSC8101.GSX      (as an example FILEX)
5  FLSC8101.CRD      (as an example coordinate file)
```

Output files produced:

```
1  FLSC8101.C##  (as an example file containing the calculated semi-
                 variances; ## refers to the column number in the
                 summary output file)
```

Note the quality of the maps produced at the end of the analysis very much depends on the quality of the semi-variogram. Spending some time in trying out different combinations of lag length and number of lags will generally result in better (more realistic) maps.

Fitting the Semi-Variogram

The next step in the spatial analysis process is to fit a model to the semi-variances. The first menu prompts for the model you want to use. Those available are as follows:

- Spherical;
- Circular;
- Exponential;
- Gaussian;
- Linear;
- Linear with sill;
- Bessel.

Criteria for selecting a particular model were discussed in the section on semi-variance and will not be repeated here. The second menu gives the user the possibility of specifying imposed model parameters. For example, you might want to force the nugget effect (C_0) to 0 or to specify a range of values. If -2.00 is selected no imposed values are assumed.

The next graphical screen displays the previously calculated semi-variances and the fitted model. At the bottom of the screen the input file, output file to which the results are written, and the semi-vario-gram characteristics (nugget, sill, and range), as well as the quality of the fit (SSD/SST), are shown. The ratio, SSD/SST, represents the Sum of Squares of the Differences over the Sum of Squares of the Total. A value of 0 represents a perfect fit, while a value of 1 indicates complete non-correlation. If a fit is satisfactory one can leave the fitting program by pressing the <Esc> key. Pressing any other key will bring you back to the first menu where the different models can be selected.

The printing options in the menu are set up for Epson and compatible printers connected to the LPT1 port of the computer.

Input files required:

```
1  DSSATPRO.FLE
2  SPATSHLL.DAT
3  FLSC8101. F##  (as an example file containing the fitted semi-
                  variogram characteristics; ## refers to the column
                  number in the summary output file.).
4  FLSC8101.CRD
5  FLSC8101.GSS
```

Output files produced:

```
1  FLSC8101.F##   (as an example file containing the fitted semi-
                  variogram characteristics; ## refers to the column
                  number in the summary output file.)
```

Again, note the quality of the maps produced at the end of the analysis very much depends on the quality of the semi-variogram. Spending some time on finding the best semi-variogram model will generally result in better (more realistic) maps.

Kriging

The first menu in the Kriging tool asks the user to specify a cut-off value. This value will be used to calculate the probability maps. Thus if a cut-off value of 7,500 kg is specified when one is analyzing final wheat yields, the probability map will present the probability that the cut-off value of 7,500 kg is exceeded. The other two items (terms in the Hermite expansion, and terms in the Hermite integration) have default values,

and these generally do not have to be changed. In the error and warning section these are discussed further. Increasing the number of terms in the Hermite expansion will generally increase calculation time. On continuing, a menu showing the mean and variance of the input data (simulation results) is presented. Press <Enter> to continue. At this stage some matrix calculations have to be performed that can take a considerable amount of time on computers without a mathematical co-processor.

The next menu presents characteristics of the semi-variogram. These values can be changed here although this is not recommended. The last item deals with the number of neighborhood points which should be included in the kriging calculations. The minimum value is 4, the maximum is 12. The default value is 8. In general, an increased number of neighborhood points will smooth the kriged maps, whereas a low number of neighborhood points can result in "noisy" maps. In choosing this number, the user should consider the objective of the map, gradients in the field, and the quality of the data.

The following menu allows you to specify the type of output produced. The options are as follows:

1 Predictions to individual X- and Y-coordinates. Selecting this menu will bring up a menu in which the X- and Y-coordinates have to be specified. Interpolated values for that particular point are calculated (prediction, variance, and probability) and written to a file, PREDICT.RES. This procedure should be repeated if more interpolations are required.

2 Predictions to a file with specified X- and Y-coordinates. This is basically the same as option 1, but now the X- and Y- coordinates have to be in a coordinate file (with extension .CRD) that can be selected from the spatial data directory. Again the results are written to an output file (PREDICT.RES).

3 Predictions to a Surfer-compatible file format. A second menu is displayed which delineates the area for the kriging. The default values shown within the menu are the minimum and maximum X- and Y-coordinates found in the CRD-file used. These values can be changed. Defining minimum X- and Y-coordinates far outside the range for which the data were obtained will lead to erroneous results because of the effects of extrapolation. The grid cell width is specified at the bottom of the menu. Selecting many, small grid cells will result in long calculation times. In general, it is better to explore your data using a wider grid space, and then use a finer grid size (if necessary) for final presentations. A graphical screen showing the value of the predicted variable will be displayed. Once finished, a bell

sound is produced. When selecting this option, three maps are produced: (1) a map with the predictions (PREDICT.GRD); (2) a map with the variance of the prediction error (VARIANCE.GRD); and (3) a map presenting the conditional probability of exceeding the specified cut-off value (CONDPROB.GRD).

4 Predictions to an IDRISI-compatible file format. This option is very similar to the previous one. One additional item, an image title, is prompted for in a separate menu. The following files are created: PREDICT.IMG, PREDICT.DOC, VARIANCE.IMG, VARIANCE.DOC, CONDPROB.IMG, and CONDPROB.DOC. The data type in all image files is Real, and the file type is ASCII.

5 Predictions to Arc/Info-compatible grid files. This again is comparable to the two previous options. The following files are created: PREDICT.ARC, VARIANCE.ARC, and CONDPROB.ARC. Again, these are ASCII files.

Input files required:

```
1  DSSATPRO.FLE
2  SPATSHLL.DAT
3  FLSC8101.F##   (as an example file containing the fitted semi-
                  variogram characteristics. ## refers to the column
                  number in the summary output file.)
4  FLSC8101.GSS   (as an example of a summary output file).
5  FLSC8101.CRD   (as an example of a coordinate file)
```

Output files produced:

Depending on the output type one has selected, the following files could be produced:

```
1  PREDICT.RES
2  PREDICT.GRD
3  VARIANCE.GRD
4  CONDPROB.GRD
5  PREDICT.IMG
6  PREDICT.DOC
7  VARIANCE.IMG
8  VARIANCE.DOC
9  CONDPROB.IMG
10 CONDPROB.DOC
11 PREDICT.ARC
12 VARIANCE.ARC
13 CONDPROB.ARC
```

Note that, since output is written to files with a fixed name, these files will be overwritten each time the same output options are selected. By using the "Re-name" option explained below, one can easily rename final or intermediate maps.

Mapping

This utility allows some basic GIS operations. It is fully described in Chapter 6 of this document.

File Rename

Output file names associated with the kriging tool are fixed and will be overwritten each time the tool is used with similar options. The "File rename" option allows you to change filenames. When accessed, you first have to select a file from the spatial data sub-directory. The second menu asks you for the new name. If you want to leave the rename option menu without changing the original file name, press <Esc> at this stage.

Exit

The exit option will bring you back to the DSSAT shell.

An Example

To illustrate the geostatistical analysis modules, one example will be discussed. It is described in the file FLSC8101.GSX. It is a hypothetical maize experiment in Florence, South Carolina. In this file, soil information was obtained from 99 locations in the field in a triangular grid. In total, four different soil types are distinguished, each with a specific soil type and set of initial conditions. Two fertilizer levels and two crop residue levels have been specified. Some details of this FILEX are presented below (only 4 of the 99 treatments and field levels are shown):

```
*EXP.DETAILS: FLSC8101GS SPATIAL ANALYSES TEST CASE
*GENERAL
@ADDRESS
 SOUTH CAROLINA
@SITE
 FLORENCE,SC,USA 34.00;0;-99;SC

*TREATMENTS                        -------------FACTOR LEVELS------------
@N R O C TNAME.................... CU FL SA IC MP MI MF MR MC MT ME MH SM
 1 1 0 0 SILTY CLAY                1  1  0  1  1  0  1  1  0  0  0  0  1
 5 1 0 0 SILTY LOAM                1  5  0  2  1  0  1  2  0  0  0  0  1
45 1 0 0 SANDY LOAM                1 45  0  3  1  0  2  2  0  0  0  0  1
61 1 0 0 SAND                      1 61  0  4  1  0  2  1  0  0  0  0  1

*CULTIVARS
@C CR INGENO CNAME
 1 MZ IB0012 PIO 3382
```

```
*FIELDS
@L ID_FIELD WSTA....  FLSA FLOB  FLDT  FLDD  FLDS  FLST SLTX SLDP  ID_SOIL
 1 FLSC0001 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000002
 5 FLSC0001 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000005
45 FLSC0001 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000008
61 FLSC0001 FLSC8101 -99      0 DR000     0     0 00000 -99   151 IB00000011
@L            XCRD            YCRD     ELEV           AREA  SLEN  FLWR  SLAS
 1           5.000           5.000     1.00         0.0000     0   1.0  90.0
 5          45.000           5.000     2.50         0.0000     0   1.0  90.0
45          45.000          45.000     4.50         0.0000     0   1.0  90.0
61           5.000          65.000     1.00         0.0000     0   1.0  90.0
*INITIAL CONDITIONS
@C   PCR ICDAT  ICRT  ICND  ICRN  ICRE
 1    MZ 81089   200   -99  1.00  1.00
@C  ICBL  SH2O  SNH4  SNO3
 1    30 0.380   0.5   0.5
 1    45 0.380   0.5   0.4
 1    60 0.380   0.5   0.4
 1    90 0.380   0.5   0.3
 1   120 0.380   0.5   0.2
 1   150 0.380   0.5   0.1
@C   PCR ICDAT  ICRT  ICND  ICRN  ICRE
 4    MZ 81089   200   -99  1.00  1.00
@C  ICBL  SH2O  SNH4  SNO3
 4    30 0.107   0.2   0.7
 4    45 0.107   0.2   0.6
 4    60 0.107   0.2   0.6
 4    90 0.107   0.2   0.5
 4   120 0.106   0.2   0.3
 4   150 0.106   0.2   0.1

*PLANTING DETAILS
@P PDATE EDATE  PPOP PPOE PLME PLDS PLRS PLRD PLDP  PLWT  PAGE PENV PLPH
 1 81097   -99   7.1  7.1    S    R   50    0  4.0  -99  -99  -99.0 -99.0

*FERTILIZERS (INORGANIC)
@F FDATE  FMCD  FACD  FDEP  FAMN  FAMP  FAMK  FAMC  FAMO  FOCD
 1 81090 FE005   -99    10    66   -99   -99   -99   -99   -99
 1 81139 FE005   -99    10    66   -99   -99   -99   -99   -99
 1 81153 FE005   -99    10    66   -99   -99   -99   -99   -99
 2 81090 FE005   -99    10    40   -99   -99   -99   -99   -99
 2 81139 FE005   -99    10    20   -99   -99   -99   -99   -99
 2 81153 FE005   -99    10    10   -99   -99   -99   -99   -99

*RESIDUES AND OTHER ORGANIC MATERIALS
@R RDATE  RCOD  RAMT  RESN  RESP  RESK  RINP  RDEP
 1 81089 RE001   300  0.50   -99   -99   -99    10
 2 81089 RE001   900  0.50   -99   -99   -99    10
```

```
*SIMULATION CONTROLS
@N GENERAL     NYERS NREPS START SDATE RSEED SNAME....................
 1 GE              1     1     S 81089  2150 N SPATIAL ANALYSES TEST
@N OPTIONS     WATER NITRO SYMBI PHOSP POTAS DISES
 1 OP              Y     Y     Y     N     N     N
@N METHODS     WTHER INCON LIGHT EVAPO INFIL PHOTO
 1 ME              M     M     E     R     S     C
@N MANAGEMENT  PLANT IRRIG FERTI RESID HARVS
 1 MA              R     R     R     R     M
@N OUTPUTS     FNAME OVVEW SUMRY FROPT GROUT CAOUT WAOUT NIOUT MIOUT DIOUT  LONG
 1 OU              Y     Y     Y     3     Y     N     Y     Y     N     N     Y

@  AUTOMATIC MANAGEMENT
@N PLANTING    PFRST PLAST PH2OL PH2OU PH2OD PSTMX PSTMN
 1 PL            155   200    40   100    30    40    10
@N IRRIGATION  IMDEP ITHRL ITHRU IROFF IMETH IRAMT IREFF
 1 IR             30    50   100 GS000 IR001    10  1.00
@N NITROGEN    NMDEP NMTHR NAMNT NCODE NAOFF
 1 NI             30    50    25 FE001 GS000
@N RESIDUES    RIPCN RTIME RIDEP
 1 RE            100     1    20
@N HARVEST     HFRST HLAST HPCNP HPCNR
 1 HA              0   365   100     0
```

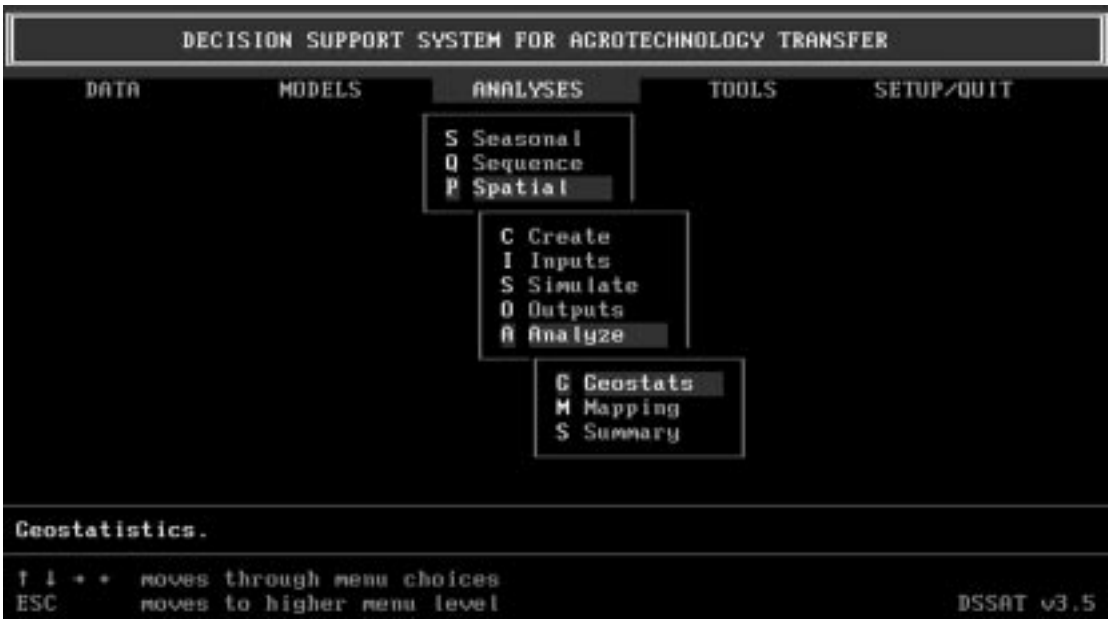
Run the experiment. During simulation with FLSC8101.GSX, the file FLSC8101.GSS will be created and OUTPUT.LST will be updated, if required.

Shell

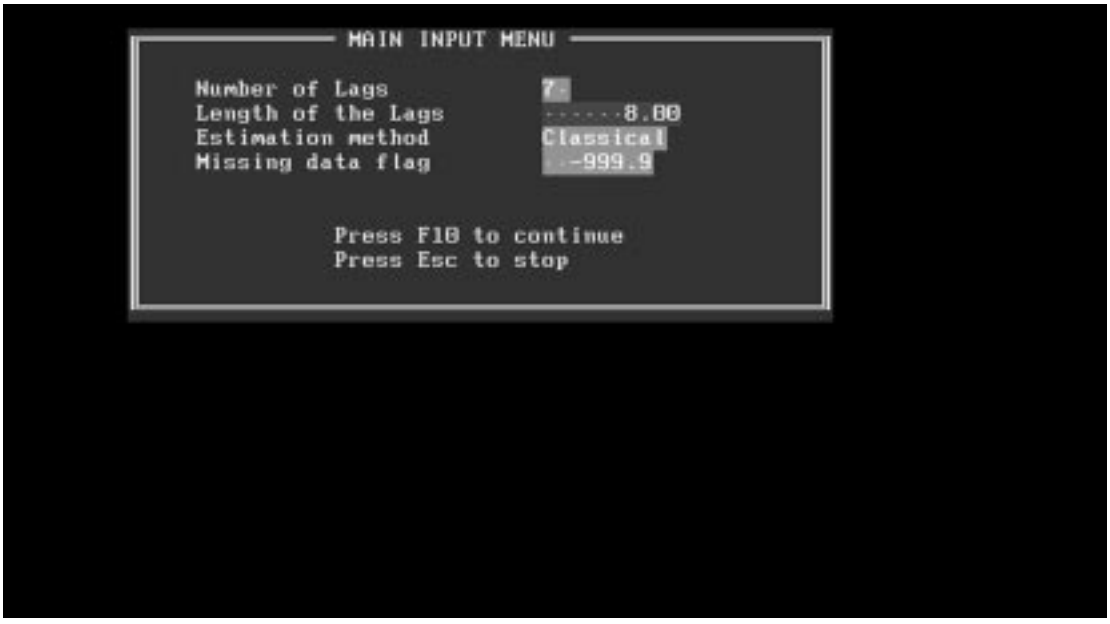
On entering the geostatistics shell, select "FLSC8101" from the output list and "Yield at Maturity" from the variable menu, by using the arrow keys or mouse (see Screen 8). Press <Enter> to make a selection. From the spatial analyses menu select the option "Calculating Semi-variance."

Semi-Variance

For the number of lags enter 10, and for the lag length, 3.00. Press <F10> to continue. Summary statistics are displayed. Press <Enter> to continue. A rather scattered plot is then presented, indicating that the combination of lag length and number was not optimal to describe the spatial dependence of yield at maturity. Press <Enter> to continue and go to the analysis menu to try another combination. By trial and error you will obtain a nice semi-variogram using 7 lags of 8.00 m. This combination is more-or-less within the restrictions mentioned in the section on semi-variance above. Continue and leave the Semi-variance tool (see Screen 9).



Screen 8.



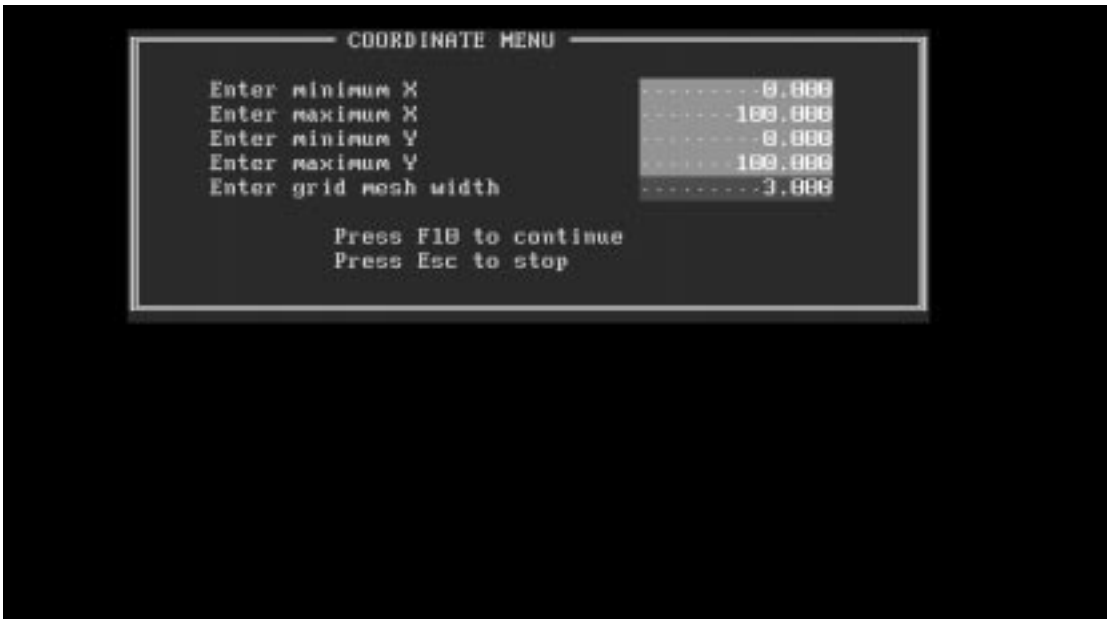
Screen 9.

Fitting the Semi-Variogram

From the "Select model" menu under the "Fitting the semi-variance model" menu, select the spherical model and continue with "No imposed parameter values". A graph displaying the fitted model and the individual data points is shown. Look at the model characteristics (nugget, sill, range, and SSD/SST ratio). Explore other model options to see if a better performer can be found (a model with a lower SSD/SST ratio). Press <Esc> to leave the tool and store the data.

Kriging

Specify a cut-off value of 8,000 kg, leave the other options at their default settings, and continue. The fitted model characteristics are displayed; continue, leaving the neighborhood size set at 8. Choose to predict to an Arc/Info-compatible file. For the minimum and maximum coordinates for X and Y, enter 0 and 100, respectively. You first have to delete the values that are the defaults (see Screen 10). Enter a grid mesh width of 10.000 m (if you have a fast computer you can select a narrower grid mesh width without having to wait too long). Pressing <F10> will start creating the various maps. A beep will sound when the program is finished. You have now created 3 maps, PREDICT.ARC, VARIANCE.ARC, and CONDPROB.ARC, that can be displayed using the mapping module. Try looking at these with "shade information" option 4 in the mapping module set to 6 or so classes (see Chapter 6 also).



Screen 10.

Geostatistical Utility Error and Information Messages

When executing the various options some general or specific error messages and warnings may show up. Most of these are related to assigning files which are not available or not open for output. The geostatistics tools check every file for its presence and status. In the various menus, acceptable ranges for input data are given, as far as is possible. The programs will not accept values outside these ranges. Specific error messages are as follows:

Calculating Semi-Variance

1. If the variance of a selected variable is 0 there is no spatial variability within the data set and execution is stopped.
2. A warning is displayed if the number of data points exceeds 1000.
3. If the number of data points is smaller than 10 there are insufficient data pairs available to calculate the semi-variances.

Fitting

1. If the graphical driver used is not a VGA or Hercules, an error message will be displayed.

Kriging

1. If a zero value in the main diagonal matrix is detected, the matrix operations cannot be performed and program execution will be stopped.
2. A warning message will be displayed if a linear model with sill is used for spatial interpolation.
3. The number of data points **SHOULD NOT EXCEED 375** in this option.
4. If the range of a semi-variogram is 0 the program stops.
5. If the iterative calculation of the Hermite normalization does not converge a warning message is displayed, including a reconstructed value. Increasing the number of terms in the Hermite expansion might lead to a solution.

Suggestions for Further Reading

Applied books

Burrough, P.A. 1986. Principles of geographical information systems for land resources assessment. Oxford, Clarendon, 193 pp.

Davis, J.C. 1986. Statistics and data analysis in geology, 2nd edition, Wiley, New York, 646 pp.

Isaaks, E.H. and R.M. Srivastava. 1989. An introduction to applied geostatistics. Oxford University Press, Oxford.

Webster, R., and M.A. Oliver. 1990. Statistical methods in soil and land resource survey. Oxford University Press, Oxford, 316 pp.

Theoretical books:

Cressie, N.A.C., 1991. Statistics for spatial data, New York, 900 pp.

Journael, A.G. 1989. Fundamentals of geostatistics in five lessons. American Geophysical Union, Washington D.C.

Journael, A. G., and C.J. Huijbregts. 1978. Mining geostatistics. Academic Press,London, 600 pp.

Ripley, B.D. Spatial Statistics. Wiley, New York, 252 pp.

CHAPTER SIX.
THE SGIS MAPPING MODULE

Introduction

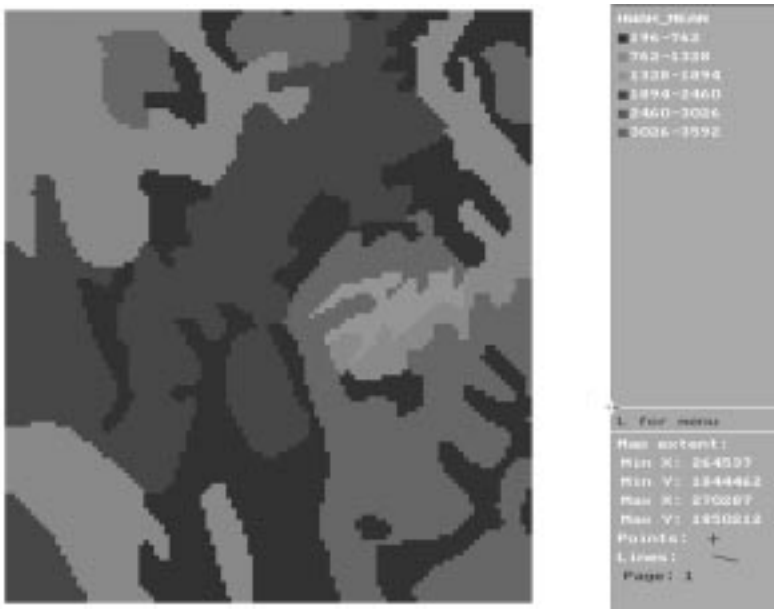
SGIS was developed to allow users to make use of GIS data and view them on their own personal computer. The program is intended for viewing existing maps. It currently supports files created by Arc/Info (exported with <UNGEN> command, only point and line coverages), Surfer (ASCII files), and IDRISI. The program can convert files from one of these formats to any other. To date, SGIS does not allow for printing; it is necessary to use a commercial software package for this purpose. However, if SGIS is started from a Windows environment, one can copy the graphics screen to the clipboard. Other packages will be able to print the contents of the clipboard.

SGIS has been developed for a wide range of different applications with a special focus on tropical environments. Although the software could be made faster, it was implemented to keep hardware requirements modest. SGIS requires a 386 (or higher) computer with a VGA monitor. It exists as one file, SGIS9801.EXE. The program requires the DSSAT setup file DSSATPRO.FLE, and it reads the entry indicating the spatial data directory (see Table 6).

Basic Mapping

SGIS deals with views and maps. A view may comprise a combination of maps, such as a raster map of soil types, a line coverage with roads, and a point coverage with observation points. Currently SGIS allows up to 10 views, and all the settings may be stored. As such, one can easily browse through the different views (see Screen 11).

Maps of all features (points, lines, and polygons) can be linked to database files and lookup tables. The database files and lookup tables should be located in the spatial data directory defined in the DSSATPRO.FLE file. A database file has an extension DAT. Database files typically contain attributes or properties of the different map units, such as the different soil properties of a soil map. Lookup tables have an extension LUT. They contain information on the symbols used for map drawing. Two alternative structures handle continuous variables (using ranges) and discrete variables, respectively. The structure of database files and both types of lookup tables are explained in the section on Formats for Database Files and Lookup Tables.



Screen 11.

After starting SGIS you will be able to draw your maps by using a menu. The different menu items are as follows:

Erase View

This will clear the present view and reset all map settings. It will result in an empty view.

Delete View

This is similar to 'Erase view'. The next view will become the current view.

View Up

Go to the previous view.

View Down

Go to the next view.

Points

Draw a point coverage using the current marker symbol.

Vectors

Draw a line coverage using the current line symbol.

Grid

Draw a grid coverage. After selecting a file, it is possible to enter shade information, such as database files, lookup tables, or the number of equidistant classes. See the section on File Linking for additional information and examples.

Legend

This will draw the legend; press "L" again to return to the menu. Note that all the hot keys for the different menu items remain active when the legend is presented.

Map Extent

When a map is drawn on an empty view, SGIS will determine the proper coordinates for viewing. These coordinates will be presented in the lower right corner of the screen. However, it is possible to enter coordinates manually. Note that SGIS will always adapt the coordinates to a square.

Symbols

Two menus are presented to enter the active marker and line symbols.

ASCII Editor

A simple ASCII editor will be called to view and edit files such as lookup tables and database files. The major hot keys are as follows:

```
Back space: Delete previous character
Del:        Delete character
Ins:        Switch insert: red cursor = insert off, green cursor
            = insert on
Enter:      Cut off present line, copy rest of line to next line
Ctrl+end:   Delete until end of line
Home:       Go to beginning of line
End:        Go to end of line
PgUp:       Page up
PgDn:       Page down
```

Convert

This allows for the conversion of files between the different supported formats.

Info

Gives general information on the author and version number of the program.

Backup Views

A file will be created containing the information on all current views. This will contain the following 5 columns: view number, map name, database file or symbol number, column name (in database file), and lookup table or number of classes. If the database file name, column title, or lookup table are absent, an '#' sign should be included. A test file in distributed with the DSSAT; try calling up the mapping module from DOS with the command "C:\DSSAT35\SGIS9801 views.lst" to operate this.

New Grid

This allows you to create a new IDRISI grid file; the program will prompt for the basic coverage (regardless of what the current view is), and it is then possible to enter shade information, such as database files, lookup tables, or the number of equidistant classes. This provides a means of saving simulation outputs in IDRISI format, which can then be converted using the "Convert" option in the mapping module or manipulated in IDRISI, for example.

Reset

All views are deleted.

Quit

Leave SGIS9801.

File Linking

It is possible to make use of database files and lookup tables. Additionally, one can indicate the number of classes, in which case the range between the minimum and maximum value are sub-divided in equidistant intervals. One can also combine the different options as indicated in Table 5. Each combination is discussed and an example given below.

Table 5. Possible Combinations of Databases, Lookup Tables and Classes.

	Database file	Lookup Table	Number of Classes
Case 1	-	-	-
Case 2	+	-	-
Case 3	-	+	-
Case 4	-	-	+
Case 5	+	+	-
Case 6	+	-	+

The different cases can be illustrated with a simple example. The basic map used in the examples is a simple soil map, which distinguishes four different soil types:

```
1111111222222222233
1111112222222222223    Soil type:
1111122222222222333    1: Aquic Humitropept
1111222222224222333    2: Oxic Humitropept
1112222222244422333    3: Typic Hapludand
1122222222444444333    4: Aquic Hapludand
1222222444444444433
2222224444444444444
2224444444444444444
4444444444444444444    (distributed as file SOILTEST.ARC)
```

Case 1: No database files, lookup tables, or classes defined.

This is the simplest case, where all the soil types are presented by their number. A standard VGA color card allows for 16 colors. The background color is not available, leaving 15 colors to be used. Numbers outside this range will be ignored, i.e. drawn in the background color. This option works only for discrete class variables.

Case 2: Using a database file

For many applications the soil properties are more relevant than soil types. To present soil properties instead of soil types, the soil map can be linked to a database containing these properties in the distributed file called SOILPROP.DAT:

```
@SOIL                               Clay%      Depth    Drainage
* Subgroup Aquic Humitropept
  1                                  20         140        3
* Subgroup Oxic Humitropept
  2                                  60         240        4
* Subgroup Typic Hapludand
  3                                  10         130        4
* Subgroup Aquic Hapludand
  4                                  10         110        3
```

Now the drainage in the area can be presented. The database file is set to SOILPROP.DAT, and the corresponding column with the drainage class can be selected (see Screen 12). The following map will be the result:

```
4444444333333333333
4444443333333333333        Drainage:
4444433333333333333        3: Moderately well drained
4444333333334333333        4: Well drained
4443333333444333333
```

```
44333333444444443333
43333344444444444433
33333344444444444444
33344444444444444444
44444444444444444444
```

Case 3: Using a lookup table

In many cases the colors corresponding to the class variables are not very appropriate. A lookup table allows the definition of user-determined colors. In the example file, the strongly weathered Ultisols could be shaded red, the poorly drained Tropaquept in blue. The structure of the lookup table may be as follows:

```
@Soil            Symbol          (distributed file SOIL.LUT)
    1               3
    2               2
    3               5
    4               4
```

Case 4: Define a number of equidistant classes

For continuous variables, such as clay content and elevation, lookup tables are not appropriate. In such cases one can define a number of classes, and the range between the minimum and maximum value will be divided into this number of equidistant classes. The following map of clay percentage may be presented between 6 classes (leaving the map as it is), 2 classes (grouping 1,2, and 3 and grouping 4,5, and 6), or 1 class (not differentiating the area).

```
11111133333333333311
11111133333333333331   Clay percentage
11111333333333333111   1: 10-26
11113333331333331111   2: 26-43
11133333311113311111   3: 43-60
11333333311111111111
13333331111111111111
33333311111111111111
33311111111111111111
11111111111111111111
```

Case 5: Using a database file and a lookup table

This combines cases 2 and 3 above, where one wants to show a property in user-defined colors. The database file in case 2 can be used in combination with the following lookup table:

```
@Drainage        Symbol          (distributed file DRAIN.LUT)
    3               2
    4               6
```

For continuous variables the lookup table can also indicate ranges and their corresponding symbols. One may select the column for clay in SOILPROP.DAT and use the following lookup table:

```
* Ranges of textures       (distributed file TEXT.LUT)
@Min   Max   Symbol
  0    25      2
 25    50      4
 50   100      6
```

Case 6: Using a database file and a number of classes

This case is similar to case 5, but here a number of classes is indicated instead of the lookup table.

Formats for Database Files and Lookup Tables

The formats of database files and lookup tables are flexible but need to follow a number of basic rules. The files need to be in ASCII format and may be either comma-delimited files or space-delimited files. The line with the column headers needs to start with the character "@". All lines before the header are ignored. In the data file, one can include comment lines, which start with the character "*" or "!".

Database Files

The first column of the database file will be linked to the map inherent in its header. One can select one of the other columns to be presented in the map, based on the column title. Note that if SGIS is started with a batch file containing view information, the header titles are not case sensitive. The database file may, for example, be used to present soil properties instead of soil types.

Lookup Table

The basic file structure of a lookup table is similar to the database files. However, the number of columns is limited. The basic structure contains two columns, one with the value (an integer or real variable) and one with the symbol (an integer variable). An alternative structure contains three columns with a minimum, maximum, and symbol column, respectively. The minimum and maximum values indicate a range to which a specific symbol is assigned.

CHAPTER SEVEN.
SYSTEM REQUIREMENTS

For all the components of the DSSAT spatial analysis to run, the usable portion of your computer system's 640 Kbyte RAM should be at least 580 Kbyte in size. All program executable files are located by default in the C:\DSSAT35 subdirectory.

For example, the program SPAXUTE.EXE requires about 250 Kbyte of free RAM to run; this is run under the DSSAT shell, which needs another 70 Kbyte, so if there is to be room for editing files from the program or for calling the mapping module (which requires another 110 Kbyte or so), it is best to have as much free RAM as you can manage. Similarly for the spatial model driver MAPDRV; it needs about 100 Kbyte of RAM, so add in the shell (70 Kbyte) and CROPGRO (360 Kbyte), and this means a total of at least 530 Kbyte of RAM is required. As for the other model drivers, the inputs program and the crop model itself are called sequentially.

The spatial analysis programs have been tested on 386-based machines with VGA graphics, and function without problems. Running many simulations without a mathematics coprocessor is possible but not recommended, as the time taken to complete complex, replicated experiments may be inordinate. A Pentium, Pentium II, Pentium III[1] processor is preferable.

In terms of the program executable files needed for spatial analysis, the names of many of these are defined in the DSSAT setup file DSSATPRO.FLE, accessible from the DSSAT shell using the SETUP options. Alternatively, this file can be edited directly. The slots, programs and directories used by the spatial programs are shown in Table 6.

If users have a copy of a GIS software package such as IDRISI, it can be set up to access the spatial data directory; the user might then use the IDRISI COLOR module to map treatment coverages and model output images.

[1] Pentium and Pentium II are products of Intel and reference to both do not imply endorsement of these products by the authors, editor, or publisher.

TABLE 6. SLOTS IN THE DSSAT SETUP FILE DSSATPRO.FLE, PROGRAMS, AND DATA DIRECTORIES ACCESSED BY THE SPATIAL ANALYSIS SOFTWARE.

Code	Program Name	Setup Description
ASP	XCREATE.EXE	Spatial – Create
ASX	SPAXUTE.EXE	Spatial – FILEX Utility
APS	MAPDRV.EXE	Spatial – Simulation
APV	SPATSHLL.EXE[1]	Spatial – Geostatistics
APM	MAPEXP.EXE	Spatial – Mapping
APA	VARAN2.EXE	Spatial – Summary
APJ	SGIS9801.EXE	Spatial – Mapping Module
APD	-	Internal spatial data directory
TOE	-	The editor specified by the user
ECD	-	The directory where economic data are stored
Mcc[2]	-	Names of model executable files specified in SETUP
-	EFM.EXE	Spatial – List Inputs
-	GFM.EXE	Spatial – List Outputs

[1] This program calls SPAT9801.EXE, DISC9801.EXE, and WLS9801.EXE, which should all be in the same directory as SPATSHLL.EXE.

[2] "cc" refers to the crop code (such as "MZ" for maize).

REFERENCES

Booltink, H.W.G., and J Verhagen. 1997. Using decision support systems to optimize barley management on spatial variable soils. pp. 219-233 in: M.J. Koppf, P.S. Teng, P.K. Aggarwal, J. Bouma, B.A.M. Bouman, J.W. Jones, and H.H. van Laar (eds), Applications of Systems Approaches at the Field Level. Volume 2. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Cressie, N.A.C. 1991. Statistics for spatial data. New York, 900 pp.

Eastman, J.R. 1993. IDRISI version 4.1 Technical Reference Manual. Clark University, Worcester, MA, USA.

Engel, T., G. Hoogenboom, J.W. Jones, and P.W. Wilkens. 1997. An agricultural and environmental geographic information system for the spatial application of crop simulation models. Agronomy Journal 89, 919-928.

Engel, T., J.W. Jones, G. Hoogenboom, and P.W. Wilkens. 1999. Agricultural and Environmental Geographic Information System for Windows. p. 103-146 in: G. Hoogenboom, P.W. Wilkens, and G.Y. Tsuji (eds), DSSAT v3, volume 4. University of Hawaii, Honolulu, Hawaii, USA.

ESRI (Environmental System Research Institute), 1992. PC Arc/Info Reference Manuals, Version 3.4D+. Environmental System Research Institute, Redlands, California, USA.

Finke, P.A., 1993. Field scale variability of soil structure and its impact on crop growth and nitrate leaching in the analysis of fertilizing scenarios. Geoderma 60, 89-107.

Finke, P.A. and A. Stein. 1994. Application of disjunctive co-kriging to optimize fertilizer additions on a field scale. Geoderma 62, 247-263.

Golden Software, Inc. 1989. SURFER Version 4 Reference Manual. Golden Software, Inc., California, USA.

Heuvelink, G.B.M. 1995. WLSFIT manual. Landscape and Environmental Group, University of Amsterdam, The Netherlands.

Hoogenboom, G., and D.D. Gresham. 1995. Integration of crop simulation models and geographic information systems for farm-level applications. p. 22-23 *in* P. K. Thornton *et al.* (eds), Applying Crop Models and Decision Support Systems. Special Publication SP-22, IFDC, Muscle Shoals, Alabama, USA.

Negabhan, B., C. Fonyo, W. Boggess, J.W. Jones, K. Campbell, G. Kiker, E. Hamouda, E. Flaig, and H. Lal. 1994. LOADSS: A GIS-based decision support system for regional environmental planning. Florida Agricultural Experiment Station Journal Series.

Singh, U., J.E. Brink, P.K. Thornton, and C.B. Christianson. 1993. Linking Crop Models With a Geographic Information System to Assist Decision-making: A Prototype for the Indian Semiarid Tropics. IFDC Paper Series P-19, Muscle Shoals, Alabama, USA.

Staritsky, I. G., P.H.M. Sloot, and A. Stein. 1992. Spatial variability and sampling of cyanide polluted soil on former galvanic factory premises. Water, Air and Soil Pollution 61, 1-16.

Stein, A. 1993. Introduction to geostatistics. Department of Soil Science and Geology, Wageningen Agricultural University, The Netherlands, 40 pp.

Stoorvogel, J. J. 1995. Linking GIS and models: structures and operationalization for a Costa Rican case study. Netherlands Journal of Agricultural Science 43, 19-29.

Thornton, P.K., H.W.G. Booltink, and J.J. Stoorvogel. 1997a. A computer program for geostatistical and spatial analysis of crop model output. Agronomy Journal 89, 620-627.

Thornton, P.K., W.T. Bowen, A.C. Ravelo, P.W. Wilkens, G. Farmer, J. Brock, and J.E Brink, 1997b. Estimating millet production for famine early warning: an application of crop simulation modelling using satellite and ground-based data in Burkina Faso. Agricultural and Forest Meteorology 83, 95-112.

Tsuji, G.Y., G. Uehara, and S. Balas S (eds), 1994. DSSAT v3 (3 volumes). University of Hawaii, Honolulu, Hawaii, USA.

Webster, R. 1985. Quantitative spatial analysis of soil in the field. pp. 1-70 in B A Stewart (ed), Advances in Soil Science, volume 3. Springer Verlag, New York.

Volume 4-3

Agricultural and Environmental Geographic Information System for Windows

T. Engel
J.W. Jones
G. Hoogenboom
P.W. Wilkens

Technical University of Munich
University of Florida
University of Georgia
International Fertilizer Development Center
International Consortium for Agricultural Systems Applications

ACKNOWLEDGMENTS

The development of this software was partially supported by the German Research Association (DFG), the Technical University of Munich and the Agricultural and Biological Engineering Department of the University of Florida.

CHAPTER ONE. INTRODUCTION

Computer-based biological simulation models are used to predict the effects of different management practices and land use systems on agricultural production as well as on the environment. AEGIS/WIN was developed for the visualization of the simulation results using farm or region maps to help improve presentation and understanding.

AEGIS/WIN stands for *A*gricultural and *E*nvironmental *G*eographic *I*nformation *S*ystem for Microsoft *Win*dows. It is an agricultural Decision Support System, that links a Geographic Information System (GIS) with the *D*ecision *S*upport *S*ystem for *A*grotechnology *T*ransfer DSSAT (Tsuji *et al.*, 1994). It runs on IBM compatible Personal Computers using the GIS software ArcView 3.0 or 3.1 (ESRI, 1994a) and has a completely mouse and menu driven user interface. AEGIS/WIN is written using the macro programming language Avenue of ArcView (ESRI, 1994b). Additional examples for AEGIS/WIN can be found in Engel et al. (1997).

AEGIS/WIN can be used for any region or so-called *coverage,* e.g., a field on a farm, region, or country scale. For each coverage, different crop management practices can be assigned to different fields or homogenous land areas, so-called *polygons.* AEGIS/WIN allows one to evaluate the effects of various practices in order to estimate crop production, to explore (optimal) land use options, and to determine some potential environmental impacts. Roughly, the user interface can be used to perform the following functions:

- Create thematic maps based on (i) polygon-characteristic features of a coverage such as soil type, total area, weather station, and climate zones, or (ii) landuse-characteristic features such as the crop and variety;
- Define and modify various crop management practices using DSSAT. These practices include crop, variety, planting management, irrigation management, fertilization management, and/or residue management;
- Perform simulations and examine results of several simulated factors. Simulation results can be graphically presented in thematic maps. Statistical results tables can be created and shown using graphs;
- Maps, graphs and tables can be printed.

AEGIS/WIN uses the crop models of DSSAT which are written in Fortran, according to the standard crop model inputs and outputs as defined by the International Benchmark Sites Network for Agrotechnology Transfer Project (Tsuji *et al.,* 1994; Jones *et al.,*

1994; Hoogenboom *et al.,* 1994). AEGIS/WIN has been developed so that it is able to use the various crop simulation models that are included in DSSAT.

This chapter provides information to on how to use AEGIS/WIN. Furthermore, it provides technical information concerning the structure of the INFO databases, the location of files, and some key programming concepts.

It is recommended that the user runs AEGIS/WIN while reading this manual, as many of the instructions and mouse/menu interactions become much clearer while viewing AEGIS/WIN on the computer screen. It is assumed users have some basic knowledge of the Windows95 and Windows98 environment. Additionally, the user should be very familiar with the DSSAT in order to able to create input files and experiment lists. It will also be helpful, if the user is familiar with the file structure of the crop management file FILEX (Hoogenboom *et al.,* 1994; Jones *et al.,* 1994). The user should also have basic knowledge about GIS; at a minimum, the installation and operation of ArcView (ESRI, 1994a).

CHAPTER TWO.
RUNNING AEGIS/WIN

INSTALLING AEGIS/WIN

AEGIS/WIN is installed as part of the DSSAT v3.5 software package.

1. Install DSSAT v3.5 and ArcView 3.0 or 3.1 before you run AEGIS/WIN. Read the corresponding instruction manuals for the installation of these programs. Test whether both programs run without any problems before running AEGIS/WIN.

2. Make sure that the file 'autoexec.bat' contains the set command for the DSSAT directory, e.g.

   SET DSSAT3=C:\DSSAT35

3. Add a new set command to the file 'autoexec.bat' for the directory where you have installed AEGIS/WIN, e.g.

   SET AEGISWIN=C:\AEGISWIN

4. Go to the *Program Manager of MS-Windows* and open the 'Properties...' from the menu 'File' for the *ArcView* icon. Change the working directory to the directory which you have created for AEGIS/WIN.

5. ARC/INFO coverages have to be stored as subdirectories of the AEGIS/WIN directory. For each coverage you have to create a subdirectory called 'DATA' which is used for all input and output files for simulations corresponding to this coverage. If the subdirectory does not exist, it is automatically created. One sample coverage and sample data is installed with AEGIS/WIN.

RUNNING AEGIS/WIN

1. Start ArcView 3.0 or 3.1

2. Use the dialog 'Open Project...' from the File-Menu to load the project 'aegiswin.apr', which is located in the AEGIS/WIN directory.

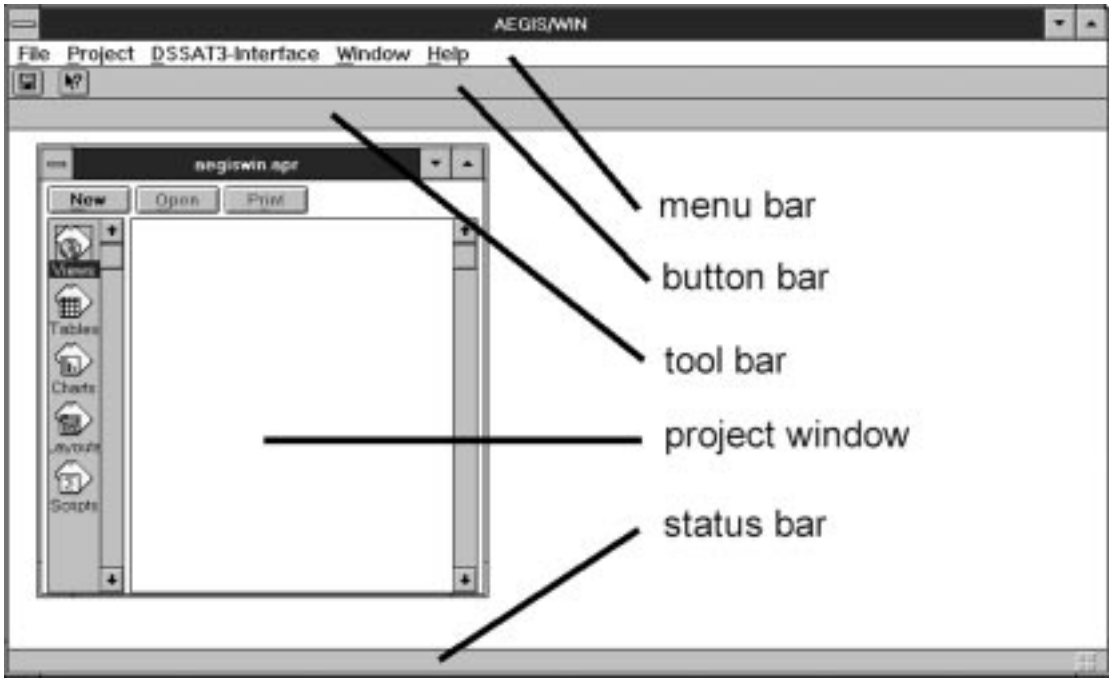
CHAPTER THREE.
THE USER INTERFACE

INTRODUCTION

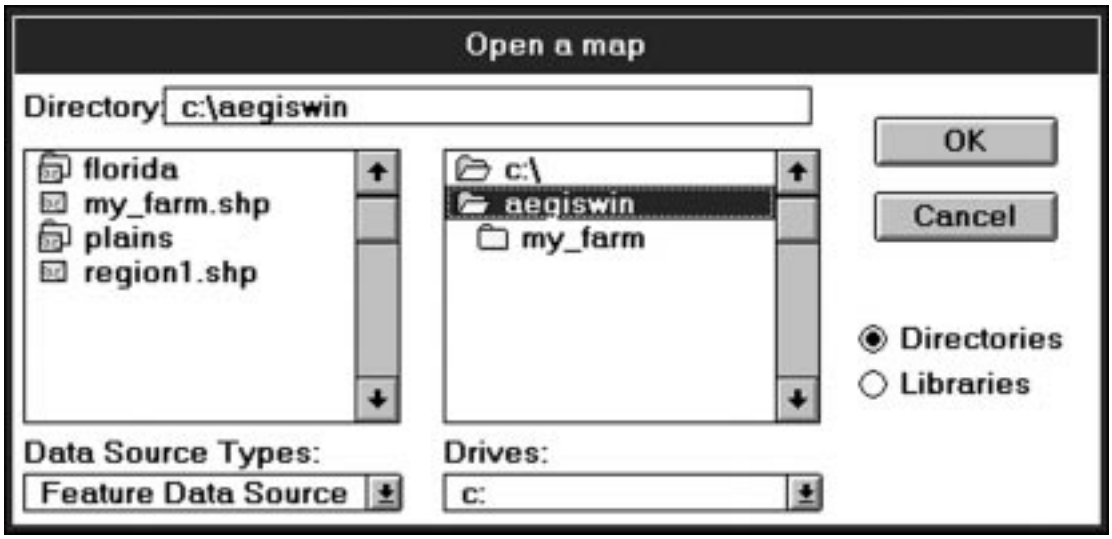
After loading 'aegiswin.apr' you obtain the window as shown in Screen 1. The AEGIS/WIN interface is similar to the ArcView interface and is made up of windows, menus, a button bar, a tool bar, and a status bar. Screen 2 shows the ArcView application window in which AEGIS/WIN runs. It serves as the framework within which you perform all operations. The Project window is titled 'aegiswin.apr' and organizes and lists the contents of AEGIS/WIN; it makes it easy to manage your work.

The *menu bar* is located at the top of the window and provides AEGIS/WIN command choices in pulldown menus. Click on each menu name to see the menu choices. If you compare the window with the original ArcView window, you will see that the Menu 'DSSAT3-Interface' was added to the *Menu Bar*.

The *button bar* is right under the menu bar. *Buttons* on the button bar are shortcuts for quickly accessing commonly used commands. On the display, the button bar



SCREEN 1.



Screen 2.

contains only two buttons, one to save the current project and one to activate the ArcView helping mode.

The *tool bar* is the second bar under the menu bar. Icons on the tool bar represent tools that allow you to modify the action of the cursor on the screen display. The tool bar in screen 1 does not contain any messages.

The *status bar* appears at the bottom of the AEGIS/WIN window. During some ArcView operations it displays a progress bar. It also displays a one-line description of the operation of a menu choice, button or tool. The status bar of screen 1 does not contain any message.

The views, tables, charts, layouts, and scripts you create in AEGIS/WIN are called documents. Menu, button and tool bars change according to what type of document is active.

A view is an interactive map that lets you display, explore, query, and analyze geographic data. A view defines how to display the geographic data you use, but it doesn't contain the geographic data files themselves.

A table stores tabular data. You can display, query, and analyze almost any kind of tabular data, such as attributes of maps, field names, crops, etc. Charts allow you to present tabular information in graphic form. Layouts allow you to place all types of documents in one window to produce a final map. Views, tables, and charts can be

referenced (instead of directly copied) so that updates are automatically reflected in the layout. You can add map elements, such as a title, legend, scale bar, and north arrow to a layout. A script contains Avenue code and is only part of your project, if you have installed Avenue on your system. It is not necessary to have Avenue for operation of AEGIS/WIN.

To list the documents contained in the project, click on each icon in the Project window: Views, Tables, Charts, Layouts and Scripts (if you have Avenue). As you click on each icon, AEGIS/WIN lists the documents available with the project of that type.

The Menu 'File' allows the handling (open, close, save, save as..) of projects. When you save a project, the links to all opened maps, tables, etc. are saved with theproject. The project 'aegiswin.apr' is write protected. Because of this, if you want to save your modifications or additions to the project, like maps or simulation result tables, use the command 'Save Project As...' and save the project under a new name—for example, 'aegiswin.rev.' The next time you open the new project, all linked maps and tables are opened again and you can access the entire functionality of AEGIS/WIN. If you want to start a new project without any opened maps or tables use 'aegiswin.apr'.

The Menu 'DSSAT3-Interface' contains the entries 'Open existing map', 'Create new map', 'Start DSSAT-Shell' and 'Run Simulation'.

In the beginning, only the first two entries are available, because the existence of a map is the requirement for all further operations. When a map is opened or created, DSSAT and its crop models are linked to the data for that map area

Maps in AEGIS/WIN

Maps show spatial data that store the geometric location of geographic features along with attribute information describing what these features represent. The locational data are stored in a vector or raster data structure, and corresponding attribute data are stored in a set of tables related geographically to the features they describe, known also as a georelational data structure.

Spatial data are central to ArcView. Using ArcView typically involves displaying spatial data, working with its existing attributes, adding your own tabular attributes to it, and using it as the basis for geocoding other data. AEGIS/WIN supports **ARC/INFO coverages** (ESRI, 1993) and **ArcView shapefiles** (ESRI, 1994a) as spatial data formats.

ARC/INFO coverages are topological data structures for geographic features. The coverage format is suitable for spatial analysis and large geographic data management applications.

The ARC/INFO coverage format is one of the most popular and widely available spatial data formats found in digital mapping and GIS applications. ARC/INFO coverages can be created using ARC/INFO, e.g., by digitizing or overlaying existing maps. ArcView does not support the creation of this format, but can use it. The user needs an external GIS program or source to convert his or her coverages into ARC/INFO format.

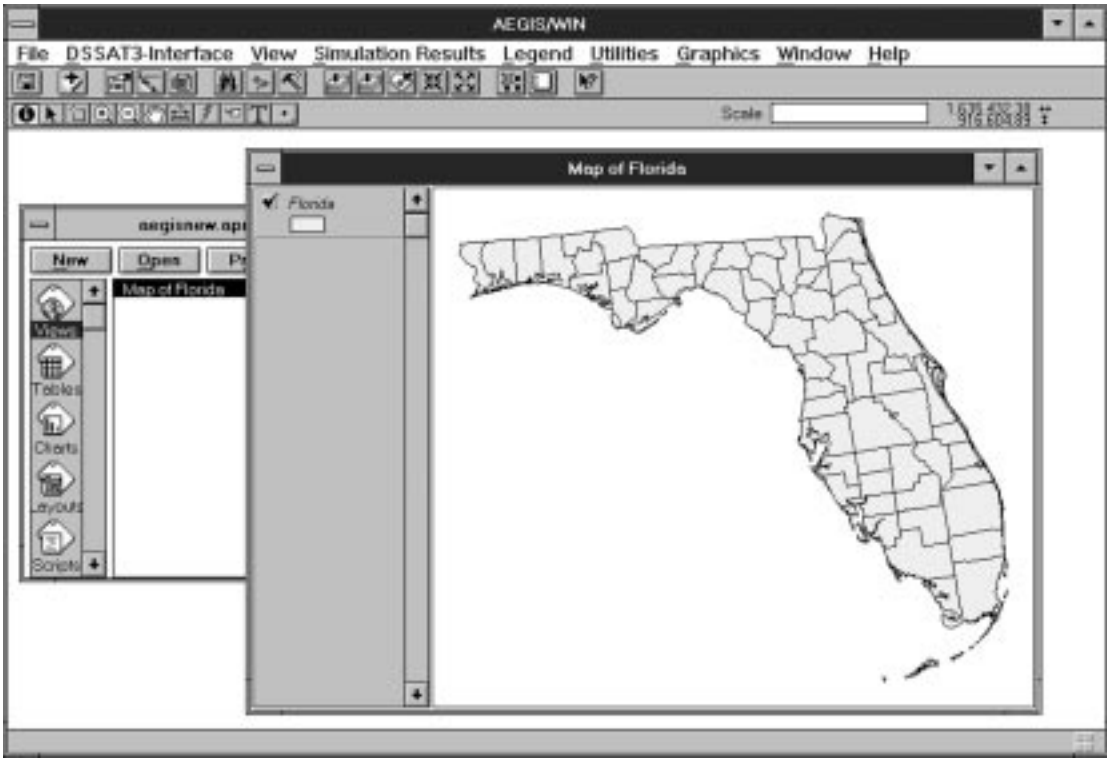
ArcView shapefiles are simple, non-topological formats for storing the geometric location and attribute information of geographic features. Shapefile features display more rapidly on a view and can be used to edit the geometry of features and their attributes, and to add and delete features of a map based on a shapefile. The most important advantage of shapefiles is that you can create your own new map, e.g., of your farm or of a watershed, that's based on this format.

To work with AEGIS/WIN, the first step is to open a map. Open either an existing map based on an ARC/INFO coverage or an ArcView shapefile, or create (draw) your own new map based on an ArcView shapefile format.

Use of Existing Maps

Open an existing map by clicking the [New] button in the project window or by choosing 'Open existing map' in the menu 'DSSAT3-Interface'. The dialog that appears gives a list of all available maps in AEGIS/WIN (see screen 2). The list contains ARC/INFO coverages and ArcView shapefiles. ARC/INFO coverages (ESRI, 1993) are recognized by the directory symbol in front of the name and by the lack of an extension, e.g., Florida, Plains. ArcView shapefiles (ESRI, 1994a) have a file symbol in front of the name and the extension '.shp' (e.g. my_farm.shp, region1.shp).

If the coverage is successfully selected, e.g., Florida, a base map is displayed showing the contours of the coverage and its polygons (see screen 3). The menu bar is changed, so that the user can now start the DSSAT-Shell, run simulations, join tables and show simulation results in the map. The button and tool bars contain many icons to modify the map (for a description of the operations of these features, see the ArcView 3.0/3.1 manual or the ArcView online help). The view list in the project window now contains a single entry 'Map of Florida'.



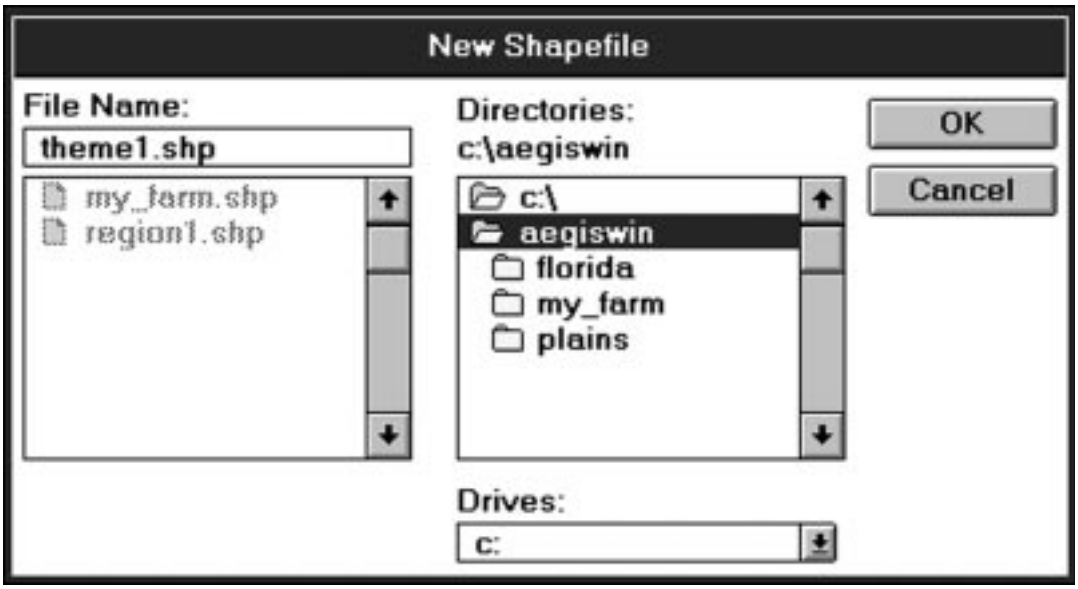
Screen 3.

Creation of New Maps

In addition to using existing maps, you can also create a new map using 'Create new map' from the menu 'DSSAT3-Interface' when the project window is your active window. When another map is already opened and this map is your active document, 'Create new map' is located in the menu 'View'. In working with AEGIS/WIN, note that the menu bar is different depending on the type of the active document. These different menu bars can be confusing initially. Remember to use only the menus that contain commands for the type of the active document.

To create a new map, enter the name of the new shapefile at the prompt. Screen 4 shows a listing of existing shapefiles and the name of the new shapefile theme1.shp. New maps have to use an ArcView shapefile for the storage of geographic information. If you select the name of an existing shapefile, this shapefile will be overwritten.

An image, e.g., a scanned map or a satellite picture, can be used as a background for drawing polygons or tracing boundaries of fields. To open an image, answer 'Yes' at the



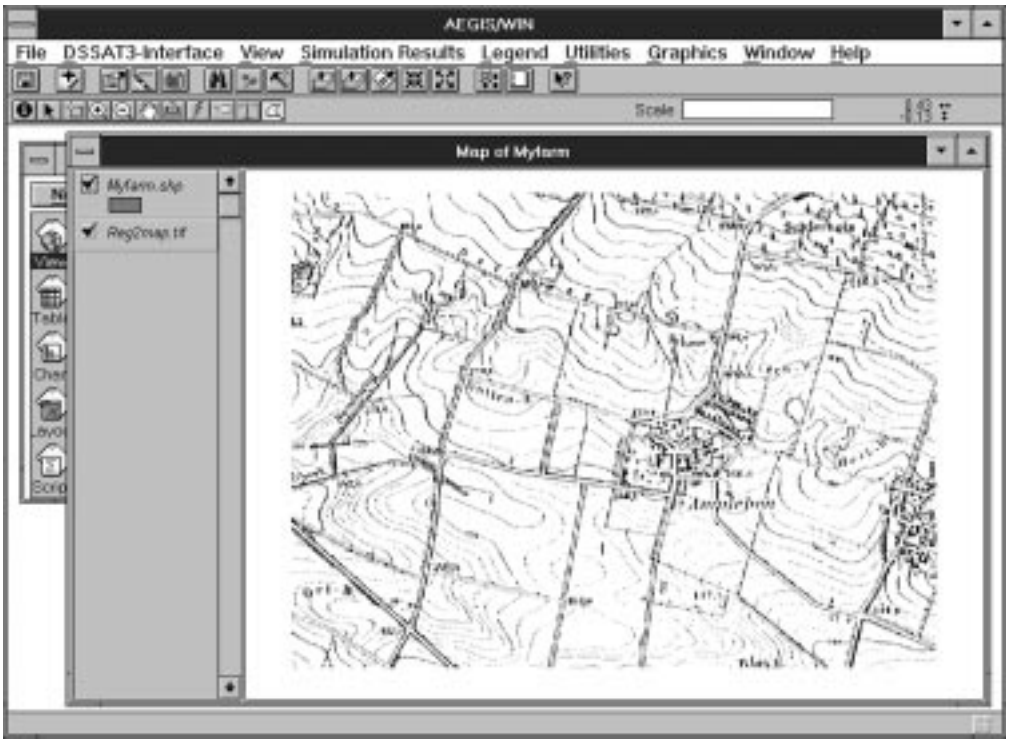
Screen 4.

prompt and the corresponding dialog will show a list of all image files that are available in the AEGIS/WIN directory.

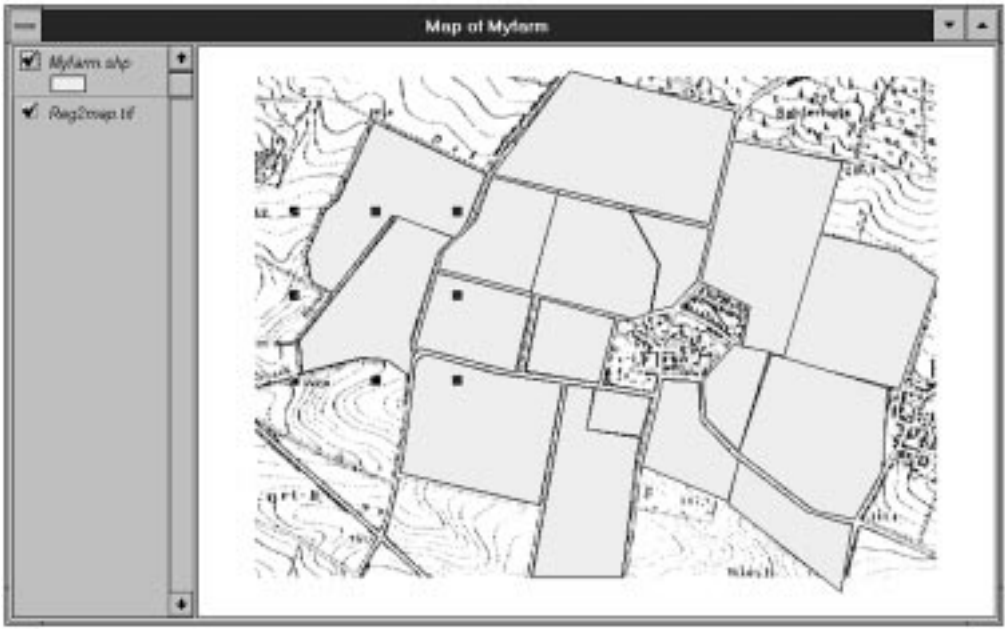
AEGIS/WIN supports different image formats, including TIFF. For more information about the supported formats see the ArcView documentation or the online help. If you have access to image files that are supported by ArcView, copy them to the AEGIS/WIN-directory. AEGIS/WIN contains 'farm_map.tif' as a sample file. After selecting an image, the opened window contains the image and your shapefile, which is still empty (see screen 5). You can now draw polygons using the different drawing tools of the tool bar (for more information see the ArcView manuals or the online help).

After drawing the polygons, your map will look similar to screen 6. Your shapefile is in the foreground with the scanned map in the background. Each map has an attribute table, which stores the geometric features. This newly created map now has an attribute table, which can be accessed by opening the 'Open Attribute Table' option from the menu 'View'.

A record is automatically added to the attribute table for each polygon you draw. Before you can run the simulation, you must fill the attribute table with information on the number of the polygon, the corresponding weather file name and soil code. If you know the area of your fields, you should enter it too. AEGIS/WIN



Screen 5.

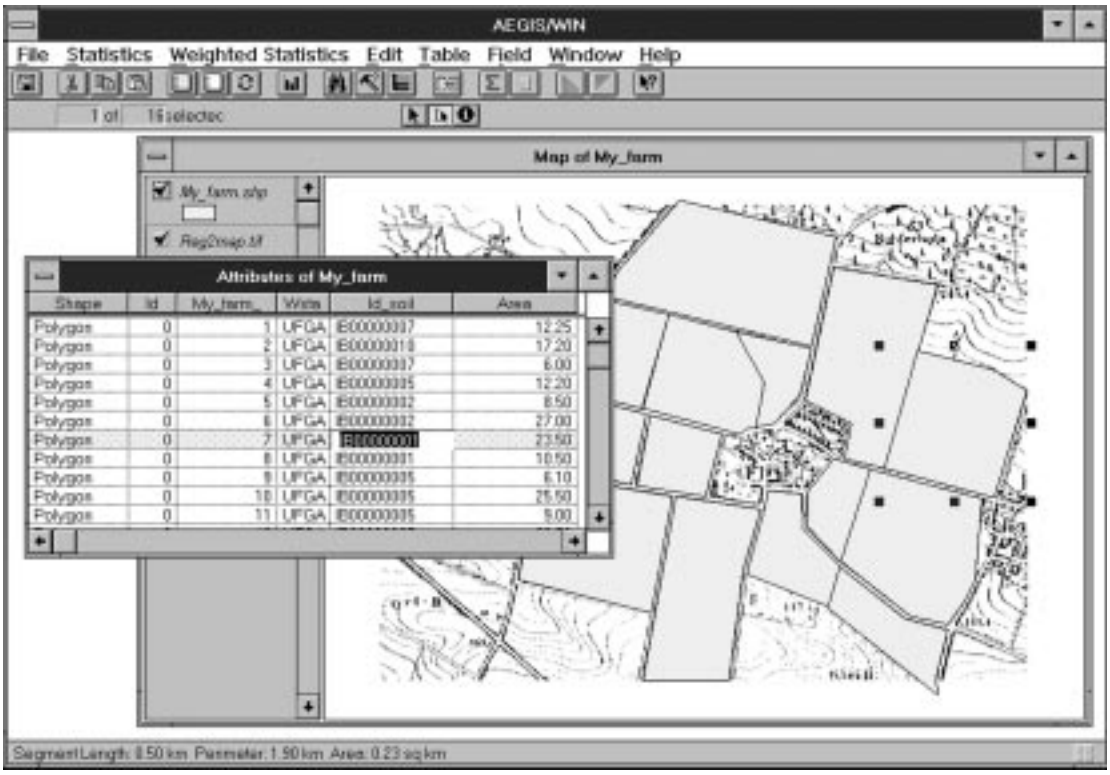


Screen 6.

needs the area for the calculation of weighted statistics of the simulation results. More detailed information about the content of the attribute table is given in the section entitled "directory structures and files" in chapter 4. Information can be added to the table only when 'Start editing/Stop editing' of the menu 'Table' is active The selected record is highlighted in the table and at the same time in the map (see screen 7).

Modification of Existing Maps

AEGIS/WIN allows the modification of existing maps. All necessary commands are included in the menu 'View'. The toolbar provides tools for changing the size of polygons (see ArcView manuals or online help). ArcView shapefiles can be modified. However, ARC/INFO coverages can not be changed. If modifications are to be made of existing ARC/INFO coverages, use ARC/INFO or convert it to an ArcView shapefile using the corresponding command in the menu 'View'. The new shapefile has the same view and the same polygon attribute table of the original ARC/INFO coverage. It is important to remember that there is no change of the attributes in the table when

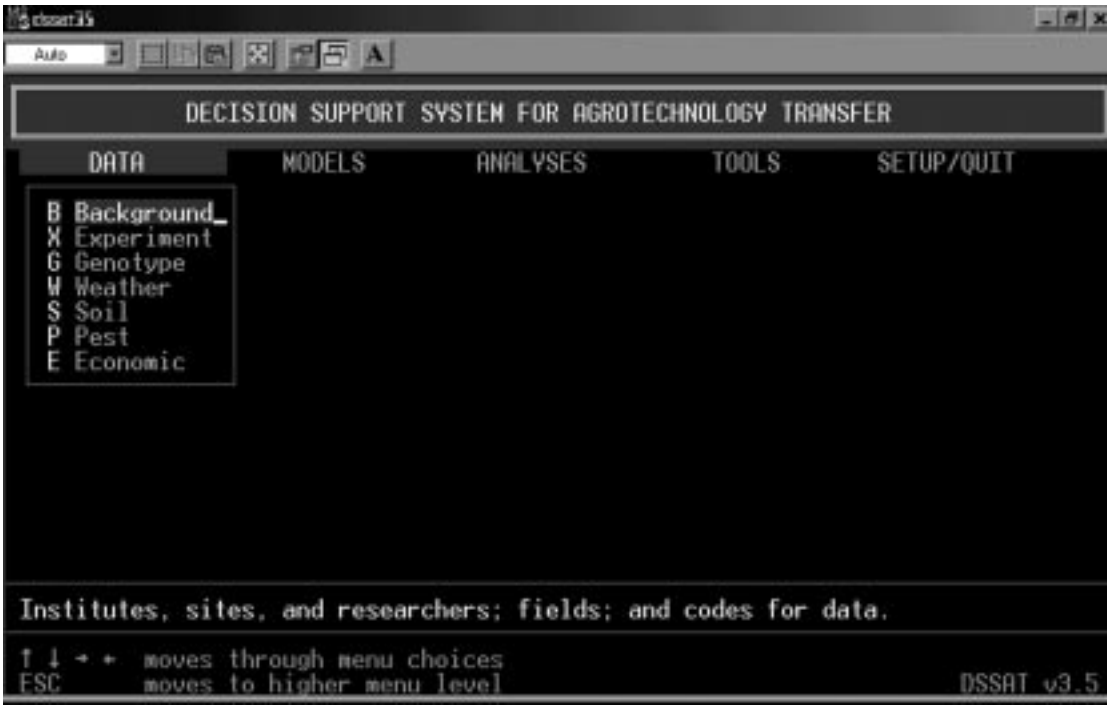


Screen 7.

you change the size of the polygon on the map. The fields with the geometric content, like area and perimeter, are not adjusted!

The menu 'View' has the following commands for the modification of maps:

View Properties	Opens properties of a view. The dialog box allows the setting of the *distance* and the *map units* of the view. As result, the actual scale is visible in the scale window of the tool bar. The units are used for the dimensions in the status bar for new polygons (for more information see ArcView online help for 'units'). When a layout to print the map is created, these units are used for the scale bar. The default values are unknown units.
Theme Properties	Opens properties of a theme (defined as a set of geographic features in a view). A dialog box appears and is used to turn on the snapping mode when maps are drawn. (see ArcView online help for 'snap tool')
Convert to shapefile	Converts existing ARC/INFO coverage into ArcView shapefile. Use the same name for the shapefile. An extension '.shp' is automatically added. If the same name is used, access to the same simulation data used for the ARC/INFO coverage is immediate.
Start editing/Stop editing	Makes a map based on an ArcView shapefile. It allows you to edit a shapefile, or it disables editing and saves changes for the map when editing is finished. A dashed line around the theme's check box in the Table of Contents indicates that editing has been allowed for that map.
Copy graphics	Copies the selected graphic or graphics to the system clipboard
Delete graphics	Deletes the selected graphic(s)
Paste graphics	Paste the data that have been copied onto the clipboard
Add Image Source	Adds an image to the existing view
Add Feature Source	Adds another map to the existing view



Screen 8.

Open Attribute Table — Opens the corresponding attribute table of the map. In order to edit the table, select 'Start editing' from the menu 'Table'.

Creating and Modifying Field Management Files

Use of DSSAT

Before you can show maps with simulation results, you have to create input files and you have to run simulations. AEGIS/WIN uses the DSSAT-Shell for the creation of the input-files. Choose 'Start DSSAT-Shell' from the 'DSSAT3-Interface' menu to access DSSAT v3.5. You can use the DSSAT-Shell only for the creation or modification of the input files, called X-files, and for the creation of experiment lists. This capability is provided in AEGIS/WIN under the menu 'DSSAT3Interface'. When you start the DSSAT-Shell, the DSSAT v3.5 menu appears (screen 8).

If you want to work with the shell in full screen mode instead of windows mode, you can use the keys [Alt] + [Enter] to switch between options.

AEGIS/WIN uses the 'Spatial' option of DSSAT v3.5 for the creation and administration of the data corresponding to the map. When you call the shell, the path for the

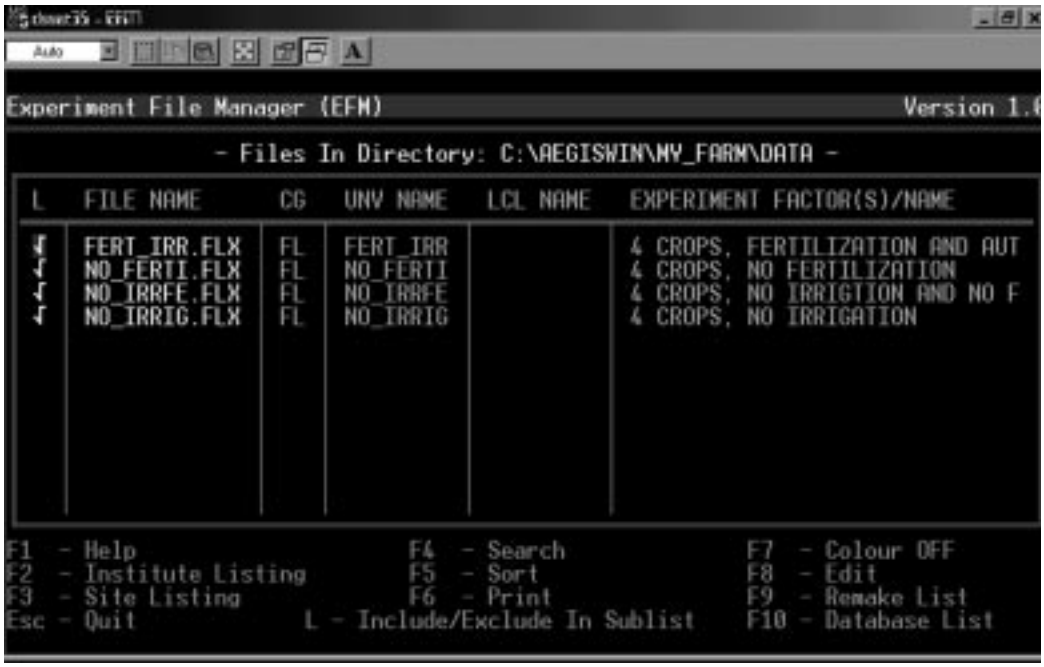
'AEGIS/WIN' data is automatically adjusted to the subdirectory '..\data' of the directory of the current map. Go to 'Spatial' of the menu 'ANALYSES' to access the data (see screen 9). The submenu has the elements 'Create', 'Input', 'Simulate', 'Outputs' and 'Analyze', but only the first two options are used by AEGIS/WIN. 'Create' allows you to start the XCREATE-program to create an input-file (see DSSAT, volume 3) for the simulation.

Select 'Input' to obtain the Experiment File Manager of DSSAT v3.5, which lists all available input files in the current directory (see screen 10) shown at the top of the window. *[F8]* opens an editor to edit a file in the list and *[F9]* creates an actual experiment list. Soil and weather files should also be located in the current directory. For more details read volume 2 of the DSSAT manuals. Important: it is necessary to create an experiment list *[F9]* from the 'INPUT' option under DSSAT v3.5 for running the simulation, even if there is only one file in the directory.

Use of ID_FIELD as Key for Linking with Map

To be able to join simulation results to a map, it is necessary to use the same key number in the polygon attribute table of the map and in the management input file

Screen 9.



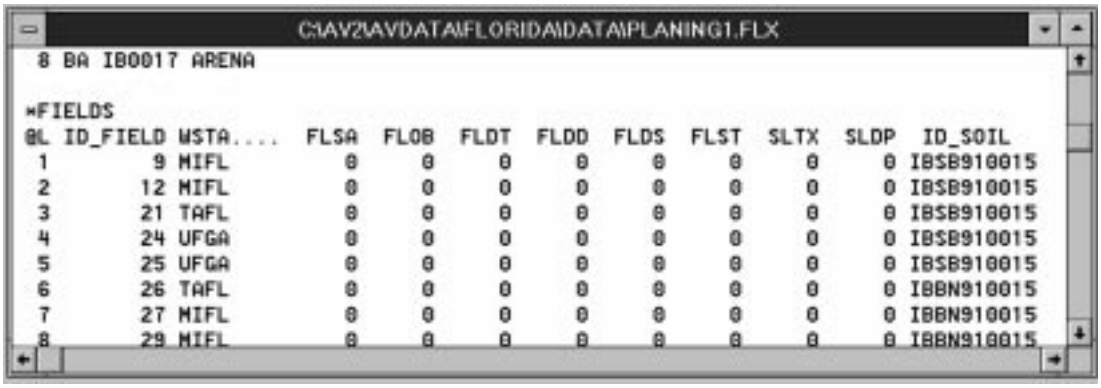
SCREEN 10.

FILEX. For a detailed description of the structure of the attribute tables of ARC/INFO coverages and ArcView shapefiles see Chapter 4, Program Structure. The important field in the attribute table used by AEGIS/WIN is the field with the name of the map and the underscore (e.g. *Florida_* or *My_farm_*). To create your FILEX (see DSSAT, volume 2 for details), use the same number in the field *ID_FIELD, located* in the *FIELDS section of the FILEX. *ID_FIELD* appears in the *Field name* column FNAM in the summary output file.

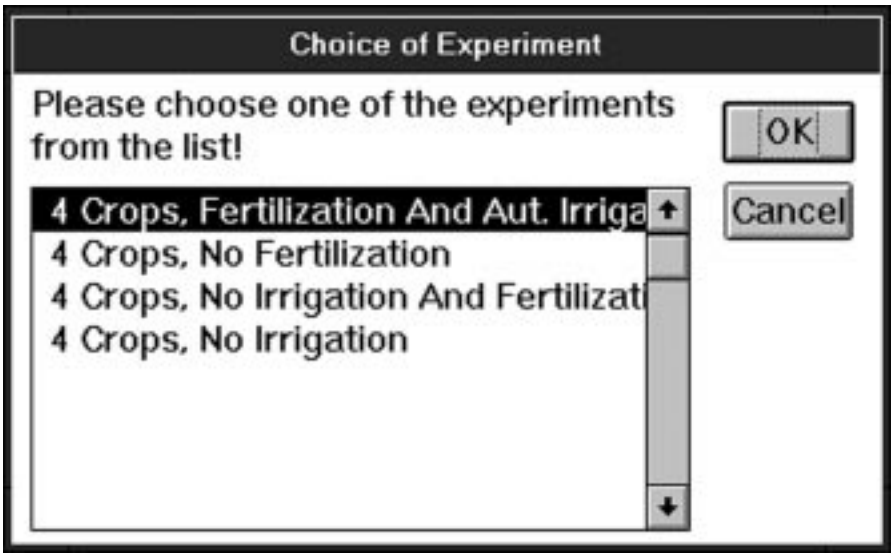
It is not necessary to add the correct soil and weather codes for each field to the FILEX. Weather and soil codes are contained in the polygon attribute table. The codes in the polygon attribute table have priority, so that they are automatically written to the FILEX. Therefore, it does not matter if the FILEX contains any values for the weather and the soil. Screen 11 shows a part of the *FIELDS section of a management input file.

RUNNING THE SIMULATION

After creating the management input files, simulation can proceed. AEGIS/WIN will either run all fields of a management input file or only the selected fields of a map or coverage.



Screen 11.

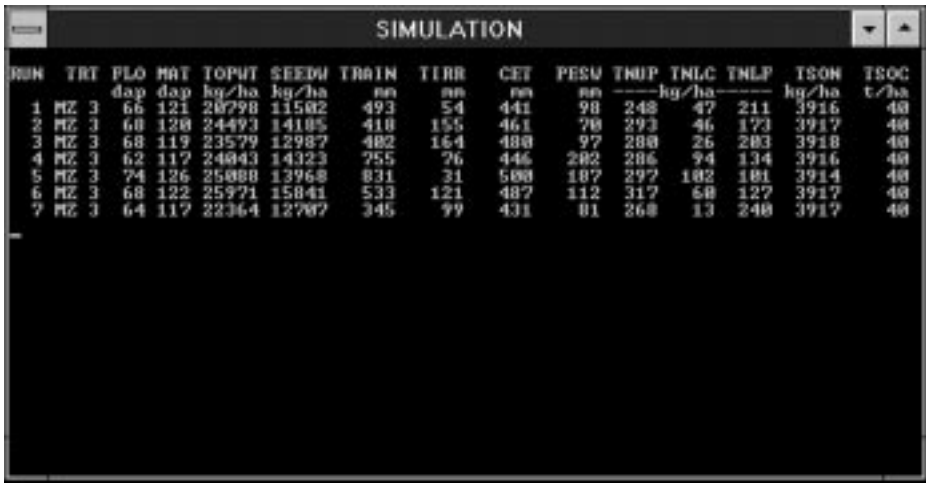


Screen 12.

Run All Fields

Choose 'Run Simulation' from the menu 'DSSAT3-Interface' and select an experiment from the experiment list (see screen 12). The list contains the description of all experiment files that are in the actual experiment list (file 'exp.lst'). The experiment list must be established before using the DSSAT-Shell. Each file defines simulations to be performed for all fields or a subset of the fields (or polygons) of the map. After creating the input files, the user can proceed with the simulation.

AEGIS/WIN will now read the selected experiment file, and first checks whether FILEX contains the same soil and weather codes as the attribute file of the coverage.



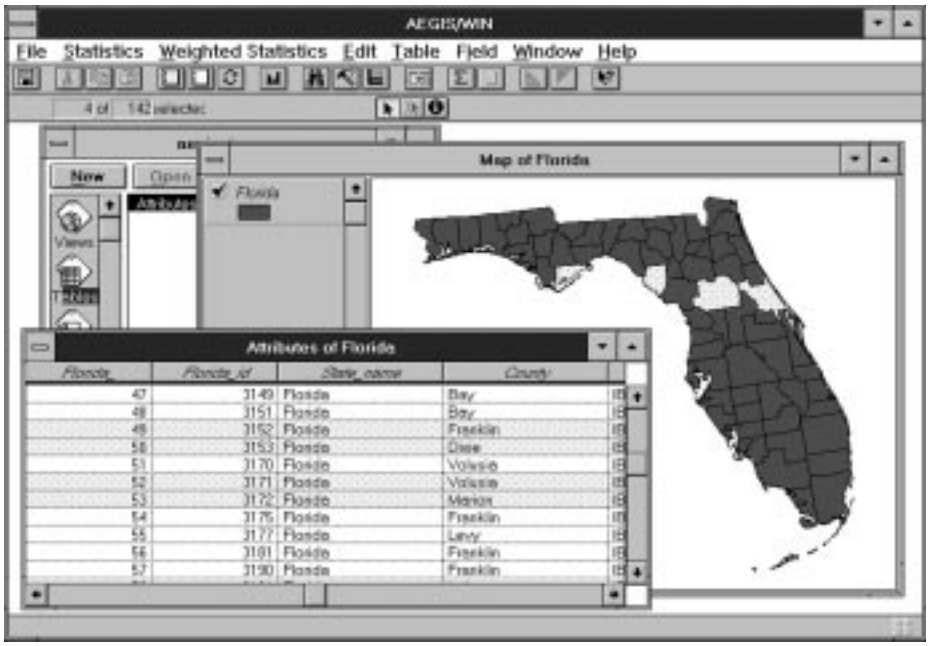
Screen 13.

The codes in the coverage file have priority, so that they are automatically written to FILEX. AEGIS/WIN then runs the simulation for all fields of FILEX using the driver program 'mapdrv.exe'. When the driver reaches a field with the same combination of soil, weather and management as a previous field, it skips the simulation and takes the results of the former simulation run. Progress of the simulation and the final results for each simulation run can be viewed in a DOS-window. Screen 13 shows the simulation window for a field with maize after a run of 7 consecutive years. The simulation is done for the number of years defined in FILEX.

Run Selected Fields

To run the simulation for only part of the fields of your map or coverage, use 'Run Selected Polygons' from the menu 'DSSAT3-Interface,' which is only enabled when fields are selected. Select the fields to be run from either the map or the attribute table. The method used to select polygons is different depending on the map format used. With ARC/INFO coverages, select a record by clicking on corresponding records in a table using the toolbar button with the one screen (see screen 7). The selected record is highlighted in the table. To select several polygons, press the [Shift]-Key and click on the records. The colors of the corresponding polygons in the map are changed to yellow to indicate which polygons have been selected (see screen 14).

With ArcView shapefiles, click on 'Start Editing' from the Menu 'View' before selecting the records directly from the map. In this case you can recognize the selected polygons by black boxes which mark the extent of each polygon (screen screen 7). After selecting the polygons, 'Run Selected Polygons' under the 'DSSAT3-Interface' menu is enabled and



SCREEN 14.

you can start the simulation. First you have to select an experiment from the experiment list (see screen 12). AEGIS/WIN now changes the weather and soil codes in the FILEX for the polygons not selected to '-99', so that they are skipped during the simulation.

SIMULATION RESULTS TABLES

CREATING SUMMARY OUTPUT TABLES

The simulation can create summary output files with the same name as the input files and the extension '.fls'. These '.fls'-files can be read into tables and linked to the maps. The following codes in the section *SIMULATION CONTROLS of the input file FILEX are necessary to get the summary output files:

1. Use Y = yes as FNAME; if N = no is selected, the system uses 'summary.out' as name. This file is overwritten after every run.

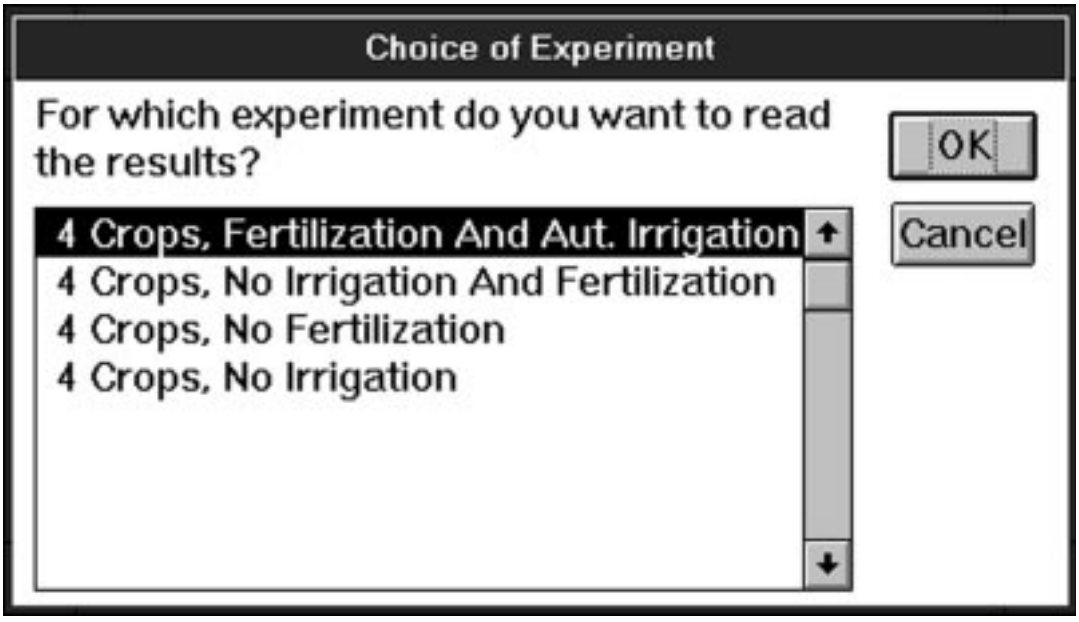
2. The field SUMRY needs A = Append, so that the simulation results are written to the output file.

Reading Simulation Results

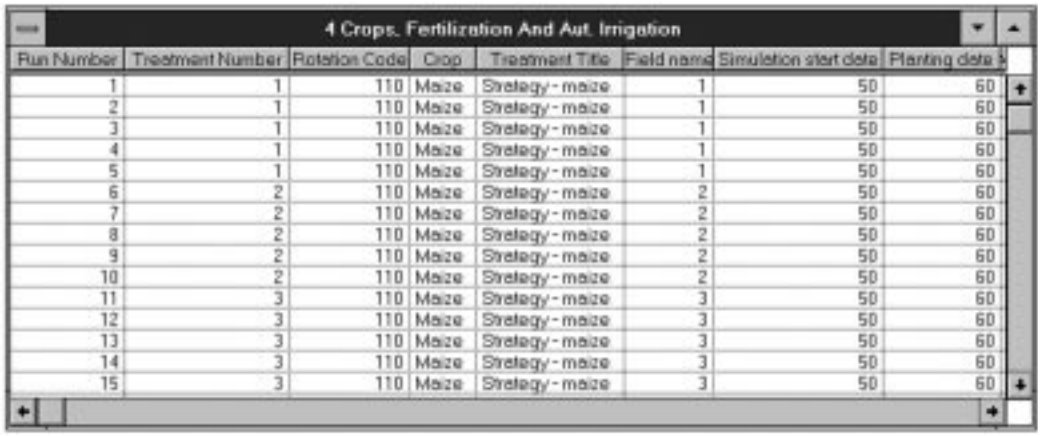
First read the summary output tables in AEGIS/WIN, which are ASCII-files in the defined DSSAT3-output format, before analyzing and mapping them. Click on 'Read Summary-Output' from the menu 'DSSAT3-Interface' to read the output files. Then select an output file from the list of the experiments for which output files exist (see screen 15). If only one output file is available, this will be automatically selected. After the selection, the system reads the output-file into an INFO-file with new, more descriptive headers (see screen 16). AEGIS/WIN uses the long description of the output-file header as a title for the corresponding table.

Averaging Simulation Results

The simulation result table can contain several results lines for one field (treatment), if the simulation is conducted for several years. For the example shown in screens 16 and 17, the simulation was run for 5 years. Therefore, another table is automatically created with the averaged simulation results for each field (see screen 17). If simulations were conducted for selected fields only, the averaged result table will have fewer records than the original one. For this case AEGIS/WIN prompts you to decide whether to overwrite or to update the old averaged result table. If you update it, the simulation results of the former simulation runs remain and only the simulation results of the selected fields are replaced.



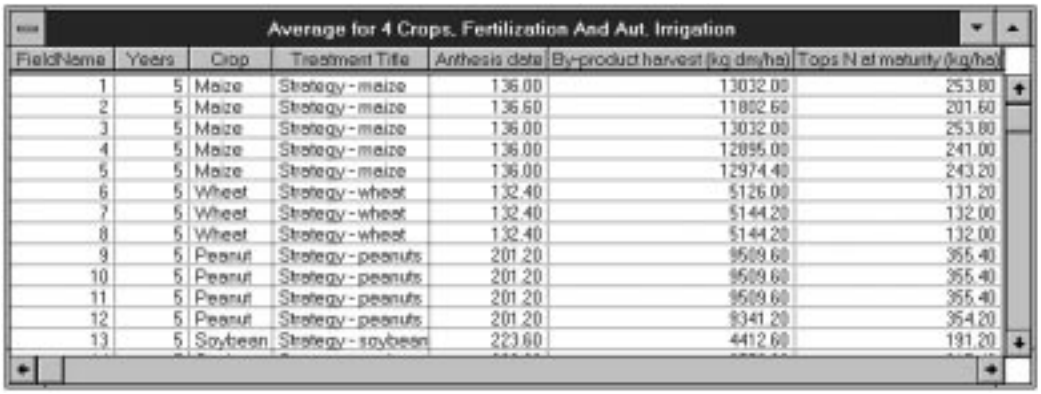
Screen 15.



4 Crops, Fertilization And Aut. Irrigation

Run Number	Treatment Number	Rotation Code	Crop	Treatment Title	Field name	Simulation start date	Planting date
1	1	110	Maize	Strategy - maize	1	50	60
2	1	110	Maize	Strategy - maize	1	50	60
3	1	110	Maize	Strategy - maize	1	50	60
4	1	110	Maize	Strategy - maize	1	50	60
5	1	110	Maize	Strategy - maize	1	50	60
6	2	110	Maize	Strategy - maize	2	50	60
7	2	110	Maize	Strategy - maize	2	50	60
8	2	110	Maize	Strategy - maize	2	50	60
9	2	110	Maize	Strategy - maize	2	50	60
10	2	110	Maize	Strategy - maize	2	50	60
11	3	110	Maize	Strategy - maize	3	50	60
12	3	110	Maize	Strategy - maize	3	50	60
13	3	110	Maize	Strategy - maize	3	50	60
14	3	110	Maize	Strategy - maize	3	50	60
15	3	110	Maize	Strategy - maize	3	50	60

SCREEN 16.



Average for 4 Crops, Fertilization And Aut. Irrigation

FieldName	Years	Crop	Treatment Title	Anthesis date	By-product harvest (kg dm/ha)	Tops N at maturity (kg/ha)
1	5	Maize	Strategy - maize	136.00	13032.00	253.80
2	5	Maize	Strategy - maize	136.60	11802.60	201.60
3	5	Maize	Strategy - maize	136.00	13032.00	253.80
4	5	Maize	Strategy - maize	136.00	12895.00	241.00
5	5	Maize	Strategy - maize	136.00	12974.40	243.20
6	5	Wheat	Strategy - wheat	132.40	5126.00	131.20
7	5	Wheat	Strategy - wheat	132.40	5144.20	132.00
8	5	Wheat	Strategy - wheat	132.40	5144.20	132.00
9	5	Peanut	Strategy - peanuts	201.20	9509.60	355.40
10	5	Peanut	Strategy - peanuts	201.20	9509.60	355.40
11	5	Peanut	Strategy - peanuts	201.20	9509.60	355.40
12	5	Peanut	Strategy - peanuts	201.20	9341.20	354.20
13	5	Soybean	Strategy - soybean	223.60	4412.60	191.20

SCREEN 17.

Both tables are then added to the table list of the project window (see screen 18); they can be opened by clicking the [Open] button. Both tables contain information for all variables included in the summary output files. Use the horizontal scroll bar to go to a variable and the vertical scroll bar for a field or a run. For manipulation of the tables, i.e. selection of records or specific queries, use the predefined ArcView functions (see the ArcView manuals or online help for more information). ArcView also provides for the creation of charts to visualize the contents of different fields.

STATISTICS FOR RESULT TABLES

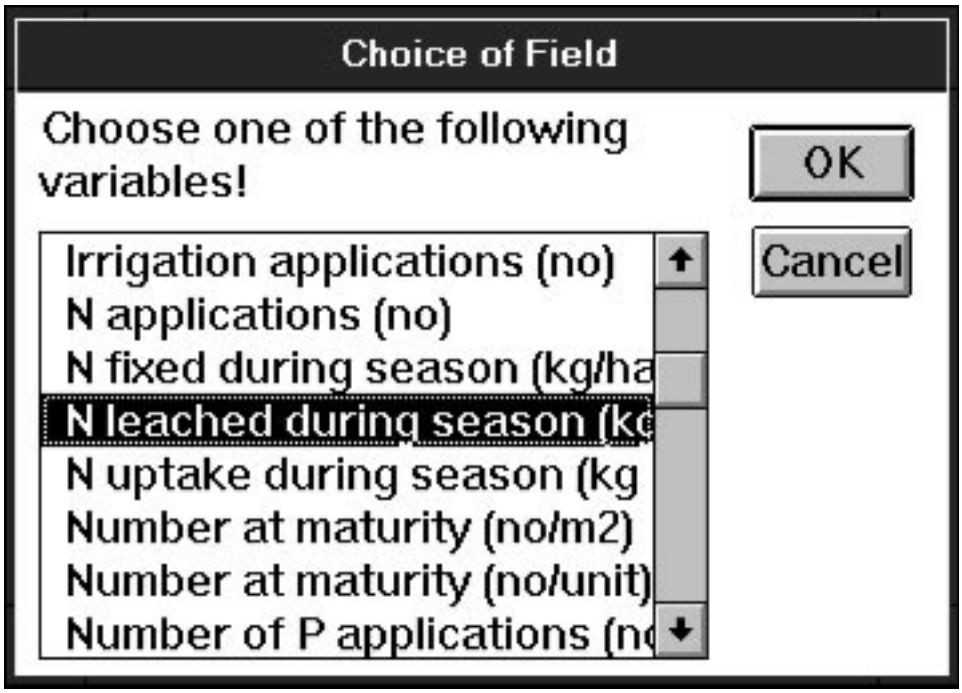
When the result tables are opened, the menu bar is changed to a table-specific menu bar. Compared with the original ArcView menu bar, new menus called 'Statistics' and 'Weighted Statistics' are now available. Both menus provide a more detailed statistical analysis of the simulation results.

Screen 18.

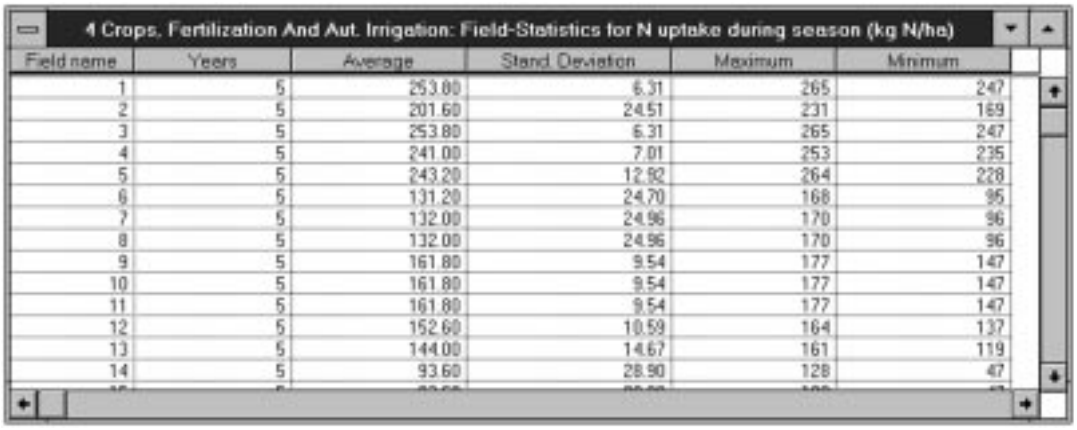
Statistics

The statistical functions for 'Fields', 'Crops' and 'Treatments' are only enabled when a table with original simulation results is opened and active. When the table with average values is active, do not use these statistics as the table already contains average results. 'Fields' calculates average, standard deviation, maximum and minimum for variable selected from the result table. Thhe statistics are calculated for each field of the coverage across all the years of a particular simulation scenario. If you choose the function 'Treatments', the system uses the field called 'Treatment Title' for averaging.

To start one of these statistical functions, first select the result variable you are interested in, e.g., 'N leached during season' (see screen 19). An example of these results is shown in screen 20.



Screen 19.



Field name	Years	Average	Stand. Deviation	Maximum	Minimum
1	5	253.80	6.31	265	247
2	5	201.60	24.51	231	169
3	5	253.80	6.31	265	247
4	5	241.00	7.01	253	235
5	5	243.20	12.92	264	228
6	5	131.20	24.70	168	95
7	5	132.00	24.96	170	96
8	5	132.00	24.96	170	96
9	5	161.80	9.54	177	147
10	5	161.80	9.54	177	147
11	5	161.80	9.54	177	147
12	5	152.60	10.59	164	137
13	5	144.00	14.67	161	119
14	5	93.60	28.90	128	47

Screen 20.

Weighted Statistics

The menu 'Weighted Statistics' allows use of the area of the polygons as a weighting factor. The areas of the polygons are not included in the simulation result table; it is part of the polygon attribute table of the map. Before using these functions, first join the map to the simulation results. Then, open the polygon attribute table by clicking

on 'Open Attribute Table' from the menu 'View'. **Important**: *All commands of the menu 'Weighted Statistics' are enabled only when the active window contains a polygon attribute table that is joined to a result table.*

The map can only be joined to the averaged simulation result table. Therefore, the joined polygon attribute table already contains averaged simulation results for the different polygons. The 'Weighted Statistics' feature uses these results for further averaging using 'Crop', 'Soil', Treatment Title' or 'Weather Station' as common variable. Similar to the use of normal statistics, first select a variable when starting one of these statistical functions. This list currently contains only variables that are related to the area so that weighted statistics has a meaning. The resulting table (see screen 21) contains no normal statistical results, such as average, standard deviation, maximum, minimum fields, for the 'Weighted Average' and the 'Total Area'. If you know the units for the area of the map, you can easily calculate total production.

My_farm: Crop-Statistics for Yield at harvest (kg dm/ha)

Crop	Count	Weighted Average	Total Area	Average	Standard Deviation	Maximum	Minimum
Maize	5	16421.50	56.15	16575.64	574.35	16983.60	15449.60
Peanut	4	4173.11	67.10	4164.85	24.68	4207.60	4150.60
Soybean	4	2853.40	99.70	2853.65	142.98	2936.20	2606.00
Wheat	3	3621.82	61.00	3623.20	5.94	3627.40	3614.80

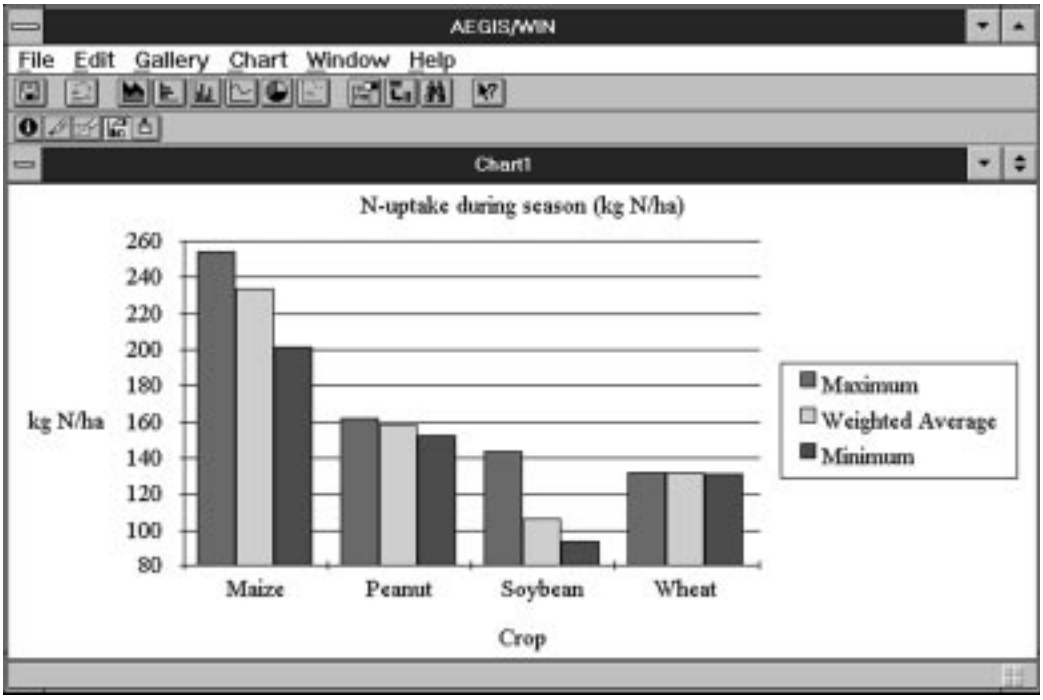
Screen 21.

Charts of Result Tables

Use AEGIS/WIN to create charts by clicking on 'Chart...' from the Menu 'Table' or by clicking on the chart button of the button bar (for more information see ArcView manuals or online help). The result tables, especially of the statistics functions, can be used for the creation of charts because they already contain condensed information. The original simulation result tables normally contain too much information to produce an easily understandable and clear chart. Screen 22 shows the results of the creation of a column bar chart using the weighted statistics result table for the N-uptake of different crops.

Creating Thematic Maps

The main purpose in developing AEGIS/WIN was the creation of thematic maps for a better spatial understanding of the simulation outcomes. This program component is



Screen 22.

therefore the most important and extensive part of AEGIS/WIN. It contains the menus 'Simulation Results', 'Legend' and 'Utilities'.

Parts of a Map in AEGIS/WIN

For the creation of maps, it is necessary to know the terminology used by ArcView. To draw maps, ArcView uses views and themes. The views are listed in the project window. Screen 3 shows the view called 'Map of Florida'.

A *view* is an interactive map that lets you display, explore, query and analyze geographic data in ArcView. A view defines the geographic data that will be used and how it will be displayed, but it does not contain the geographic data files themselves. Instead, a view references these source data files. This means that a view is dynamic, because it reflects the current status of the source data. If the source data changes, a view that uses these data will automatically reflect the change the next time the view is redrawn. The same data can be displayed on more than one view.

A view is actually a collection of themes. A *theme* is a set of geographic features in a view. The themes in a view are listed in its *Table of Contents*. For example, a view of a farm might have one theme representing the fields, one theme representing buildings,

and one theme representing the soils. Each theme has its own *legend* displayed in the Table of Contents. A theme's legend controls how the theme is displayed in the view. The view in screen 3 contains only one theme called 'Florida' with a simple legend. A few of the operations you can perform on a view work on the *active* theme(s), e.g., deletion. To make a theme active, click on the theme's name or legend in the Table of Contents. The theme will highlight to indicate that it is active.

Other operations work only on *visible* theme(s), e.g., changing of the legend. The visible themes are the themes that are drawn in the map. You can control which themes are drawn in your view by simply turning them on or off. Use the check box that is next to a theme in the Table of Contents to turn the theme on or off. The theme 'Florida' in the view 'Map of Florida' (screen) is active and visible. For more information, see the ArcView manuals or online help.

Joining Maps to Result Tables

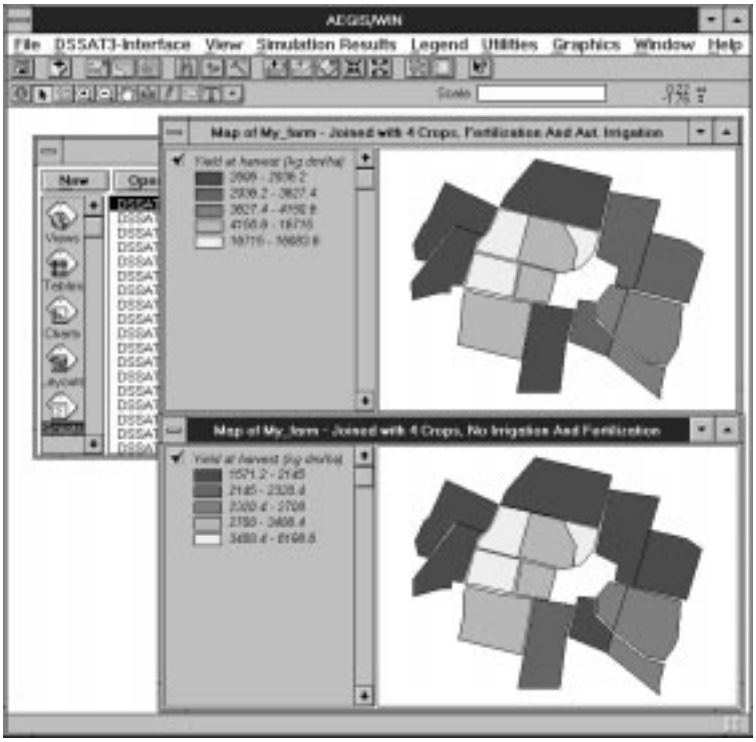
Before simulation results can be viewed in the map, the map or the attribute table of the map must be joined to the simulation result table. To do this, use 'Join to Result table' from the menu 'DSSAT3-Interface'. The menu 'Simulation results' is then used to show simulation results in the map. The pop up menu displayed after opening the map will indicate that most of the menu entries are disabled. Only those entries that are not simulation results or are entries from the attribute table of the opened map can be accessed.

The system allows you to only join result tables that have been averaged and contain simulation results for the active map, e.g., the map which is in the foreground, using the dialog in screen 23. You can only join result tables that are included in the table list of the project window, INFOtables. If you have different summary output files, have the system read all the files before joining them. If only one result table is available, the system automatically joins this table. After joining the tables, the title of the active view is changed; it informs you which table was joined (see screen 24). At the same time all menu entries of the menu 'Simulation results' are enabled.

To join another available table or to read another summary output file, click on 'Unjoin table' from the menu 'DSSAT3-Interface'. The menu items 'Read summary output' and 'Join to Result table' are then enabled and disabled depending on the joining status of the active map. To compare the results of different simulations in two different ways, open one map, join it to the first table, and study the results. Then unjoin the table, join it to the second table, and study these results. It is, however,

Screen 23.

more convenient to open the same map twice and to join one with the first and the other with the second result table (see screen 24).



Screen 24.

Modifying Themes

Adding Defined Themes

After joining a map to a result table all entries of the menu 'Simulation results' are enabled (see screen 25). The first twelve commands of the menu add defined variables as new themes to the view. The new theme is automatically set active and visible. All other themes are no longer visible. Screen 26 shows the result of the 'Add crops' command. A new active and visible theme called 'Crop' is added.

If the selected variable contains character data, the number of classes is set to the number of different elements in the field of the corresponding table. If the field contains numbers, the legend depends on the number of different values. If the field contains no more than 7 different values, the legend shows these values. When there are more than seven different values in the data field, AEGIS/WIN builds five classes that contain an equal range of values in each class. The new theme appears every time on the top of the Table of Contents (TOC). When more than one theme is visible, the

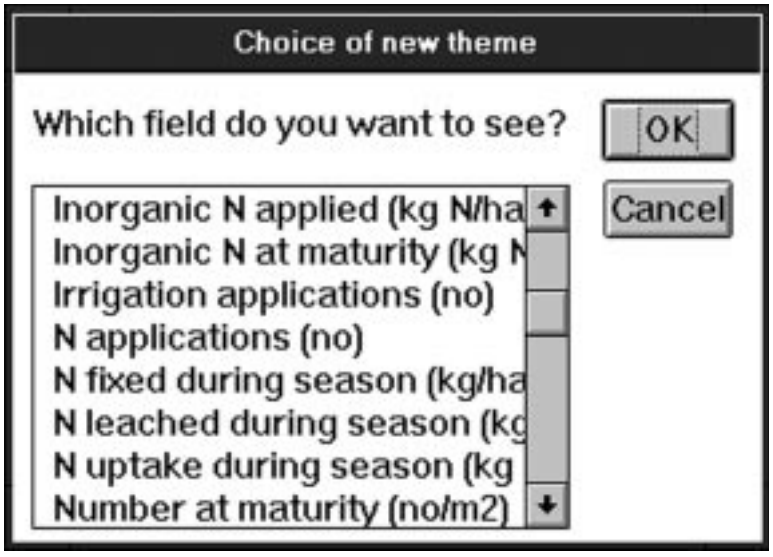
Screen 25.

theme on top is drawn last. Therefore, it can not be covered by other visible themes. The series of the themes can be changed by drawing themes in the TOC with the mouse.

Adding Other Theme(s)

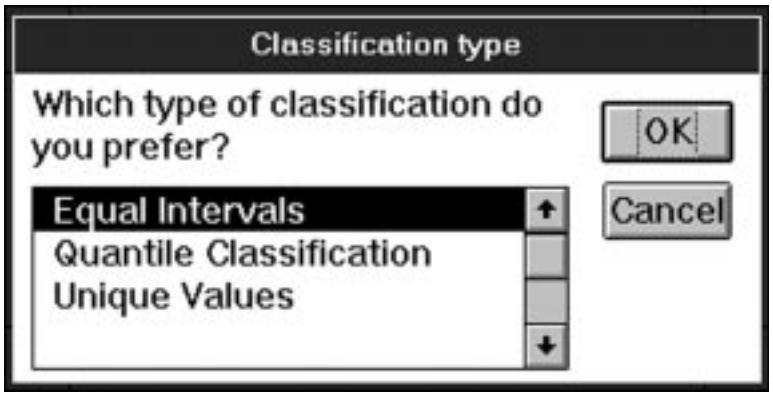
The entries in the second part of the Menu 'Simulation results' allow a more flexible means of adding or changing of themes. 'Add other theme' offers a list of all available variables in the corresponding table (screen 27) and the list is alphabetically sorted. Scroll through the list to search for the field of interest. After selecting one field, AEGIS/WIN checks the type of this field. If there are numbers in the table, the system

Screen 26.



Screen 27.

prompts you for the number of classes to use for the classification. The default value is a tenth of the number of records in the table. Use either the default value or enter your own number of classes. The system also offers you two types of classifications (see screen 28). Equal classification builds classes that contain an equal range of values in each class. Quantile classification creates classes that contain an equal number of records in each class. For fields with character data, the system performs a classification so that every unique value is placed in a separate class. If the number of unique values exceeds 64, AEGIS/WIN will not perform the classification.



Screen 28.

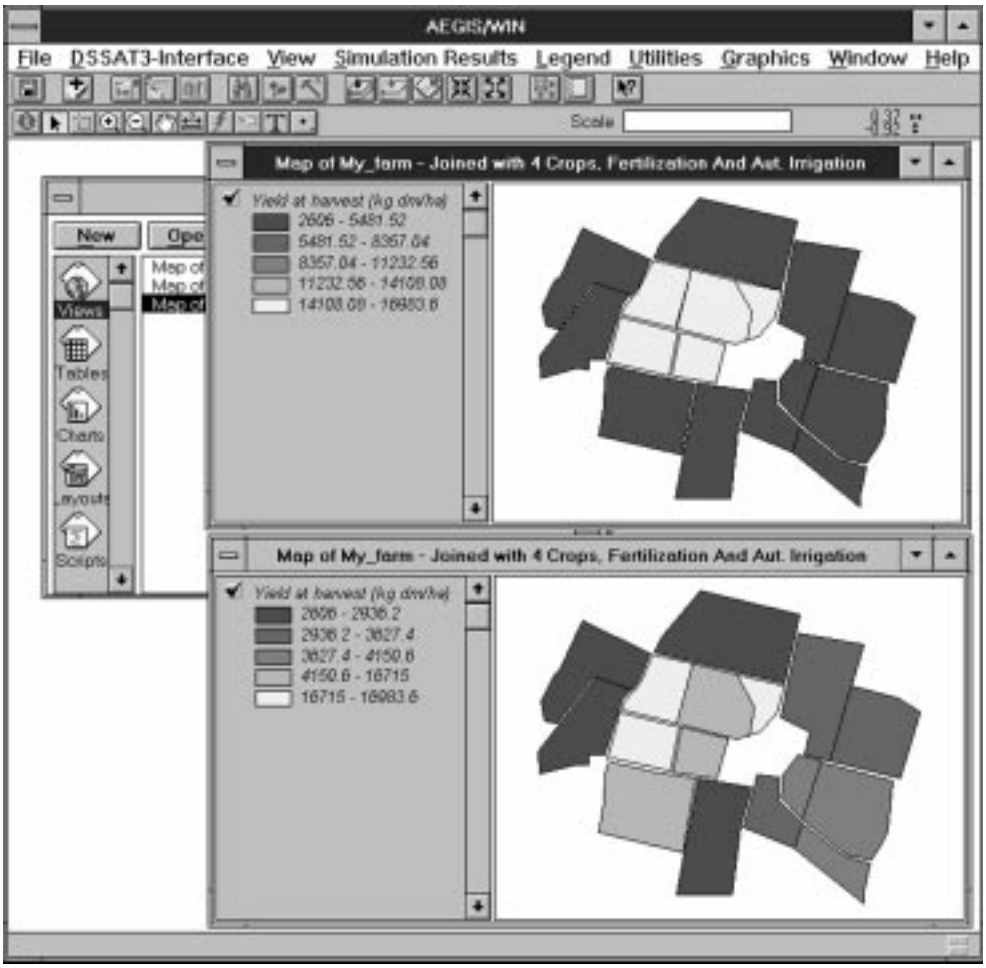
Screen 29 shows the results of different classification methods. The upper map shows the use of equal intervals of around 3,000 kg/ha. The problem here is the second, third and fourth classes do not appear in the map because there were no yield values for these ranges. The lower map shows the quantile classification. The fourth class (4,150 - 16,715) is very large in comparison with the range of all other classes. All classes contain a similar number of polygons in the map.

Changing Visible Theme

'Change visible theme' allows the change of the actual visible theme. Use this command when there are many themes in the view or when the actual visible theme is no longer of interest. It is important to have only one theme visible. The system prompts you for the variable you want to see, the number of classes, and the type of classification. No theme is added to the Table Of Contents in this operation, only the actual visible theme is changed. This command prevents overloading the view with themes no longer needed.

Deleting Themes

Views that are no longer needed can be deleted from the Table Of Contents using 'Delete active theme(s)' from the menu 'View'. The command deletes all active



SCREEN 29.

themes. To make a theme active, click on the theme's name or legend in the Table of Contents. The theme will be highlighted to show that it is active. To make more than one theme active, click on the themes and press [Shift] Key at the same time.

MODIFYING THE LEGEND

The menu 'Legend' provides different commands to change the legend. Each command, except for 'Hide/Show Legend', modifies only the visible theme.

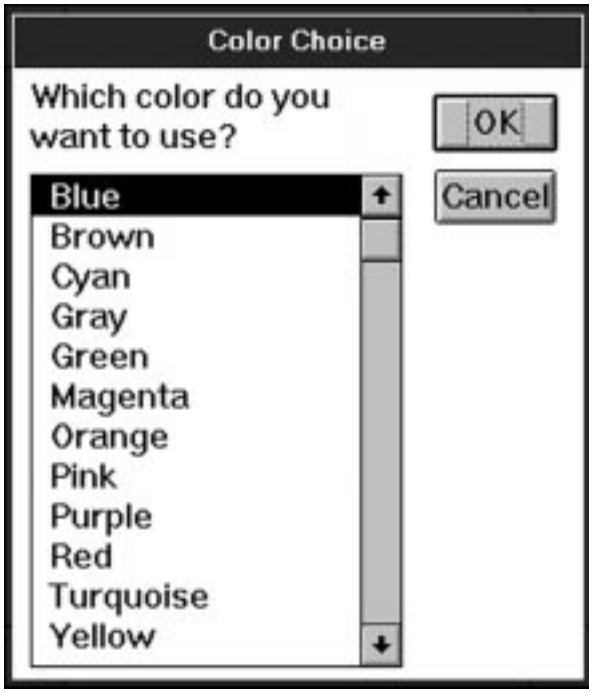
Change Title Modifies the title of the theme. The old title as shown as a default value.

Change Classification Changes the classification of the theme. It is only possible to make modifications when the field of the corresponding table contains numbers. You can change the number of classes and the type of classification.

Colors - Random The original legend contains randomly created colors. Use this command to create other colors randomly.

Colors - Ramp Creates a ramp of colors from the first color to the last color of the existing legend.

Colors - List If the randomly created colors are not satisfactory, select from an alphabetical list of 12 colors from 'Blue' to 'Yellow' (see screen 30). Select only one color for a simple legend with one element. Select the first and the last color for a legend with more elements. AEGIS/WIN then creates a ramp using the selected colors.



Screen 30.

Hide/Show Legend This command hides and shows the legends of all themes in the Table Of Contents. After hiding, only the titles of themes are visible and provides a better overview of many themes with long legends in your view.

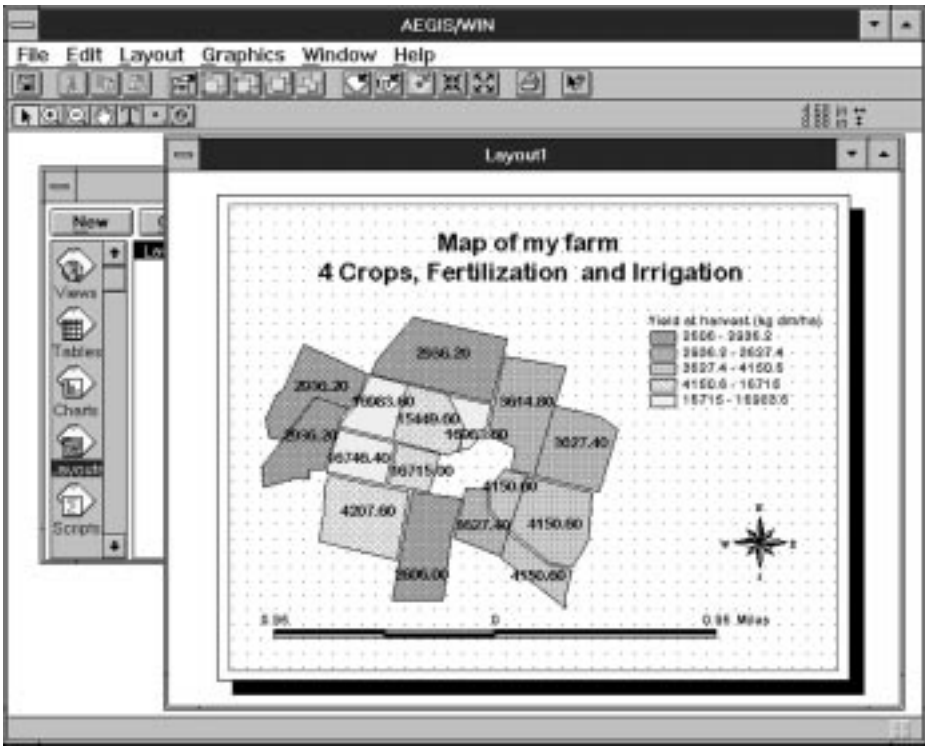
Additional Utilities

The menu 'Utilities' contains two commands.

Label visible theme Labels the map according to actual visible theme. To change fonts and size, use the Symbol Palette of ArcView. To open it, click on 'Show Symbol Palette' from the Menu 'Window'

Create Layout Opens the Layout-Editor of ArcView to create printing outputs of the view. For more information, see the ArcView manuals

Screen 31 shows a simple layout that was created with the Layout-Editor. Before the Layout Editor was started, the map was labeled. The scale bar will have only values



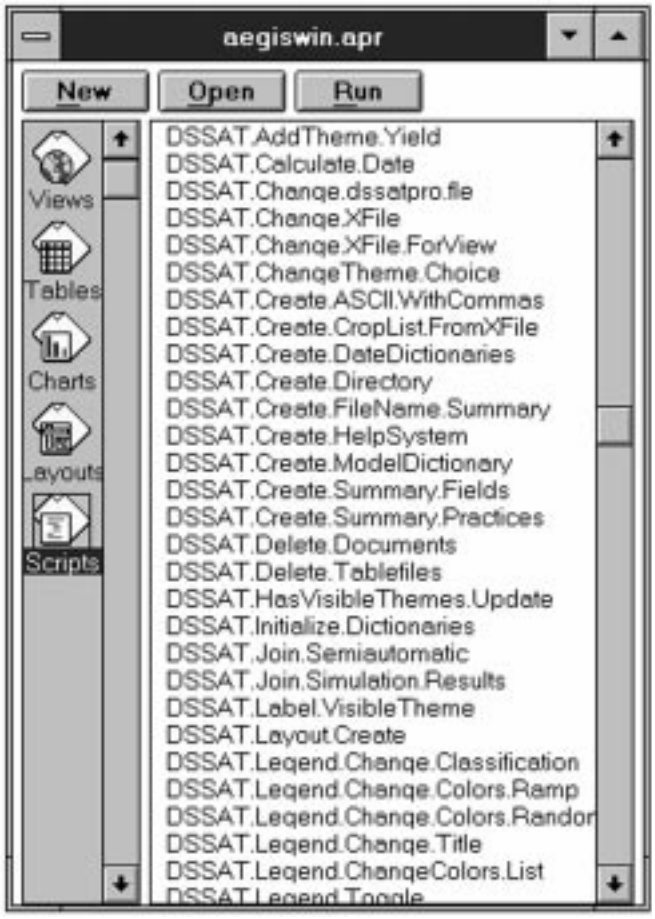
Screen 31.

when the map units in 'View Properties' are set. The layout can be modified, e.g., change of font and font size, and addition or removal of text. When the layout has the view you like, you can print it using 'Print' from the menu 'File'.

CHAPTER FOUR.
PROGRAM STRUCTURE

USE OF AVENUE SCRIPTS

The source code of AEGIS/WIN was written in modular form, using the macro programming language Avenue of ArcView. The Script Editor of ArcView was used to write Scripts with Avenue code, which can be compiled and executed. Each unique procedure or activity is written as a separate Script. Screen 32 shows a partial list of the written scripts. In total, AEGIS/WIN contains more than 80 scripts. All scripts are included in a project called '*aegiswin.apr*'. There are no separate files that contain the source code. When you load the project, all scripts are read and compiled and you can use it even if Avenue is not installed on your system.



SCREEN 32.

Directory Structure and Files

Geographic Data

The handling of geographic information depends on the format you use (ARC/INFO coverages or shapefiles) even though these formats are similar

Arc/Info Coverages

Arc/Info manages coverages, using sub-directories for every coverage. The name of the subdirectory is also the name of the coverage, e.g., **\aegiswin\florida**. Each subdirectory contains several files that contain the coded information necessary to create the map.

The most important file is the so-called polygon attribute table (**pat.dbf**) that stores information which quantifies and describes the geographic features of a coverage. The **pat.dbf** of every coverage used by AEGIS/WIN must contain the following set of items:

Shape	The type of shape of the corresponding geographic element.
Area	The area of the polygon. This real value is automatically calculated by ARC/INFO when the coverage is created or cleaned. *Do not change this value.*
Perimeter	The perimeter of the polygon. This real value is automatically calculated and assigned by ARC/INFO when the coverage is created or cleaned. *Do not change this value.*
coveragename_	The number of the polygon (e.g., florida_). This integer value is automatically calculated and assigned by ARC/INFO when the coverage is created or cleaned. *Do not change this value.*
coveragename_id	The polygon identification number, e.g., florida_id. This integer value is automatically set to zero by ARC/INFO when the coverage is created or cleaned. The user can replace it with another number as a linkage with other data files. It is not needed in AEGIS/WIN.
Wsta	The 4-character weather station code for the polygon. This code is normally composed of a 2-character institute code and a 2-character site code, or a 2-character site code and a 2-character code for the state or county. This item and its values must be added by the user (see ArcView manuals for modification of

tables). For example UFGA indicates a weather station at Gainesville (GA), managed by the University of Florida (UF).

Id_soil The 10-character soil mapping unit ID of the polygon. This code is normally composed of 2-character institute code, a 2-character site code, the 2-character year number, and the 4-character soil number. Examples are IBSB910015 and GAPL930001. The values of this item must be added by the user.

Attributes of Florida

Shape	Area	Perimeter	Florida_	Florida_id	State_name	County	Id_soil	Weth
Polygon	2.662435e+009	329527.800000	15	3079	Florida	Santa Rosa	IBSB910015	MFL
Polygon	1.830031e+009	193175.800000	16	3407	Florida	Madison	IBSB910015	TAFL
Polygon	5.321886e+007	43390.620000	17	3081	Florida	Duval	IBSB910015	UFGA
Polygon	1.676309e+009	287623.900000	18	3082	Florida	Escambia	IBSB910015	MFL
Polygon	1.607895e+009	216901.900000	19	3408	Florida	Jefferson	IBSB910015	TAFL
Polygon	1.590041e+009	207281.700000	20	3084	Florida	Washington	IBSB910015	MFL
Polygon	1.822378e+009	209208.500000	21	3409	Florida	Leon	IBSB910015	TAFL
Polygon	1.344788e+009	172685.000000	22	3085	Florida	Gadsden	IBSB910015	TAFL
Polygon	1.114302e+008	78907.620000	23	3088	Florida	St. Johns	IBSB910015	UFGA

Screen 33.

The first five items are automatically written in the PAT by ARC/INFO, the other two items must be added by the user. The unique key item, which is used by AEGIS/WIN for linking the map to the simulation results, is the polygon number *coveragename_*, e.g., *florida_*. Screen 33 shows a part of the pat.dbf for the coverage of Florida.

ArcView Shapefiles

The shapefile format defines the geometry and attributes of geographically-referenced features in several files (minimum of 3, maximum of 7) with specific file extensions. These files are stored in the AEGIS/WIN directory. The three most important shapefiles are:

***name*.shp** stores the feature geometry.

***name*.shx** stores the index of the feature geometry.

***name*.dbf** a dBASE file that stores the attribute information of features. When a shapefile is added as a theme to a view, the file is displayed as the feature table.

Four other files (***.sbn, *.sbx, *.aih, *.ain**) additionally can exist to store different indices, e.g., the actual selected polygons. They are automatically created by ArcView if necessary.

The most important file is the so-called polygon attribute table ***name.dbf***, e.g., *my_farm.dbf* which stores information that quantifies and describes the geographic features of a coverage. The ***name.dbf*** is used by AEGIS/WIN and must contain the following items:

Shape The type of shape of the corresponding geographic element (usually polygon)

Id The polygon identification number. This field is automatically set to zero by ArcView when the shapefile is created. The user may replace it by another number as a linkage with other data files. Actually it is not needed in AEGIS/WIN.

name_ The number of the polygon. You have to add this number after drawing the polygons. It is used for defining field management practices in FILEX and for the linkage with the simulation results. It has to be a number to conform to the the ARC/INFO format, where the number is automatically generated when the coverage is created.

Wsta The 4-character weather station code for the polygon.

Id_soil The 10-character soil mapping unit ID of the polygon.

Area The area of the polygon. In contrast to ARC/INFO coverages, the area is not automatically created, but it can be added by the user. This value is

Attributes of My_farm

Shape	Id	My_farm_	Wsta	Id_soil	Area
Polygon	0	1	UFGA	IB00000007	12.25
Polygon	0	2	UFGA	IB00000010	17.20
Polygon	0	3	UFGA	IB00000007	6.00
Polygon	0	4	UFGA	IB00000005	12.20
Polygon	0	5	UFGA	IB00000002	8.50
Polygon	0	6	UFGA	IB00000002	27.00
Polygon	0	7	UFGA	IB00000001	23.50
Polygon	0	8	UFGA	IB00000001	10.50
Polygon	0	9	UFGA	IB00000005	6.10
Polygon	0	10	UFGA	IB00000005	25.50
Polygon	0	11	UFGA	IB00000005	9.00

SCREEN 34.

not necessary for joining the map to simulation results. However, if the value is available, enter it for use in the calculation of area weighted statistics. Other program parts do not require this information.

The first two items are automatically written to the file ***name*.dbf** by ArcView; the other four items must be added by the user. The unique key item, which is used for the linkage of the map to the simulation results, is the polygon number *name_*, e.g., My_farm_. Additional attributes can be added by the user if desired, e.g., slope, name of the field. Thematic maps can then be created for these attributes. Screen 34 shows part of the attribute table of the shapefile *my_farm.shp*.

Simulation Data

AEGISWIN uses '..*data*'-subdirectories for the handling of the data related to the simulation runs. The subdirectories are automatically created if they do not exist. For *ARC/INFO coverages,* only the subdirectory '.. *data*' is required since the coverage information is already located in a directory with the name of the coverage.

If you create a new *ArcView shapefile,* a directory with the name of the shapefile and a subdirectory '..*data*' is created. The directory with the name of the shapefile usually contains no files, because all data are stored in the subdirectory '..*data*'.

If you use the path '**\aegiswin**, you have the following directory structure:

\aegiswin
\aegiswin*name* (for every coverage or shapefile)
\aegiswin*name***\data** (for every coverage or shapefile)

The **'..\data'** subdirectory should contain the following files:

***.flx**	management input files for running DSSAT crop models
soil.sol	soil information file
***.wth**	weather files for the coverage or shapefile
***.fls**	summary output files
start.bat	batch file to run the simulation *(Note: This file is created from AEGIS/WIN and is overwritten with each simulation run.)*
exp.lst	experiment list file, which has to be created using the DSSAT-Shell
arc*.*	INFO-files which are created from AEGIS/WIN

Compatibility of AEGIS/WIN and DSSAT v3.5

Attention must be given to the following details in order to successfully use a DSSAT v3.5 management input file in AEGIS/WIN:

1. In order to conduct any simulations with a FILEX in AEGIS/WIN, it must be assigned to a coverage. The selected coverage must have enough polygons assigned to all specified practices in FILEX.

2. AEGIS/WIN does not use the soil type (ID_SOIL) and weather station (WSTA) listed in the section *FIELDS of FILEX, but those specified in the polygon attribute table of the selected coverage. If they do not agree, the original values of WSTA and ID_SOIL in the FILEX will be overwritten.

3. You have to add the name of a new FILEX to the experiment list file *exp.lst* using the DSSAT-Shell (*F9*), before you can run the simulation. Otherwise the experiment will not be listed in the list of experiments.

CHAPTER FIVE.
KEY PROGRAMMING CONCEPTS

It is impossible to explain the full functionality of AEGIS/WIN programming in this document. However, key programming concepts on the operation of AEGIS/WIN are presented here.

CHECKING AVAILABLE PROGRAM MODULES

AEGIS/WIN uses the spatial driver 'mapdrv.exe' of DSSAT v3.5 to run the crop model programs. Additionally it requires the files '*data.cde*', '*detail.cde*' and '*dssatpro.fle*', that are part of DSSAT v3.5. During the initialization session, AEGIS/WIN checks the availability of the required DSSAT v3.5 files. If one of the above mentioned files does not exist in the DSSAT v3.5 directory, an error message will appear and AEGIS/WIN will not run.

ACCESSING DSSAT V3.5

AEGIS/WIN uses DSSAT v3.5 for the creation and administration of the input files and for the run of the simulations. Before accessing DSSAT v3.5, open a map, so that DSSAT v3.5 can access the correct directories and data. Before starting the DSSAT-shell, AEGIS/WIN sets the path for the 'Spatial' data in the setup-file ('*dssatpro.fle*') to the subdirectory ..*data* of the current coverage or shapefile. This setup-file is read during the initialization of DSSAT v3.5. Thereafter, each time the DSSAT-Shell is started, the user will have access to the corresponding input files.

Each time a simulation is run, AEGIS/WIN changes the current directory to the subdirectory ..*data* of the current coverage or shapefile. DSSAT is called, using a pif-file (*dssatv35.pif*), which starts *dssat35.exe*.

RUNNING THE SIMULATION

For each simulation the batch file *start.bat* is created. The file contains the call for the spatial model driver and the selected experiment number as attribute, e.g., c:\dssat35\mapdrv.exe 2.

PROGRAMMING WITH AVENUE

Avenue is the programming language and development environment of ArcView (ESRI, 1994b). ArcView provides the necessary customization and language environment tools in an easy-to-use framework. The graphical user interface can be created

to establish some initial properties for the graphical controls that a user will interact with, to fine tune the behavior and appearance of those controls, and to write Avenue code that responds to the interface operation. In addition, scripts written in Avenue can be linked to events, such as starting up and shutting down a project.

Avenue is an object-oriented scripting language. The emphasis in Avenue, as in all object-oriented systems, is on identifying objects and then sending them requests. Think of an object as a package that is comprised of tightly-coupled data and functionality. Contrast this with procedural programming practices, in which the language places an emphasis on the function (procedure). In Avenue, instead of calling functions explicitly with arguments, send a request to an object. When an object receives this request, it performs some action. ArcView's objects are members of a class hierarchy that are organized into functional categories related to all aspects of the application.

Avenue's statements are used to organize and structure when and how requests are made. Requests are Avenue's counterpart to the traditional function call. A function call and its implementation are in one-to-one correspondence, whereas a request can trigger one of any number of methods. A request specifies what an instance of a given class will do and a method specifies how it is done. Therefore, programming in Avenue consists of writing object requests rather than calling functions. By sending a request to an object, a method is activated, appropriate to the class of which the object is an instance. An object in Avenue always responds to a request by returning an object; in some cases, the request creates a new object and in other cases, the original request returns an existing object.

The code in Table 1 is from the script *'DSSAT.Change.dssatpro.fle',* which is part of AEGIS/WIN, and demonstrates how Avenue operates. The following explanations use **bold** characters for objects and ***bold/italic*** characters for classes and requests. The nomenclature of Avenue requires a dot between object and request. All lines starting with ' are comment lines.

SELF represents the object which is used as an argument in the call of the script. It is assigned to the object named **coverageName**. The command ***AsFileName*** in line 11 converts the object of the class ***String*** "$dssat3\dssatpro.fle" in an object of the class ***FileName*** with the name **nameOfFile.** Line 12 creates on object of class ***LineFile*** named **setupFile**. The class ***LineFile*** manages lines of characters stored on disk. After creating a ***List*** object called **lineList** (line 15), line 18 reads the lines of the file in the **lineList**. The request ***Read*** is a method of the class ***LineFile*** which reads a defined number of lines from a ***LineFile*** and places them in a list. The first attribute (**lineList**) is the list object, the second attribute defines the number of lines to be read. In this case

the second attribute is the result of the request ***GetSize,*** which is sent to the object **setupFile**. This returns an object with the number of lines of the ***LineFile*** object **setupFile**. Therefore, all lines of the file are read into the **lineList**. The lines 24 to 31 define **numOfLine,** which is the number of the line in the file which starts with the string "APD".

The ***for each*** statement of Avenue applies a set of requests to all elements of the **lineList**. The object **line List** contains a list of strings which represent the single lines of the read file. The request ***Left(3)*** returns the first three characters of the string. The ***if...then*** statement evaluates the expression (code = "APD"). If the expression returns the object **True**, the ***break*** statement is executed, which ends the loop. The lines 34 to 37 define the new string and exchange it in the list of lines. The variable **coverageName** represents an object of the class ***String,*** with the name of the active coverage. The requests + concatenates the three String objects and returns the resulting string. ***AsFileName*** converts the resulting string into an object of class ***FileName***. The request ***GetFullName*** to the resulting ***FileName*** object returns the full path including the drive as object of class ***String***. This new string is assigned to the variable **newDir.**

In line 35/36 **"APD"** and parts of the **newDir** are concatenated using the request **++,** which separates the strings with one space character. The resulting string is assigned to the new object **newLine**. The request ***Set*** replaces the element with the number **numOfLine** in the **lineList** with the object **newLine**. The lines 40 to 42 create a new object of the class ***LineFile*** with the name **nameOfFile** for writing. All elements of the **lineList** are written to the new file and the file is closed. The last line returns the ***String*** object **newDir** to the script, which has called this script.

For more information about object-oriented programming and the use of Avenue see the ArcView 3.0/3.1 manuals.

TABLE 1. AVENUE SCRIPT 'DSSAT.CHANGE.DSSATPRO.FLE'.

```
01 'Modifies the file "dssatpro.fle" in the dssat3-directory
02 'Changes the directory for the spatial data to the subdirectory
03 '..\data of the current coverage-directory
04 'SELF is the argument which has to be defined with the call
05 'of the script.
06
07 coverageName = SELF
08
09 'Defines FileName and opens file for reading as object
10 'of class LineFile
11 nameOfFile = "$dssat3\dssatpro.fle".AsFileName
12 setupFile = LineFile.Make(nameOfFile, #FILE_PERM_READ)
13
14 'Creates object of class List
15 lineList = List.Make
16
```

```
17 ‘Reads lines from file into lineList
18 setupFile.Read(lineList, setupFile.GetSize)
19
20 ‘Close the file
21 setupFile.Close
22
23 ‘Loop which looks for line which has to be changed
24 numOfLine = 0
25 for each element in lineList
26 code = element.Left(3)
27 if (code = “APD”) then
28    break
29 end
30 numOfLine = numOfLine + 1
31 end
32
33 ‘Changes line to new directory
34 newDir = (“$AEGISWIN\”+coverageName+”\DATA”).AsFileName.GetFullName
35 newLine = “APD”++newDir.Left(2).Ucase
36                ++newDir.Right(newDir.Count - 2).Ucase)
37 lineList.Set(numOfLine, newLine)
38
39 ‘Writes the new file dssatpro.fle
40 newSetupFile = LineFile.Make(nameOfFile, #FILE_PERM_WRITE)
41 newSetupFile.Write(lineList, lineList.Count)
42 newSetupFile.Close
43
44 return newDir
```

REFERENCES

Engel, T., G. Hoogenboom, J. W. Jones, and P. W. Wilkens. 1997. AEGIS/WIN - A computer program for the application of crop simulation models across geographic areas. Agronomy Journal 89(6):919-928.

ESRI. 1993. Understanding GIS - The ARC/INFO Method. Environmental Systems Research Institute, Inc., Redlands, CA.

ESRI. 1994a. ArcView, The Geographic Information System for Everyone. Environmental Systems Research Institute, Inc., Redlands, CA.

ESRI. 1994b. Customizing ArcView with Avenue. Environmental Systems Research Institute, Inc., Redlands, CA.

Hoogenboom, G., J.W. Jones, P.W. Wilkens, W.D. Batchelor, W.T. Bowen, L.A. Hunt, N.B. Pickering, U. Singh, D.C. Godwin, B. Bear, K.J. Boote, J.T. Ritchie, and J.W. White. 1994. Crop Models. p. 95 - 244. *In:* G.Y. Tsuji, G. Uehara, and, S. Balas (eds.). DSSAT v3. Vol. 2-2. University of Hawaii, Honolulu, HI.

Jones, J.W., L.A. Hunt, G. Hoogenboom, D.C. Godwin, U. Singh, G.Y. Tsuji, N.B. Pickering, P.K. Thornton, W.T. Bowen, K.J. Boote, J.T. Ritchie, and S. Balas. 1994. Input and Output Files. p. 1 - 94. *In:* G.Y. Tsuji, G. Uehara, and, S. Balas (eds.). DSSAT v3. Vol. 2-1. University of Hawaii, Honolulu, HI.

Tsuji, G.Y., G. Uehara, and S. Balas (eds.). 1994. DSSAT v3. University of Hawaii, Honolulu, HI.

Volume 4-4

Displaying Soil Water and Nitrogen Graphs

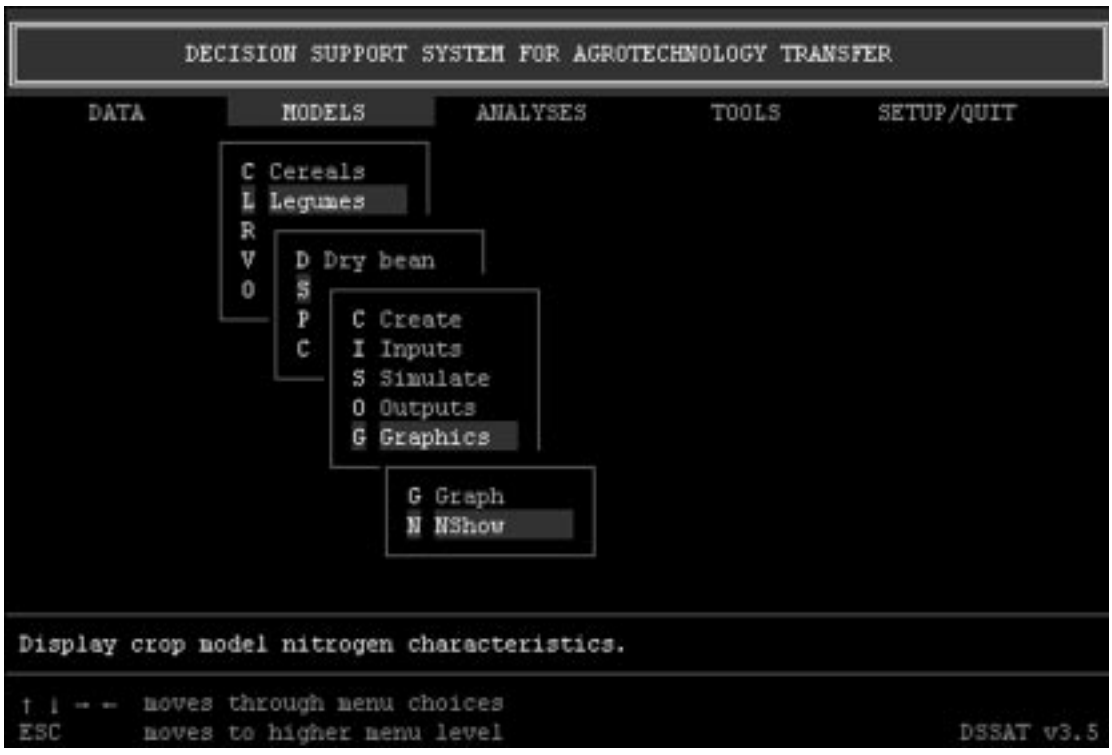
M.L. Cabrera
B. Baer

University of Georgia
Michigan State University
International Consortium for Agricultural Systems Applications

CHAPTER ONE. INTRODUCTION

N-SHOW (Cabrera, 1994) is a tool used within DSSAT used to assist in understanding aspects of nitrogen cycling in cropping systems. N-Show was developed to visualize changes in nitrate concentration that occur in a soil profile, resulting from fertilizer N applications and from its interaction with soil water. It demonstrates the potential for nitrate leaching when excessive N rates are used. N-SHOW is accessed from within the DSSAT shell under the GRAPHICS option for each crop model (Screen 1).

N-SHOW displays dynamic graphs of nitrate concentration with depth, plant nitrogen, leached nitrate, and soil water. The graphs can be selected from the output of selected model runs with DSSAT v3.5. The data may be viewed in a dynamic nature, in which time can be changed from the start of the simulation until final harvest, following the course of the simulation. The data can also be examined for any specific day and treatments can be compared.

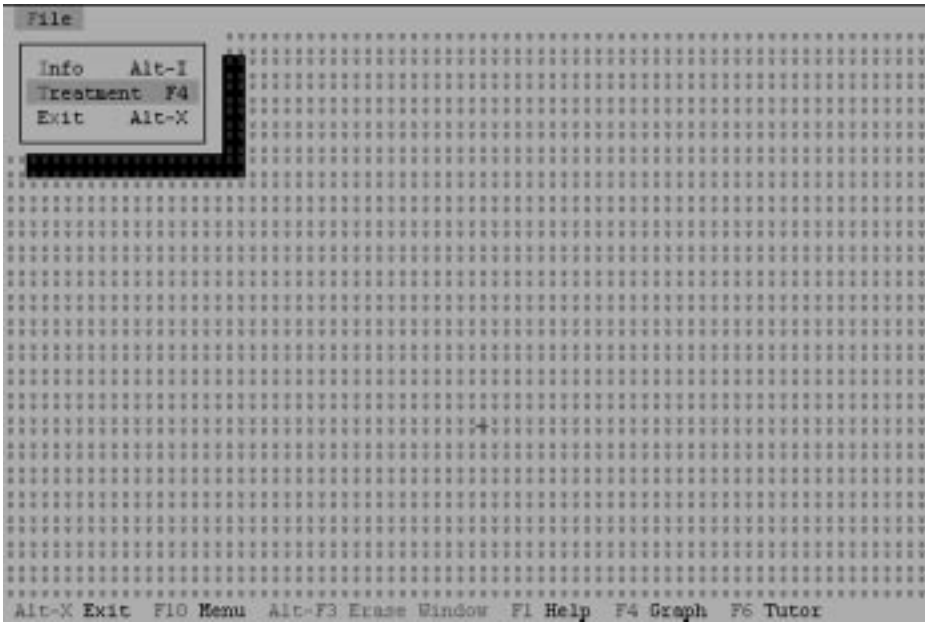


SCREEN 1.

CHAPTER TWO.
OPERATION OF N-SHOW

Select the FILE option followed by the TREATMENTS option to access different simulation runs (Screen 2). The treatment to be displayed is selected by choosing the option FILE in the top menu bar, followed by the option TREATMENT in the popup menu. The top menu bar and corresponding options can be accessed and selected by pointing and clicking a mouse. If a mouse is not available, the menu bar can be accessed with the special function F10 key; options can be accessed with the arrow keys and selected with ENTER. If you are not familiar with the type of graphs displayed, press the F6 key for a tutorial on them. Then double click on a run, or press tab and use the up and down arrow keys to browse the runs, using the spacebar to select the desired run. When the desired run is selected, click OK or press enter. In this example, we have chosen run 4, which represents a maize crop grown at the University of Florida under irrigated and high nitrogen management conditions.

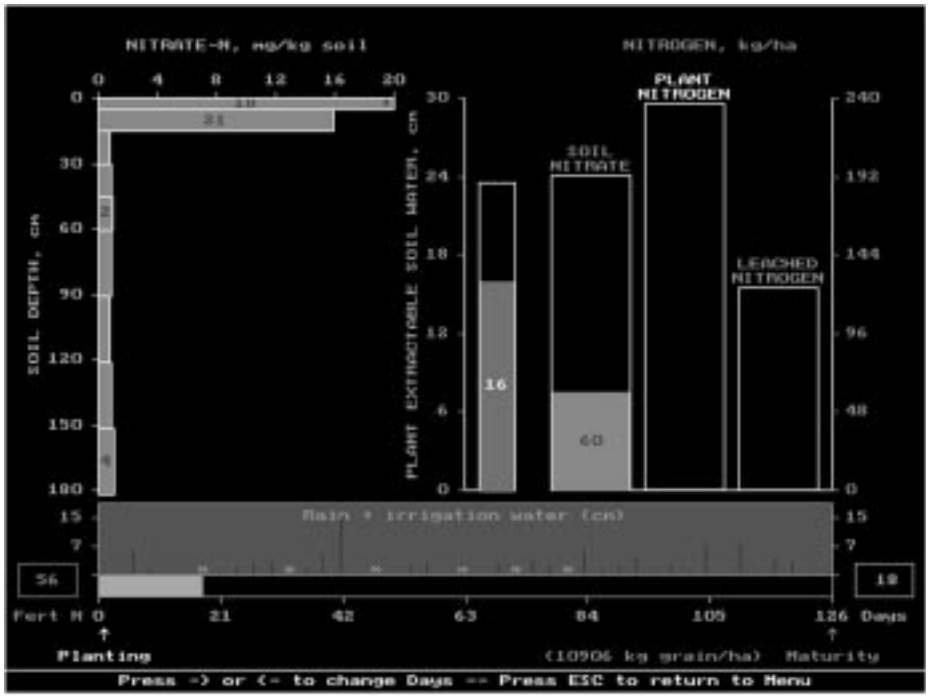
After selecting a treatment to plot, a screen is displayed summarizing the treatment selected (Screens 4, 5). Pressing the right arrow key (->) or the left arrow key (<-) displays the flux of water and N in the soil profile through time. Screen 4 shows a snapshot of conditions on day 18 of the simulation. Note the high levels of nitrate in



SCREEN 2.

layers 1 and 2 from the fertilizer application on that day. At this point, there has been no N leached from the soil profile into the groundwater. Screen 3 shows the same simulation on day 108 of the simulation. There are several things of note in this

Screen 3.



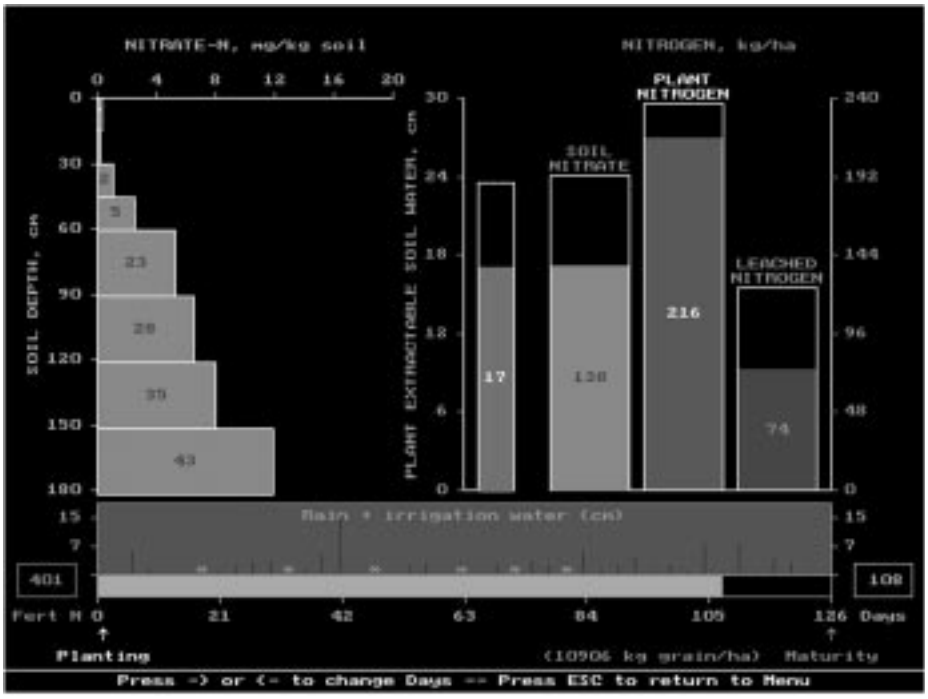
Screen 4.

graph. First, plant nitrogen uptake has reached a level of 216 kg/ha; second, nitrogen leached is at 74 kg/ha. Note the large rainfall event on day 105; at this point (day 105) there was only 58 kg/ha N leached. N-SHOW in this case illustrates clearly how N is lost from the profile under conditions of high N and excessive rainfall.

An analysis and description of the elements displayed on the screen follows. The slide bar at the bottom of the screen represents time. The displayed period starts at the first day of the simulation run and extends to the last day, which in most cases represents the harvest of a crop. The number in the box on the right shows the time (in days) corresponding to the position of the slide bar, expressed as days after the start of simulation.

The blue vertical bars placed above the time bar represent the combined total of rain and irrigation water received during the period. The green stars indicate when fertilizer N was applied. The number in the box to the left shows the cumulative amounts of N fertilizer applied at any point in time.

The plot on the left-hand side of the display shows nitrate concentration in the soil layers. The center vertical plot shows plant extractable soil water in the profile. The three bar charts on the right show nitrate-N in the soil, plant N, and total leached N.



Screen 5.

After you have finished viewing a plot, you can press the escape key and a summary screen is displayed (Screen 6). You may then access additional treatments to graph or exit N-SHOW via the menu bar.

Screen 6.

REFERENCE

Cabrera, M. L. 1994. N-Show: an educational computer program that displays dynamic graphs of nitrogen in soil. Journal of Natural Resources and Life Sciences Education 23(1):43-45.

Volume 4-5

Creating Experimental Performance Files

P. Wilkens
G. Hoogenboom

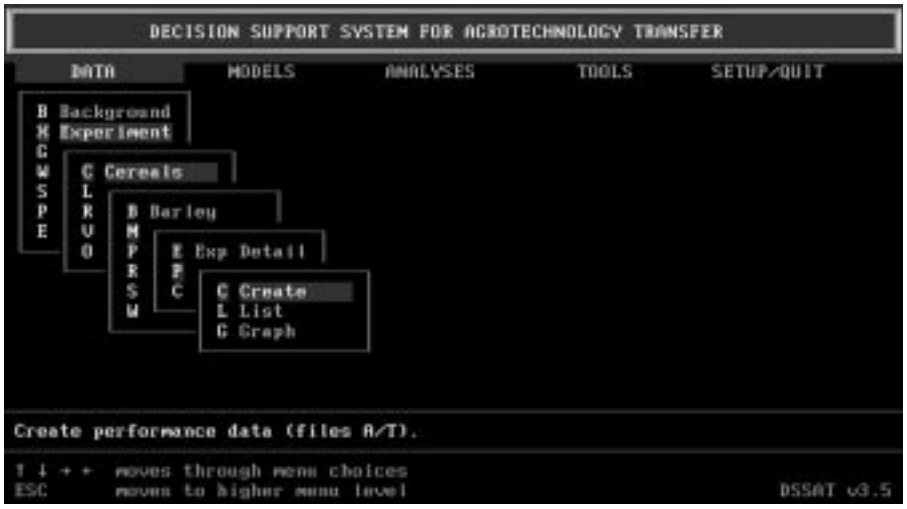
International Fertilizer Development Center
The University of Georgia
International Consortium for Agricultural Systems Applications

CHAPTER ONE. INTRODUCTION

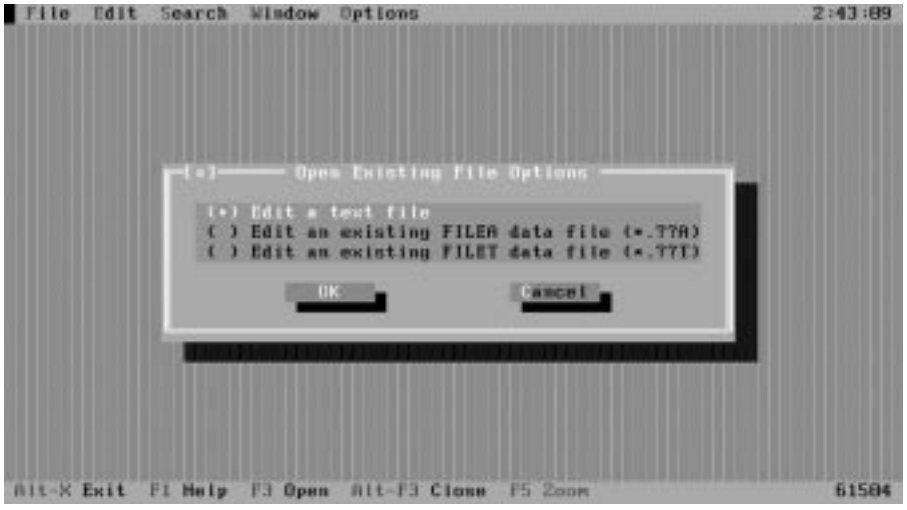
The creation of the summary average file FILEA and the time series file FILET is necessary for calibration and evaluation of the DSSAT v3.5 crop simulation models. FILEA contains the average performance of each treatment in an experiment (such has harvest yield, grain weight, etc.). FILET contains the measured time course data over the time course of the experiment (such as weekly biomass measurements). These files are especially important when calibrating new varieties, either through sensitivity analysis or with GenCalc (Hunt and Pararajasingham, 1994).

CHAPTER TWO. OPERATION

A utility has been developed for DSSAT v3.5, called ATCREATE. It can be accessed through the "Data-Crop-Performance-Create" in the DSSAT v3.5 shell (Screen 1). Note: This version is a developmental version with limited capabilities. Updates, when available, will be posted at the ICASA web site - http://agrss.sherman.hawaii.edu/icasa.

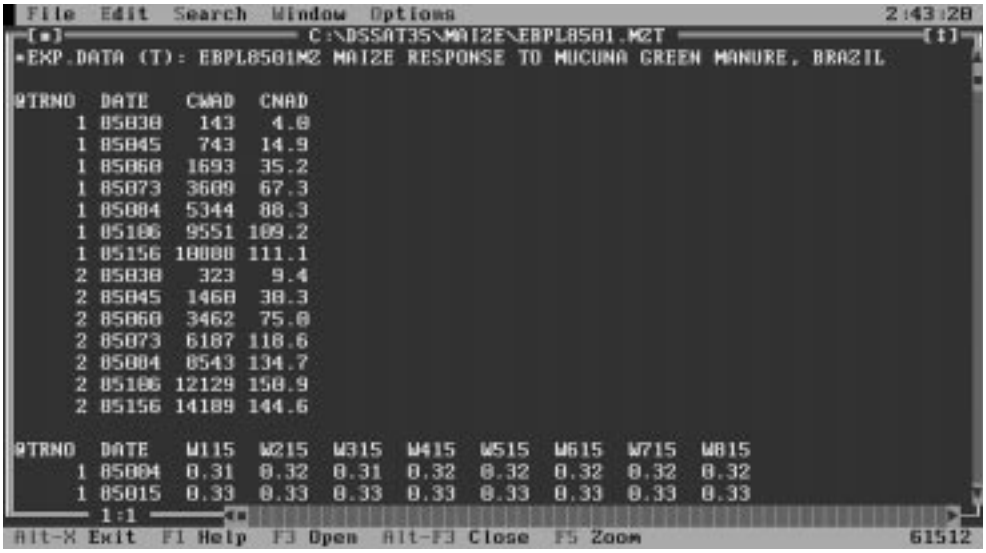


SCREEN 1.

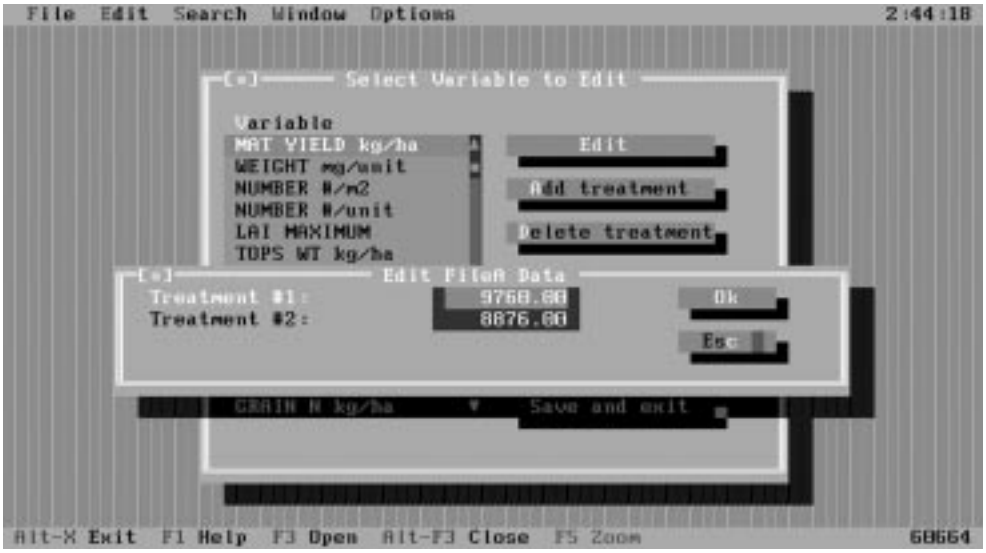


SCREEN 2.

The menu structure is accessed via the keyboard or with a mouse. Under the file menu, options are available to "Open" an existing file (see Screen 2) or create a new file. By opening an existing text file or an existing FILET, the default built-in editor is invoked (Screen 3). From this editor, you can create any observed data file you choose, including the ability to open more than one file simultaneously and to cut and paste



Screen 3.



Screen 4.

between files. It is important to maintain the ICASA standard for the data files regarding spacing and labeling to avoid spurious results (Jones et al., 1994).

If you elect to edit an existing FILEA, a dialog box displaying all of the observed data variables from the file is presented (Screen 4). By clicking on the appropriate button, you can edit the data, add or delete treatments, and add or delete variables. After modifying the data of your choosing, save the edited data file.

As noted above, ATCREATE is not yet complete. Enhancements being planned for ATCREATE include a full implementation of the ability to edit both FILEA and FILET files in their native format (rows and columns) rather than as a text file; rigorous data checking for data errors; and a flexible capability to read "foreign" format data files (spreadsheet or other ASCII formats) directly into the editor.

REFERENCES

Hunt, L. A. and S. Pararajasingham. 1994. Genotype coefficient calculator. *In:* G.Y. Tsuji, G. Uehara, and S. Balas (eds.). DSSAT v3. Vol. 3. University of Hawaii, Honolulu, HI.

Jones, J.W., L.A. Hunt, G. Hoogenboom, D.C. Godwin, U. Singh, G.Y Tsuji, N.B. Pickering, P.K. Thornton, W.T. Bowen, and J.T. Ritchie. 1994. Input and output files. *In:* G.Y. Tsuji, G. Uehara, and S. Balas (eds.). DSSAT v3. Vol. 2. University of Hawaii, Honolulu, HI.

Volume 4-6

Concepts for Calibrating Crop Growth Models

K.J. Boote

University of Florida
International Consortium for Agricultural Systems Applications

CHAPTER ONE. INTRODUCTION

Assume that you have a crop model that was developed elsewhere and you would like to calibrate that model for a new cultivar in a region where that crop has not been planted before or where the model has not been tested. Or maybe you may even want to adapt the CROPGRO model or one of the other models in DSSAT for a new species. Where do you begin?

We propose a systematic approach to calibrate a model for a new crop, new cultivar, or for a particular data set. The procedure assumes access to good growth analysis data, phenological observations, weather data, and soils data.

CHAPTER TWO. DATA REQUIREMENTS

Requirement for In-Season Growth Analysis Data

The first and foremost requirement is access to good growth analysis data, phenological observations, and the necessary weather and soil data required for running the model. Growth studies should be from trials conducted under well-watered conditions to minimize constraints due to water deficits. Information on the timing of phenological stages, particularly anthesis and maturity, are needed to calibrate the developmental timing. In-season information on dry matter accumulation in various plant parts and leaf area index is very helpful to calibrate aspects of photosynthesis, leaf expansion, termination of leaf expansion, and intensity and duration of reproductive growth.

With in-season information, you can evaluate state variables over time in comparison to observed values. The criteria for agreement between simulated and observed outcomes probably should be whether the model stays within one standard deviation of the mean of the data. Evaluation is a key to identify which model parameters and relationships are incorrect.

What if You Have Only Final Yield Information?

What if the only observed data collected is on final seed yield, seed size, and maturity and there is no data describing growth over time? A few parameters defining attributes such as total life cycle and seed size may still be calibrated. However, there are probably a half-dozen different coefficients that can be adjusted and will lead to the same final seed yield. Calibration in this way can be unreliable. The GENCALC model attempts to improve this situation by including anthesis date and biomass at maturity as input data, traits not routinely measured by plant breeders (Hunt and Pararajasingham, 1994).

"Species" File Information From Literature and Model Developers

Before starting, realize that the "species" file for the model already contains a large amount of information that model developers were able to obtain from the literature and from their calibration experiences. This includes definitions of conversion costs to synthesize tissue, typical composition of various tissues, initial protein concentrations of N-fertilized crops, minimum protein concentration of senesced tissues, cost of N assimilation and N-fixation, initial plant conditions upon emergence (initial weights of leaves, stems, and roots, and leaf area which depend on seed size), initial rooting depth, rate of

root depth progression as a function of physiological days, and many other coefficients. Not all coefficients are resolved conveniently from the literature; up to half of them were developed by the modelers from model calibration experiences. A typical example is partitioning as a function of vegetative stage (V-stage) and reproductive stage (R-stage), used to allocate dry matter among plant parts, and the rate of increase in height and width of the crop canopy as a function of V-stage. Other parameters were developed by optimization to data sets, for example the base and optimum temperatures for reproductive development were derived by Grimm et al. (1993) for soybean. Lastly, there are a number of functions based upon rather incomplete information from the literature that describe the relationships of processes to environment, such as temperature sensitivity of photosynthesis, N-fixation, respiration, leaf area expansion, pod addition, seed growth rate, partitioning intensity, and so on.

The author suggests a conservative approach to changing "species" parameters be taken by making changes only if there is a solid basis for it. While changes are necessary to adapt the model to a new species, do not modify the "species" files just to calibrate a new cultivar. Calibrating the "species" parameters may "uncalibrate" the model for other cultivars, unless the problem persists across several cultivars and soil types. The focus in this chapter will be on calibration of the parameters in the "cultivar" file, "ecotype" file, and "soil fertility" factor.

CHAPTER THREE. SYSTEMATIC APPROACH AND ORDER FOR CALIBRATION

A systematic approach is proposed for model calibration that evaluates parameters the model developers know the model is sensitive to. For example, think of the problem as involving a number of unknowns or "degrees of freedom" that affect dry matter accumulation and yield. Rather than randomly changing various parameters, we want to reduce the number of degrees of freedom by knowing where to start, what sequence to follow, and what is really most important to change. In effect, these are like expert system rules to follow. The steps shown below apply to calibration of the CROPGRO model only. Calibration for other models will be similar with respect to the systematic approach, but different with respect to the parameters that are involved.

STEP 1. CROP LIFE CYCLE

The first step should be to simulate crop development (flowering date and maturity date) of the cultivar being calibrated using the actual weather data. Start by selecting the general Maturity Group of the cultivar (or market type for peanut, or ecotype for common bean). Assume initially that the critical short daylength (CSDL) and slope of photoperiod sensitivity (PPSEN) are correct for that general or generic type. Adjust the EM-FL parameter (photothermal days between plant emergence and flower appearance) until the simulated date of flowering matches the observed. Then, adjust SD-PM (photothermal days between first seed and physiological maturity) until the simulated date of maturity is correct. If either of these two parameters is grossly larger or smaller than values for other cultivars, the cultivar will need to be made more or less sensitive to daylength (CSDL or PPSEN). Maturity date is also affected by FL-SD, but DO NOT change FL-SD at this time.

STEP 2. DRY MATTER ACCUMULATION

The next step involves simulating rate of dry matter accumulation and comparing simulated to observed values. Also, look at predicted versus observed leaf area index (LAI) and specific leaf area (SLA). Don't worry about final seed or pod yield yet, although when photosynthesis is low, there will be a general underprediction of biomass, LAI, seed weight, and seed number. If the slope of model predicted-dry matter accumulation is too rapid or too slow, adjust the parameters that affect leaf and canopy photosynthesis. The two primary ones to adjust are SLPF (soil fertility function in the soils file) and LFMAX (leaf photosynthesis rate in the cultivar file).

SLPF should be adjusted to give the correct slope to dry matter accumulation, under the assumption that site differences in fertility have created the difference. Adjust LFMAX only if there is data showing two cultivars that give different rates of dry matter accumulation when grown on the same soil.

Figure 1 illustrates that increasing SLPF from 0.92 to 1.00 gives a better fit to total biomass and pod weight for the 1984 irrigated peanut treatment. The field was in a good rotation and had a good soil fumigant program as well. Did this make a difference? An SLPF of 0.92 worked fine for ten other peanut experiments at Gainesville.

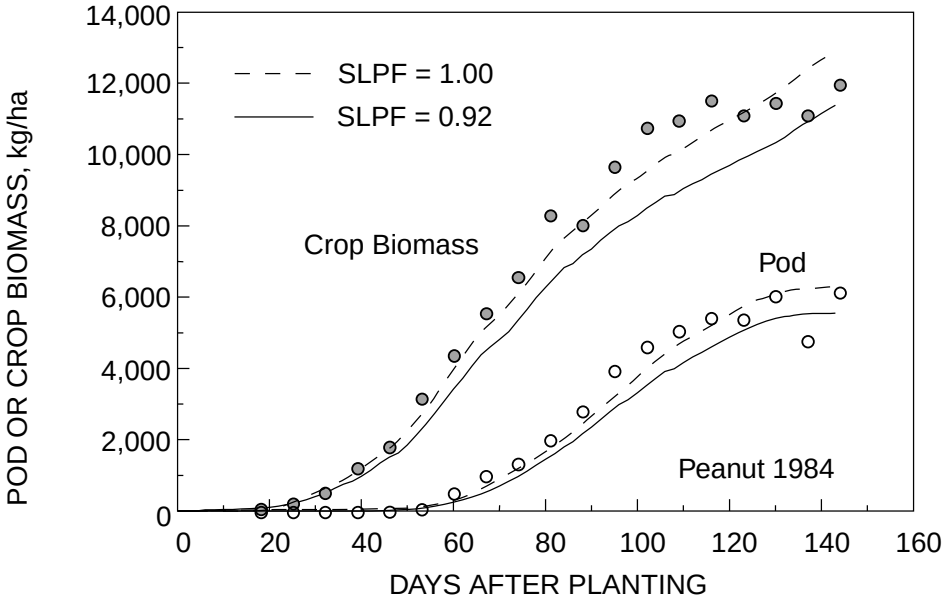


Figure 1. Crop and Pod Biomass as a Function of Days After Planting for Two Different Values of the Soil Fertility Function (SLPF).

SLPF can be adjusted with the Management Sensitivity Analysis Options, by selecting 5. for Soil, "select" 1 for your soil, but don't change soils. Then select option 13. "Growth reduction factor due to poor soil fertility". Adjust this to temporarily change the value of SLPF.

Step 3. Leaf Area Index and Specific Leaf Area (SLA)

Several "cultivar" parameters have a small impact on dry matter accumulation via their effect on LAI and light interception. These include specific leaf area (SLAVR), time to cessation of leaf area expansion (FL-LF), SIZLF (early "sink" limited leaf area expansion), and timing of pod addition (FL-SH) and seed growth (FL-SD). If you have data on LAI and SLA, plot these. If predicted SLA is too high, decrease SLAVR.

If SLA is too low, increase SLAVR. Now LAI should be closer to the observed (at least if leaf weight was correctly predicted).

Look at LAI next. When does simulated LAI reach its peak relative to observed data? Increase FL-LF if you want to delay the modeled peak in LAI and obtain a greater total LAI. The "ecotype" file contains a relative width (RWDTH) or relative height (RHGHT) variable that may need to be changed if the cultivar is a dwarf type. Incorrect prediction of LAI will give incorrect predictions of biomass accumulation, at least early in the life cycle when LAI is low. For example, if SLA is too high, early dry matter accumulation

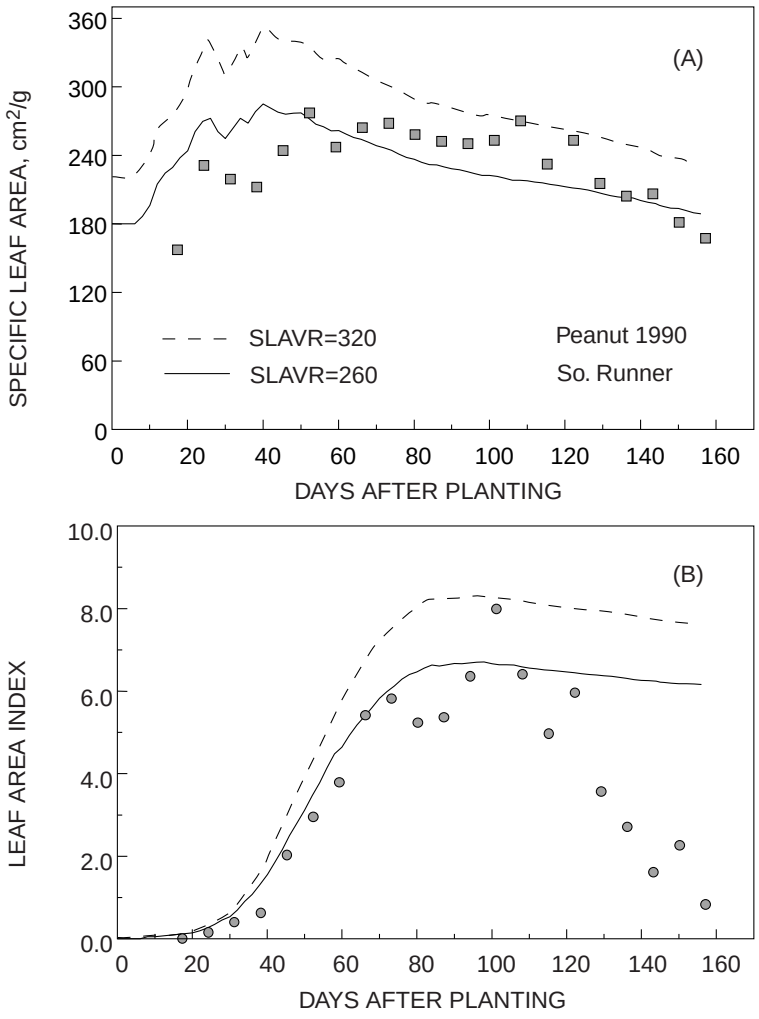


Figure 2. Specific Leaf Area (A) and Leaf Area Index (B) as a Function of Days After Planting for Two Values of the Specific Leaf Area Parameter (SLAVR).

may start faster than the real data. If SLA is low, the "lag phase" before rapid dry matter accumulation will be longer, yet the slope of dry matter accumulation will be almost the same. The parameter SIZLF can be used to limit leaf area expansion (and "hold down" SLA) during early V-stage development. The premise is that leaf area expansion is partially "sink-limited" during this phase.

Figure 2A and 2B illustrate the fit of SLA and LAI for the Southern Runner peanut cultivar grown in 1990, using an incorrect SLAVR of 320 compared to 260 cm^2/g. Note that using an SLA that is too large causes LAI to be overpredicted. The point of peak LAI is related to the value selected for FL-LF as well as when rapid pod growth begins.

Step 4. Re-Calibrate Dry Matter Accumulation

Use the calibrated SLA and leaf area timing aspects from step 3 and recalibrate dry matter accumulation as in step 2.

Step 5. "Species" Parameter Effects on Photosynthesis and Dry Matter Accumulation

"Species" parameters that affect photosynthesis and dry matter accumulation include the functions for temperature effect on photosynthesis, photosynthetic response to N concentration, rate of canopy expansion in height and width, degree of vertical layering of specific leaf weight in the canopy, partitioning to root, nodule growth and N-fixation parameters, and rate of protein remobilization from vegetative tissue.

Do not adjust these "species" parameters unless there is solid information on which to base these changes, as the changes will apply for all soils used and cultivars grown. The parameters influencing photosynthetic response to N concentration probably have the largest effect. These species parameters have their signature effects at different times and on different processes. The "photosynthetic response to N" parameters and N mobilization coefficients have more effect during the latter part of seed fill. Partitioning to root, rate of canopy expansion, and nodule growth traits, will impact primarily the early part of the life cycle. Temperature effects on photosynthesis will have more effect during cool times of year (during seedling growth or near maturity).

The signature effect of partitioning to roots is as follows. If too much assimilate is partitioned to roots early, then dry matter accumulation and leaf area development will appear to have a long lag phase. If less assimilate is partitioned to roots early, but more to leaves, then leaf area (and dry matter accumulation) will increase more rapidly and

have a short lag phase. If there is a water deficit, the plant with too little partitioning to roots will encounter earlier water deficit, which will cause a shift in partitioning to root.

Step 6. Initial Calibration for Seed Size, Seeds per Pods, and Seed Filling Duration

Set the correct seeds per pod and seed size. In the CROPGRO models, enter values for WTPSD (Maximum weight per seed, g), SDPDV (average seeds per pod), and SFDUR (seed filling duration for pod cohort, photothermal days) in the "cultivar" file, so as to match the final seed size and seeds per pod. The seed filling duration is often known from the literature. SFDUR should be proportional to but less than SD-PM, the duration from beginning seed to maturity. For the CERES models, final seed size is set by parameters that define the growth rate per seed and the expected duration of grain fill. We will come back to seed size later, because the initial timing may be incorrect.

Step 7. Initial Timing and Initial Rise in Pod and Seed Dry Weight

Now adjust the timing from flowering to first pod (FL-SH) and the timing from first flower to first seed (FL-SD) and duration of pod addition (PODUR), to get the correct timing to the initial rise in pod dry weight and seed dry weight. Examining pod or seed numbers versus time may be helpful, but these typically have much more error in measurement and questions of defining a "pod." Generally, shorten FL-SH and FL-SD to start pod and seed growth sooner, and the opposite to delay. Be cautious with FL-SH and FL-SD, because the length of time from FL-SH to FL-SD sets the lag from start of pods to start of seeds for *all subsequent* cohorts of pods. Thus, trying to make FL-SH early, but keeping FL-SD late, may slow everything down because of the long lag phase. In general, move FL-SH and FL-SD together, in the same direction. The PODUR (duration of pod addition) parameter also has an influence on the initial rise in pod and seed mass. Creating a short pod adding phase (determinate plant) will independently cause a more rapid rise in pod and seed mass without the need to adjust FL-SH and FL-SD (see example below). The PODUR parameter can sometimes be evaluated with data on increase in pod or seed numbers over time.

Figure 3 illustrates how increasing the value of FL-SH and FL-SD, each by 5 photothermal days, results in an improved fit to timing of pod mass for the 1976 peanut experiment. Not shown but also similarly improved, was the fit to seed mass, average weight per seed, shelling percentage, and LAI (latter was increased). The genetic coefficient values for Florunner peanut are those giving the best fit for approximately 12 data sets, so creating a set of genetic coefficients for each of the 12 different growth analyses is not valid.

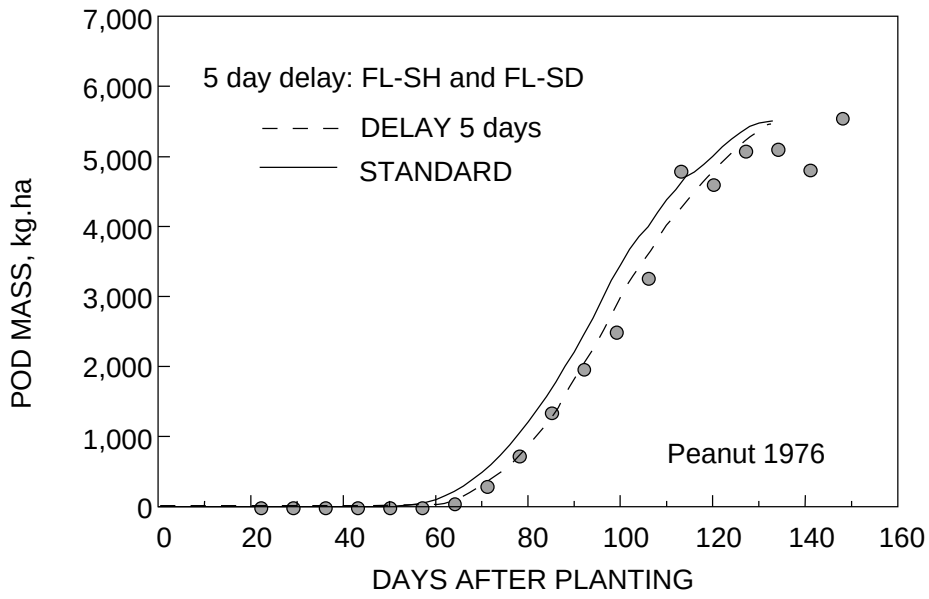


Figure 3. Pod Biomass as a Function of Days After Planting for a Short and Long Phase from Flowering to Beginning Pod (FL-SH) and Beginning Seed (FL-SD).

Step 8. Re-Calibrate Time From First Seed to Maturity

If FL-SD (time to beginning seed) was calibrated, adjust SD-PM, time between first seed and physiological maturity, in order to again correctly predict the observed date of physiological maturity.

Step 9. Seed Size, Shelling Percentage, and Seed Filling Duration

Now, go back to fine-tune final seed size and final shelling percentage which may have changed because the initial timing was changed. Increase or decrease WTPSD to correctly predict final seed size. Also, look at final shelling percentage. Remember that in the CROPGRO models, total seed growth will become limited as cohorts reach their maximum shelling percentage. To be safe, the final shelling percentage should be at or somewhat below the observed. If no data are available, be sure the predicted shelling percentage is a few units below the maximum THRESH value in the ecotype file. If predicted shelling percentage is too high, increase SFDUR (seed filling duration), and if it is too low, decrease SFDUR. The CERES models have no shelling percentage limit on seed growth or size.

Looking at average weight per seed and shelling percentage over time can be very helpful, if those observed data are available. For the CROPGRO models, if predicted shelling percentage and seed size over time are too high and reach a maximum long before maturity, then you will need to increase SFDUR (seed filling duration). Increasing SFDUR (keeping WTPSD the same) will reduce the slope of dry matter increase per average seed. Figure 4 illustrates, for the 1979 irrigated soybean study, how the average weight per seed can be graphically evaluated, to determine initial timing of seed set, slope (single seed growth rate), and seed filling duration. Increasing SFDUR from 22 to 25 days resulted in a better fit to the slope of seed growth (and shelling percentage).

Step 10. Re-Evaluate Total Dry Matter Accumulation and Relative Partitioning Between Vegetative and Reproductive Stages

Are seed size, shelling %, pod addition, and phenology timing correct? If so, re-evaluate dry matter accumulation in seed mass and total aboveground biomass, and the relative partitioning between seed and shoot mass.

If both seed and total biomass accumulation are too rapid or too slow (and the yield is high or low, respectively), then adjust SLPF or LFMAX to attain the correct slopes of aboveground dry matter and seed dry weight, as well as final seed yield. If the fit is good, you have completed your calibration.

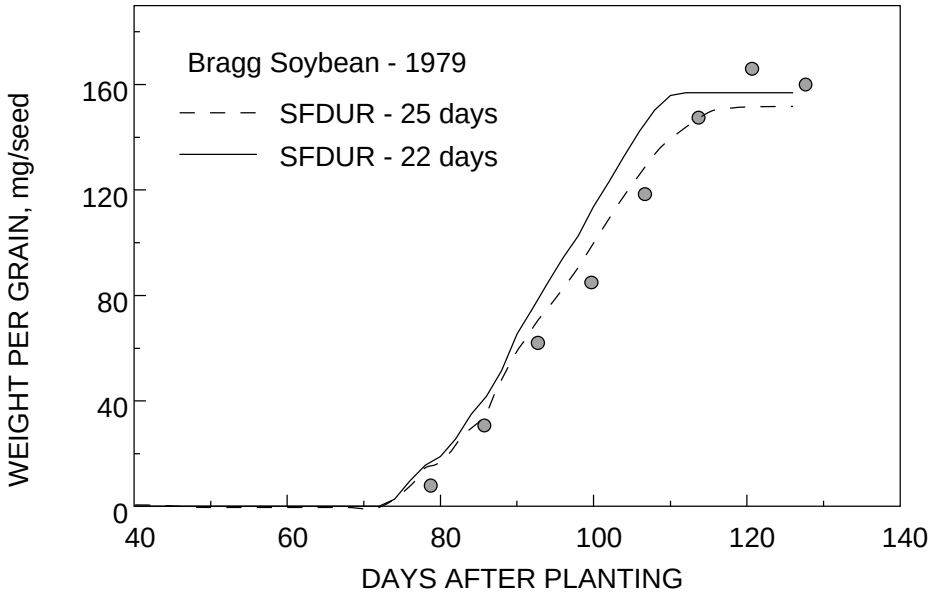


Figure 4. Seed Size as a Function of Days After Planting for Two Different Values of Seed Filling Duration (SFDUR).

Partitioning Factor for Indeterminate Crops

If, during reproductive growth of indeterminate crops such as peanut, the slope of total biomass accumulation is correct, but that of seed and pod is too steep and that of leaf and stem do not increase enough, then decrease XFRT (maximum fraction of daily growth that is partitioned to seed + shell). XFRT is generally assumed to be 1.0 for soybean and drybean, but older peanut cultivars/land races types/leafspot resistant types have lower partitioning. Some of these peanut cultivars actually continue slow vegetative growth (leaf and stem) during rapid pod fill. If pod or seed harvest index over time have been computed, calibrate partitioning (XFRT) to improve the prediction of the slope and final value of harvest index. By contrast, the initial timing of harvest index increase is dependent on FL-SH, FL-SD, and PODUR.

Possible Causes for Mismatch in Slopes of Biomass During Vegetative Versus Reproductive Stages

If the simulations cannot match the proper slope of biomass accumulation during both vegetative and reproductive growth phases, there are several possible problems that may relate to species file parameters. If biomass accumulation does not decline enough during reproductive growth, try to increase protein re-mobilization and leaf senescence. In addition to enhancing leaf loss, a more rapid protein mobilization will also limit photosynthesis more toward the end of seed fill. The NMOBMX parameter is in the "species" file. If protein mobilization and decline in photosynthesis during seed fill are correctly predicted, then there are other reasons for mismatched slopes of dry matter accumulation during vegetative and reproductive growth. The model may be partitioning too much or too little to root growth during early growth. If too little is partitioned to roots, then biomass accumulation will be overpredicted during the vegetative phase, yet the slope will be correct during reproductive growth (when we assume partitioning to roots goes to nearly zero anyway). Effect of incorrect maintenance respiration also has a unique signature. If the maintenance coefficient is too high, biomass accumulation will adequately predicted early, but underpredicted later in the life cycle.

CHAPTER FOUR. SUMMARY OF MODEL CALIBRATION

The calibration approach described above is systematic in the sense of starting at the beginning: learn what is known from the literature, ask what is known or calibrated by model developers (species file), describe initial conditions and compositions, predict life cycle timing, then calibrate traits to predict dry matter accumulation and LAI, roughly set seed size, then calibrate timing of pod and seed, then fine tune seed size and filling period and shelling percentage, then fine-tune dry matter accumulation in biomass and seed.

The Visual Art of "Solving" Unknowns for "Time Equations"

By making liberal visual use of the graphics option with the model runs, you are really solving for parameters to fit a number of "time equations", i.e. total mass versus time, LAI versus time, SLA versus time, seed mass versus time, pod mass versus time, seed size versus time, shelling percentage versus time, partitioning to vegetative vs. reproductive versus time, N concentration change over time (if such data is available). This is similar to solving for a number of unknowns with many equations, except we, as crop modelers, know that parameters should not be randomly modified, but are cognizant as to what parameter is most influential. While some of the "time equations" are somewhat interdependent, we think this approach reinforces the value for parameter estimation, of taking in-season measurements of leaf areas, plant parts, total mass, average seed size, N concentration, etc. over time. When you have only final yield, there are many (possibly incorrect) ways to modify model parameters to get to final yield.

Calibrating N-Fixation and N Uptake Parameters

These traits are best evaluated in experiments that compare N fertilized with unfertilized, and nodulated versus non-nodulated plants. The problem here is often one of poor quality information on initial soil conditions, and soil parameters needed for predicting N-supplying rate. Rooting data would also be nice to confirm that N uptake efficiency per unit root length works as intended in the model.

Just a note of caution. Both the CROPGRO and CERES models will benefit from continued improvements in N uptake functions, soil N supplying functions, and N effects on processes, and CROPGRO, by testing its approach for N-fixation. Particularly for CROPGRO, we have not had access to sufficient data to test these features

(other than data on N concentration, N-fixation, and nodule mass from Florida, Hawaii, and Iowa). Too often the model has been run assuming no N limitations.

Calibrating Parameters Influencing Response to Water Deficit

One should compare water-deficit treatments to well-watered treatments. Again, soil initial conditions and water supplying traits are very important to have; otherwise, the experiment is not very helpful. With water deficit treatments, you can evaluate relative rate of root depth progression and rooting profiles needed to create the observed soil water profiles or to maintain observed growth during stress. With sufficient and appropriate data, the relative sensitivity of leaf senescence to water deficit and the relative delay or acceleration of life cycle phases could be computed.

CHAPTER FIVE.
STATISTICAL APPROACHES FOR MODEL TESTING

An important objective in development of crop models is the determination of how well simulations predict observed phenomena. Calibration and validation are two different modes in which this determination of fit can be made. For calibration, one changes model parameters and even coding in order to obtain the best fit or prediction versus observed data. On the other hand, validation is the process whereby the model is run against independent data, without any modification of model parameters or code. Independent data is data collected in a different year or location, to which the modeler did not have access, for the original model development. Until a model predicts independent data well, one cannot say the model has broad validity to situations outside of the original data set for which it was developed. In the author's opinion, holding back some of the replicates of a data set for validation does not give you legitimate independent data. Independent data must be a different season or year.

Statistical criteria for determining goodness of fit for a simulation model include: 1) simulated line passing within a certain standard deviation about the observed data points, 2) minimizing sums of squares of deviations of predictions from observation, 3) maintaining a high r^2 of predicted versus observed, preferably with a slope of 1:1 and an intercept near zero.

The exact type of statistical criterion to use depends on the nature of simulated prediction. For example, you may want to test how well the model is predicting a point in time, such as the date of flowering (R1) or date of physiological maturity. The same applies to test how well the model predicts a given value, say yield (Y) at maturity. In these cases, you can determine the error sum of squares over *i* model runs with *i* different experiments or weather years:

$$SS = (\text{sim } Y_1 - \text{obs. } Y_1)^2 + \ldots + (\text{sim } Y_i - \text{obs } Y_i)^2$$

This sum of squares approach can also be used as an objective function by which an automated optimization procedure can systematically search for parameters to minimize the sum of squares. This approach was used by Grimm et al. (1993) to solve for temperature and photoperiod parameters of a phenology model to minimize the sums of squares about the observed and predicted start of flowering dates. A downhill simplex search technique was used.

Once the sum of squares is computed, a standard error of treatment mean and coefficient of variation can be computed as:

s_x = SQRT(SUM OF SQUARES/N)

CV = s_x * 100 / x

where N is the number of observations and x is the mean observed yield.

Another way of testing the above type of predictions, is to do a regression of predicted versus observed. With this regression, you will obtain an intercept and a slope (for simulated versus observed). It is important to have a wide range in the observed data to make valid conclusions. With sufficient data points, you should be able to compute the standard error of the estimated intercept, and the standard error of the estimated slope, as well as a residual standard error (*rse*) and an r value. Example graphs are shown in Figure 5 and Figure 6 from Huda (1988), where simulated yield is compared to observed yield at maturity for sorghum. Good predictability by the model is indicated if the regression slope is near 1.0, if the intercept is close to 0.0, and if the *rse* is low and the r value is high. A plot of predicted versus observed should fall near the 1:1 line (draw in a 1:1 line and the regression line). If the slope is significantly different from 1.0 (testable by t test and standard error of the estimated slope), then the model is either too sensitive or insufficiently sensitive to the environments represented by the observed data. An intercept significantly different from 0.0 is also a problem, because it signifies that the model would predict a positive or negative yield in an extremely poor environment when yield should be zero. If the slope is 1.0 and intercept is 0.0, yet the *rse* is high and r value low, then suspect natural variability in the field data.

A more complex approach is to determine how well the model predicts <u>all</u> of the in-season data points of growth analysis for a given treatment (maybe many treatments). In this case, then a different approach with sum of squares (SS) could be taken, where there may be one or several state variables (X, Y, Z) being fitted at one time. For just one variable (Y), the SS can be computed about the simulated and observed Y at each date (*i*) during the season. As before, a standard error of treatment mean and coefficient of variation can be computed.

SS = (sim Y_1 - obs. Y_1)^2 + . . . + (sim Y_i - obs Y_i)^2

In actual fact, the variance will differ as the amount of biomass, LAI, or numbers increase during the season. Thus, the use of some type of weighted sum of squares may be needed. This also would minimize the attempt of the model to fit outlying data.

Batchelor et al. (1994) used the above approach to solve for parameters of peanut pod detachment, where the objective function was based on predicted versus observed pod mass, but using all sampling dates, and over 18 experiments. They also used a cross validation technique, whereby one experiment at a time was left out of the optimization procedure, and parameters from 17 experiments were used to predict the experiment not included in the optimization. The process was repeated 17 more times to obtain mean square error of prediction for each experiment. Then, an average mean square error of prediction was computed.

A more complex situation could exist if one were to evaluate the SS associated with several state variables at one time. Young et al. (1979) used a weighted sum of squares

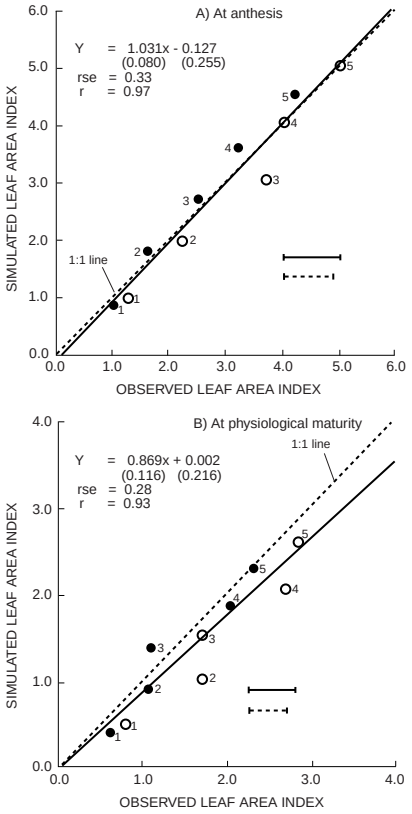


Figure 5. Relationship Between Observed and Simulated Leaf Area Indices at Anthesis (A) and Physiological Maturity (B) of Two Sorghum Cultivars (After Huda, 1988).

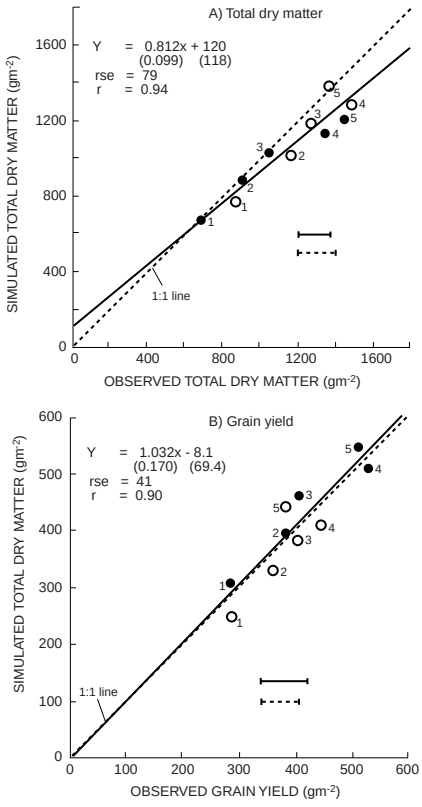


Figure 6. Relationship Between Observed and Simulated Total Dry Matter (A) and Grain Yield (B) of Two Sorghum Cultivars (After Huda, 1988).

approach based on three state variables (total biomass (B), pod mass (P), and flower count (F)) as part of an optimization package to develop the PEANUT model. They weighted or normalized the sums of squares from each state variable as follows:

$$SS = 1000\ n_1 \sum_{i=1}^{n_i} \frac{\left[(\text{sim } B_1 - \text{obs. } B_1)^2 + \ldots + (\text{sim } B_i - \text{obs } B_i)^2\right]}{(\sum_{i=1}^{n_i} \text{obs } B_i)^2}$$

$$+ \quad 1000\ n_2 \sum_{i=1}^{n_i} \frac{\left[(\text{sim } P_1 - \text{obs. } P_1)^2 + \ldots + (\text{sim } P_i - \text{obs } P_i)^2\right]}{(\sum_{i=1}^{n_i} \text{obs } P_i)^2}$$

$$+ \quad 1000\ n_3 \sum_{i=1}^{n_i} \frac{\left[(\text{sim } F_1 - \text{obs. } F_1)^2 + \ldots + (\text{sim } F_i - \text{obs } F_i)^2\right]}{(\sum_{i=1}^{n_i} \text{obs } F_i)^2}$$

The weighting done by Young et al. (1979) is not the same as the typical weighted sum of squares, but is a way of normalizing the sum of squares of a given state variable to the absolute magnitude of that state variable.

Young et al. (1979) used this sum of squares approach with an optimization package to develop parameters for his model. The author does not know of anyone who has used the approach to test how well a given model works for a new data set, but the principles should be the same. If such an optimization tool were to be used in the CROPGRO environment, the author suspects that one would need to be very careful in the decision of which parameters to adjust during the optimization. The author suggests to optimize on certain state variables to solve for specific parameters, whereas other parameters such as photosynthesis could be optimized against several state variables including total biomass and seed biomass. If the model has 10 to 20 parameters which can be varied, caution is needed to vary the correct parameter(s) in order to improve model fit. There is a real danger here in the sense of having many "degrees of freedom", i.e., parameters to adjust, but not knowing the correct one to modify.

Despite all this information relayed concerning statistical criteria and their use in a optimization environment, an experienced model developer can often look at the simulated curve relative to observed points and know just as quickly which parameter must be adjusted. The adjustment is made by trial and error. This works fine when dealing with a single experiment or treatment, but an optimization approach has real merit when dealing with 10 or more and upto 40 data sets.

REFERENCES

Batchelor, W.D., J.W. Jones, K. J. Boote, and G. Hoogenboom. 1994. Carbon-based model to predict peanut pod detachment. Transactions of the ASAE 37:1639-1646.

Grimm, S.S., J. W. Jones, K. J. Boote, and J. D. Hesketh. 1993. Parameter estimation for predicting flowering date of soybean cultivars. Crop Science 33:137-144.

Huda, A.K.S. 1988. Simulating growth and yield responses of sorghum to changes in plant density. Agronomy Journal 80:541-547.

Hunt, L.A., and S. Pararajasingham. 1994. Genotype coefficient calculator. p. 201-233. *In:* G.Y. Tsuji, G. Uehara, and S. Balas (eds.). DSSAT v3. Vol. 3. University of Hawaii, HI.

Young, Y.H., F.R. Cox, and C. K. Martin. 1979. A peanut growth and development model. Peanut Science 6:27-36.

Volume 4-7

Data Requirements for Model Evaluation and Techniques for Sampling Crop Growth and Development

K.J. Boote

University of Florida
International Consortium for Agricultural Systems Applications

CHAPTER ONE. INTRODUCTION

Before using a crop model for a particular production region, it is important that a minimum amount of crop growth and performance data be collected to allow evaluating model performance for that region's cultivar types, and in some cases for calibration of specific parameters. We will focus here on crop performance data (for comparison to model outputs). Readers are referred to other DSSAT material for needed input data for running the model, such as weather, soils, management, and genetic traits (Hunt and Boote, 1998; Jones et al., 1994).

CROPGRO, CERES and the other crop simulation models of DSSAT have a number of model state variables and rate processes which are available to output in order to compare with field measured values to allow validation of the model. Typical state variables available on a daily basis are dry weights (kg ha^{-1}) of leaf, stem plus petiole, podwall, seed, root, total crop weight, total mass of crop N, leaf N concentration, leaf area index, plant height, vegetative growth stage (V-stage), reproductive growth stage (R-stage), and numbers (per m^2) of pods and seeds. Simulated root length density and soil water content are available by soil layers. Typical rate variables of daily total photosynthesis, daily water uptake, drought stress, and daily light interception are also available. Although many of the above variables are not routinely output, they are internally computed and could be output by simple code changes. The above variables, final seed yield, harvest index, and seed size are available at final harvest maturity.

CHAPTER TWO.
LEVELS OF MODEL EVALUATION

What is the minimum set of crop variables needed for model validation and how should they be sampled? The answer to this question depends on your objective and how well you want the model to be tested. We propose three levels of model testing and evaluation.

What is model validation? Although some say one can only "invalidate" a model, a practical definition is that validation is the comparison of modeled outputs to field-measured variables collected in experiments not used to develop or calibrate the model. A more appropriate term to use might be model evaluation.

Model Evaluation

One must distinguish between "informational" model testing and model development. For model testing, the model inputs (weather, management, soils, and varietal information) must be collected or obtained.

As a minimum, varietal information on the final yield and seed size should have been collected. Of course, this means if the model incorrectly predicts the final yield, it is not possible to determine the cause. Adding three or four minimal growth samples over the growing season would help considerably in determining the cause resulting in differences between predicted and observed yields. Taking even more frequent and intensive samples would be needed to develop the model, especially with reference to physiological traits such as pod addition and maturation. For these reasons, we propose three levels of model testing and evaluation:

- **Informational Model Testing**
  (Presumes model has been developed).
  A. Model inputs (weather, management, soils, and genetics)
  B. Final seed yield, seed size, shelling %, harvest index.

- **Minimal-Model Validation**
  (Collecting enough in-season data to check model performance for a new region or cultivar.)
  A. Model inputs (weather, management, soils, and genetics)
  B. Final seed yield, seed size, shelling %, harvest index.
  C. Minimum plant growth measurements (sampling at 2 to 4 week intervals, or at selected growth stages as recommended in the appendix). Mass of plant parts

must be determined on oven dried weights at 60° C. Express mass and numbers on a land area basis.

1. Leaf weight
2. Stem-plus-petiole weight
3. Pod weight
4. Vegetative and Reproductive stages

- **Maximum-Model Calibration or Detailed Validation.**
  A. Model inputs (weather, management, soils, and genetics)
  B. Final seed yield, seed size, shelling %, harvest index.
  C. More intensive and extensive plant growth measurements (sampling at 7 to 10 day intervals). Mass of plant parts must be determined on oven dried weights at 60° C. Mass and numbers to be expressed on a land area basis.
     1. Leaf area index
     2. Leaf weight
     3. Stem-plus-petiole weight
     4. Podwall weight
     5. Seed weight
     6. Number of pods (>0.5 cm for soybean; size for other crops will vary)
     7. Number of seeds (also compute average mass per seed)
     8. Root mass or root length (optional)
     9. Vegetative and Reproductive stages
     10. Crop Canopy Height and Width

Pests

In all of the above sampling, certain information on pest parameters should be taken to evaluate their impact even if you are not doing an intensive analysis of pest effects. Typical information are suggested here for four types of "damage" or growth reduction caused by pests:

- Record percent defoliation caused by insect defoliators,
- Record species, heights, and weights per unit land area of weeds (shading pests),
- Record disease progress according to standard disease rating scales, hopefully representative of the whole canopy,
- Record either numbers or consequences (root damage, etc.) caused by soil-borne pests and nematodes. First model comparisons (for model development) should be made under conditions not limited by water deficit or soil fertility.

Water and Nitrogen

If testing includes modeled soil water balance under rainfed conditions, then volumetric soil water contents and root length density should be measured. Likewise, if testing involves modeled N balance aspects, then data from tissue analysis for N concentrations, and initial soil nitrate, ammonium, and organic matter for various soil layers should be collected.

CHAPTER THREE. HOW TO COLLECT DATA ON CROP GROWTH TRAITS FOR MODEL EVALUATION

The following large area sample plus subsample approach is proposed as a way to minimize work load and to minimize sample variability. It is based on techniques the author learned while working with W. G. Duncan and D. E. McCloud in the late 1970s at the University of Florida (Duncan et al., 1978). The principle proposed includes taking a large sample (about 1 m^2 land area) for land-area dependent biomass to reduce the coefficient of variation (CV), and then taking a small subsample (1 to 3 plants, 1 plant is enough) for plant internal traits which have low CV's. The CV of the computed trait, e.g., leaf weight (g $m^{-2)}$ or leaf area index will be only slightly greater than the CV for land area dependent samples.

Subsampling and Ratio Technique

The subsampling and ratio technique described here results in lower CV's than sampling random plants for mass, leaf area, etc. and multiplying by plant population.

Step 1. A minimum of four field *replications* are recommended in a randomized complete block design. Sampling within each replication can be either systematic (say from the northeast end of the plot and going to successive positions) or random (pre-locate sampling areas after emergence and randomly select among them). Realize that all sampling locations must be completely bordered on all four sides for the entire season; random sampling causes more problems in this regard. The amount of border required is a function of plant height (the width of the border should be about the height of plants to reach a zero "advective" light-stealing effect). The width or importance of border is also a function of how long your future sample area will remain "unbordered."

Step 2. *Large biomass sample*: Sample 1 meter square of land area for total dry weight (g m^{-2}). This sample should have border rows on either side and plants on either end. A given length of row (say 1 m length) should be sampled using a 1-meter stake and moving the stake down the row at the sampling location until each end of the stake falls midway between two plants (mid-gap to mid-gap). Follow this procedure to reduce variability and to avoid the human tendency to squeeze more plants into the sample. Most crops compensate quite well for gaps of up to 0.3 m. If there are many gaps of this size between plants, do not use this plot for sampling. Stand uniformity is one key to good sampling data. For important growth analysis studies, overplanting and hand-thinning is worth the additional effort.

Step 3. *Subsample* at least one representative plant (not the largest, not the smallest, but representative one in height, etc. Absolute weight is not important, but equal height and light exposure are important.) On this subsample, measure:

A. Leaf area
B. Leaf weight
C. Stem and petiole weight
D. Podwall weight
E. Seed weight
F. Number of pods (fully expanded, having seeds)
G. Number of seeds
H. Vegetative stage
I. Reproductive stage

Subsample size is usually one plant for peanut and between 1 and 3 plants for soybean and dry bean. For small grain cereals, a larger number of subsampled plants should be harvested (say 6 plants, see cereal literature), being certain to take whole plant units rather than just partial tillers.

Step 4. *From the subsample, calculate the fractions* of plant dry weight in various components: leaf, stem and petiole, podwall, and seed. Compute the *specific leaf area* (ratio of leaf area to leaf dry weight). Compute the *specific pod number* or *specific seed number* (ratios of pod number or seed number to plant dry weight).

Step 5. *Compute weights or numbers of crop components per unit land area* using the biomass (g m^{-2}) sampled in step two above and the ratios computed from the subsample in step three above. If the subsample was part of the large area sample be sure to add its weight back into the biomass. Data below is shown as g m^{-2} for easy computation of LAI, but all weights for DSSAT models must be reported as kg ha^{-1} (multiply g m^{-2} by 10).

LAI = biomass (g m^{-2}) * fraction leaf * specific leaf area (m^2 g^{-1})
Leaf = biomass (g m^{-2}) * fraction leaf
Stem = biomass (g m^{-2}) * fraction stem
Podwall = biomass (g m^{-2}) * fraction podwall
Seed = biomass (g m^{-2}) * fraction seed
Pods m^{-2} = biomass (g m^{-2}) * specific pod number (# per g plant mass)
Seeds m^{-2} = biomass (g m^{-2}) * specific seed number (# per g plant mass)

Effect of Subsampling on Coefficient of Variations

The reason subsampling works well is because the CV's are large (14 to 20 %) for land area dependent traits like total biomass or final seed yield, but lower (4 to 14 %) for plant internal traits (fraction leaf, fraction stem, fraction seed, harvest index, seed threshing fraction, specific leaf area, ratio pod #/plant weight, and ratio seed #/plant weight. When two variables are multiplied, the variance of the computed parameter is dominated by the variance of the variable having the largest variance. Thus, if the CV for biomass is 16.5 % and that for fraction stem is 4.4 %, the CV for computed stem mass per unit land area may be 16.8 %.

The reason for the difference in CV's between land area dependent samples and plant internal traits is that the land-area dependent traits depend critically on how properly the sample was bordered and how well that sample represented that land area. Taking a randomly selected single plant for mass is a highly variable land area dependent sample, and averaging mass over several single plants and then multiplying by plant population does not reduce the CV much. The CV for land area dependent samples decreases as the sample size increases (Figure 1). This is why the CV for final seed yield on a 10 m^2 area may be as low as 10 %, thus the need to take a fairly large sample for biomass or final seed yield.

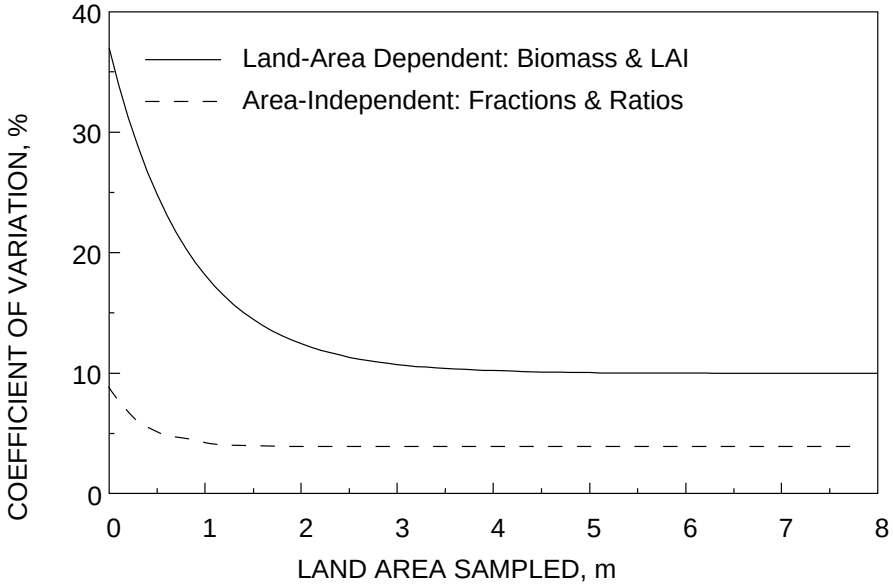


Figure 1. Effect of Land Area Sampled on the Coefficient of Variation.

The CV's of plant internal traits are low, because these traits are under greater genetic control and/or respond to "laws of plant allometry." For example, a plant twice as large will still have nearly the same ratio of leaf mass to plant mass or leaf area to leaf mass. The CV's for plant growth stage sampling are also quite low (7 % for V-stage and 1 to 3 % for date of R-stage). The same low CV's apply to other plant traits known to be under strong genetic control, such as percent protein in seed or weight per seed.

Sampling Crop Growth Stage

Information on crop growth stage is needed to develop or validate crop growth models. The vegetative stages are continuous in nature and sufficiently unimportant that they can be observed weekly during the season. By contrast, the reproductive stages are more important to dry matter allocation and they do not progress continuously. The R-stages should be observed on a 2-2-3 day schedule or on a 3-4 day schedule. Stages R2 (beginning of pegging for peanut) and R3 (beginning of pod formation) are the most critical to the CROPGRO models because they signify the start of pod addition and signify a shift in assimilate allocation to reproductive growth. The R4 stage (first full-size pod) is valuable because it occurs at a fairly defined time after R3 (6 to 12 days depending on species). The R5 (beginning seed) stage is valuable because it indicates the beginning of seed growth. Practically speaking, the R5, beginning seed stage, occurs very closely after the R4, first full-size pod stage, especially for peanut. A frequent sampling interval is especially helpful if you are developing relationships of V- or R-stage progression versus photoperiod or heat units accumulated. Severe drought can reduce rate of early and mid-season V- or R-stage progression by up to 50 %. Record observations of such conditions. On the other hand, drought during seed-fill will accelerate maturity, so try to quantify this response if your study has an irrigation variable.

Flower Counts and Aborted Pods

Numbers of flowers are less important to mass balance models, because the relative energy used to produce flowers is minimal. Moreover, the flower number of soybean is considerably in excess of the number of pods that can be carried by the daily photosynthetic production. Although the internal code of CROPGRO predicts flower production, we calibrate the model only to pod numbers. Thus flower number is of minimal value. For soybean, you may encounter the need to record the number of elongated pods (greater than 0.5 cm) which abort after the plant achieves its full pod load. These pods should definitely not be counted as good pods with seed, thus you may need to establish this category. You can distinguish aborted pods because they turn yellowish and begin to dry early. If you count pods as anything over 0.5 cm length for soybean and dry bean, expect the number to peak and decline.

Measuring Light Interception

Light interception is commonly measured with a line quantum sensor, and most typically measured within 2 hours of mid-day. One measurement of incident photon flux is made above the canopy followed by two measurements under the crop canopy and one more measurement of incident light above the canopy. To compute light actually absorbed, invert the line sensor and measure reflectance from the canopy. Remember that the sensor is 1 m long, thus for row spacing less than 1 m, the sensor must be angled under the canopy to go from mid-row to mid-row.

REFERENCES

Duncan, W.G., D.E. McCloud, R.L. McGraw, and K.J. Boote. 1978. Physiological aspects of peanut yield improvement. Crop Science 18:1015-1020.

Hunt, L.A., and K.J. Boote. 1998. Data for model operation, calibration, and evaluation. p. 9-39. *In*: G.Y. Tsuji, G. Hoogenboom, and P. K. Thornton, (eds). Understanding Options for Agricultural Production. Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Jones, J. W., L. A. Hunt, G. Hoogenboom, D. C. Godwin, U. Singh, G. Y. Tsuji, N. B. Pickering, P. K. Thornton, W. T. Bowen, K. J. Boote, and J. T. Ritchie. 1994. Input and output files. p. 1-93. *In* : G.Y. Tsuji, G. Uehara, and S. Balas (eds). DSSAT version 3, Volume 2. University of Hawaii, Honolulu, HI.

Volume 4-8

Field and Laboratory Methods for the Collection of the Minimum Data Set

R.M. Ogoshi
B.G. Cagauan, Jr.
G.Y. Tsuji

University of Hawaii
International Consortium for Agricultural Systems Applications

PREFACE

Field and Laboratory Methods for the Collection of the Minimum Data Set was developed as a quick reference guide for users who wish to evaluate the DSSAT models.

Field procedures and descriptions related to the layout of plots are only recommendations. We use illustrations to depict areas within experiment plots that the user would need to collect the complete minimum data set. The experimental design and the area available to the user will dictate the actual layout.

This document is a revision of IBSNAT Technical Report 2, Field and Laboratory Methods for the collection of the IBSNAT minimum data set, Department of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, University of Hawaii).

CHAPTER ONE. RATIONALE & PURPOSE: MODEL EVALUATION EXPERIMENTS

A model evaluation experiment can be an independent or adjunct part of any planned field experiment that will allow for the collection of the minimum data set. It should not differ greatly from agronomic experiments being conducted at research stations today. These will require monitoring the crop's phenological development, daily weather, and crop and soil management inputs. Because of the destructive sampling required for biomass harvesting, adjustments to plot size will most likely be necessary. **Plot sizes for modeling experiments tend to be larger than those for other agronomic experiments.**

It is anticipated that the complete minimum data set will permit comparisons of crop model simulation with observed field measurements at various intermediate phenological stages of growth and at final harvest. The DSSAT crop models are process-based growth models that compute photosynthesis on a daily basis and partition the supply of carbohydrates to the root, stem, leaf, and reproductive organs in ratios governed by the crop's phenological program. In principle, grain yield alone is sufficient to test the crop models. Nevertheless, biomass information is collected during stages of crop development as part of the minimum data set (MDS). With this information we hope to increase our understanding of the relationship between plant growth and development, and final yield.

This document attempts to answer many of the "how" questions associated with methods needed to collect the MDS. The procedures discussed should provide researchers with sufficient information to help them plan and manage experiments that will yield accurate minimum data sets in the most resource-efficient manner.

General Considerations for Modeling Experiments

A model evaluation experiment requires that measurements of dry matter production be taken at different phenological stages of growth, and thus necessitates the systematic destruction of plants within a finite population. Therefore, the variability of the population within the experimental unit must be minimized if the growth data are to represent the development of a crop from emergence to maturity. The many causes of variance in field experimentation can be grouped in the following general categories: mechanical errors, plant competition effects, and soil heterogeneity. The researcher must be able to cope

with each of these causes in order to ensure that a satisfactory level of uniformity among plants will exist within the plot.

Table 2 lists recommended or acceptable plot size dimensions, row spacing, plant density, etc. This list should be used only as a *guide*.

Despite all the precautions a researcher may take, some degree of variability will undoubtedly remain in the plant population within a plot as well as among plots of replicated treatments. The researcher must, therefore, develop a good sampling technique that will enable him to collect precise data to achieve the objectives of the experiment.

Sampling and processing of harvested plant materials are both labor intensive and time consuming. The researcher needs to critically evaluate the available resources and be able to fit the experiment within these resources with the overall goal of maintaining a high degree of precision at reasonable cost.

CHAPTER TWO.
FIELD METHODS

Daily Weather Data

In order to record daily weather needed for the MDS, reliable instrumentation located in a representative and unobstructed position is required. The following paragraphs provide information to help select, locate, and install these instruments.

Selection and Installation of Weather Instruments

The required minimum weather data are daily records of rainfall, maximum and minimum air temperatures, and solar radiation. Instruments used to measure these data should be monitored regularly and the data recorded daily. This is especially critical during the crop growth period—from the time the seed is sown (or when pre-plant soil samples are taken) until final crop harvest.

Selection

Automatic recording weather stations and manually read weather instruments are now commonly available. Either system is sufficient to collect the MDS. Battery-operated recorders must be closely monitored to ensure continuous recording of data, as failure can result in irreversible loss of data. Installation of a manually maintained station as a back-up to an automatic one is recommended to minimize loss of data.

The following table lists instruments and sensors normally available for either type of weather station:

Table 1. Weather Instruments & Sensors.

Parameter	Automatic	Manual
Rainfall	Tipping bucket, 1mm	Standard 8-inch rain gauge with metric dipstick
Air temperature	Thermocouple sensors	Max/min thermometers
Solar radiation	Silicon cell pyranometer	Sunshine hour recorder or pyroheliometer

For manually maintained instruments, it is essential to record events at nearly the same time each day. At readily accessible sites, the instruments should be read and

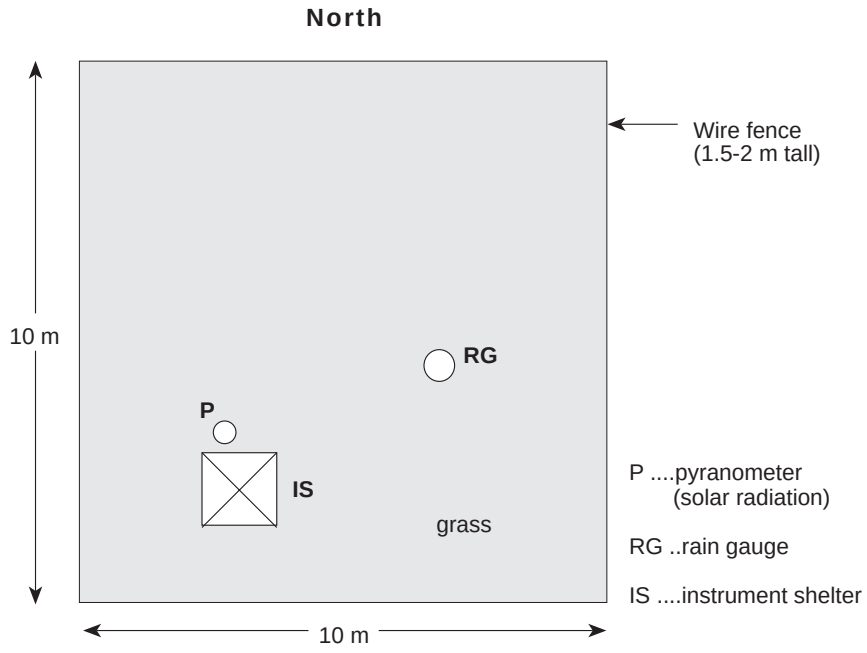


FIGURE 1. SUGGESTED LAYOUT OF INSTRUMENTS AT A WEATHER STATION.

data recorded as the first activity of the day. For remote sites, the instruments may be read at the end of the working day.

LOCATION AND LAYOUT OF WEATHER STATION

If an existing weather station that records the minimum weather data set is available at the experimental site, the distance from that station to the experiment plots should be noted.

If no weather station exists, select an area that is most representative of the experimental site. The weather station site should be on level ground and planted with grass. Avoid locations where natural or man-made structures may obscure sunlight. As a general rule, the weather station should be positioned at a distance at least five to ten times the height of the obstruction.

An example of layout and installation of a weather station is shown in Figure 1. A minimum area of 100 m^2 should be allotted for the installation of the station.

INSTALLATION

Max/min thermometers and thermographs should be housed in a well-ventilated instrument shelter. It is best to have the door to the instrument shelter open to the north or south to minimize direct sunlight on the thermometers. Height of the shelter should be at least 1.25 m above the ground. Legs of the shelter should be secured on a concrete footing or otherwise well-anchored to prevent vibration.

Rainfall

The rain gauge should be mounted on level ground or a level platform with a concrete footing. For tipping bucket rain gauges, the electrical lead to the event recorder should be sufficiently insulated and buried as a safety precaution. The event recorder should be housed in an instrument shelter.

Solar radiation

Solar radiation instruments, and sunshine duration recorders should be mounted on a well-anchored level platform. Both must be positioned in a location clear of any east-west obstructions or shadows.

Experimental Plots

Estimation of Dry-weight Crop Residue

Data on initial crop residue is reported as part of the minimum data set. It is used in the nitrogen and water balance subroutines of the crop models. The procedure for estimating this variable is described below.

1. Just before plowing, select ten 1 m^2 sections, uniformly distributed across the experimental area.

2. From each 1 m^2 sampling area, collect all aboveground residue and weigh. Determine the average weight of all ten samples. This is the average fresh weight measurement in the formula in Equation 1 (below). Be sure that residue does not include soil.

3. From each of the ten samples, obtain a subsample of 100–200 g. Weigh these, and place in a marked bag.

4. Oven dry each residue subsample at 70°C for 24 hours or more until a constant weight is obtained.

5. Determine average moisture content (M.C.) of each of the subsamples by the following:
   Equation 1:

$$\text{M.C.} = \frac{\text{fresh weight} - \text{dry weight}}{\text{fresh weight}}$$

NOTE: The value of M.C. should be <1.

6. Determine moisture content of the ten subsamples.

7. Express residue amount in kg/ha, on a dry basis by the following:
   Equation 2:
   Crop residue = (average fresh weight in kg/m^2) (1- average M.C.) (10,000). (kg/ha)

Soil Water Contents

Volumetric Water Content

The following procedure should be used to determine volumetric soil water content.

1. Collect a core sample or auger sample from each soil layer two weeks before planting.

2. Thoroughly mix and composite the auger samples in one bag for each layer.

3. Take a duplicate set of subsamples from the composite. Weigh (W_i) both sets of samples in soil moisture cans.

4. Oven dry the samples at 105°C for 24 hours or until a constant weight (W_d) is obtained.

5. Calculate the gravimetric moisture content (θ_g) as follows:
   $$\theta_g = (W_i - W_d)/W_d$$

6. Determine the bulk density separately and convert gravimetric moisture content into volumetric moisture content (θ_v), as follows:
   $$\theta_v = (\rho_b/\rho_w) \times \theta_g$$
   where ρ_b is the bulk density and ρ_w is the density of water equal to 1.

Bulk Density

Core Method

A core sampler can be a single cylindrical tube or a tube with several cylindrical rings of known dimensions. The sampler can be driven mechanically or manually into the soil either horizontally or vertically. Extruded soil from core sampler or soils and rings are weighed and oven-dried to constant weight at 105°C. The oven-dry weight divided by the volume of the core sampler or cylindrical ring is the bulk density usually expressed in g/cm^3. This method should not be used for wet (e.g., aquic, histic) or dry, sandy soils.

Clod Method

Natural clods about one inch in diameter are coated with saran and weighed in air, then in water, using the Archimedes principle to determine the soil volume.

Apparatus
Beam balance accurate to 0.01 g; beaker or container to accommodate clod while being weighed in water.

Reagent
Methyl ethyl ketone solvent; Dow Saran S310 resin (as an alternative, paraffin can be used). Saran-solvent ratios of 1:6 and 1:7 are used, dependent on porosity of the soil. Coating solution: fill a weighed container to 3/4 of its total volume. From the weight of solvent, calculate the weight of resin to obtain derived resin-solvent ratio and add to solvent. Mix vigorously under an exhaust hood; solution is highly volatile and flammable, so store in metal cans tightly closed to prevent evaporation.

Procedure
1. Tie approximately 30 cm of fine thread around each clod and weigh (W_f).

2. Dip in coating solution several times until the clods are fully coated. Allow to solidify between each coating.

3. Weigh coated clods in air (W_{cp}).

4. Weigh coated clods in water (W_{cw}) that are hanging from balance pan and suspended in water. Make sure that during weighing the clod is fully submerged in water but not resting at the bottom of the water container. Remove or dislodge air bubbles sticking to the side of the clod during submergence. If bubbles continuously appear out of the clod while submerged, the clod is not fully coated with the saran. Discard clod and start all over again with new clod.

5. Carefully peel off saran coating, weigh and dry clod at 105°C. Weigh dry clod and calculate moisture factor (MF) of clod as follows:

$$MF = 1 + \frac{\text{wet weight - dry weight}}{\text{dry weight}}$$

6. Calculate the clod dry weight:

$$W_d = \frac{W_f}{MF}$$

7. The bulk density, ρ_b, is then calculated by the following formula:

$$\rho_b = \frac{W_d}{\frac{W_{cp} - W_{cw}}{\rho_w} - \frac{W_{cp} - W_f}{\rho_p}}$$

where W_d = clod dry weight
W_{cp} = weight of clod and wax in air
W_{cw} = weight of clod and wax in water
W_f = weight of clod in air
ρ_w = density of water at temperature of determination
ρ_p = density of resin (about 0.9 g/cubic centimenter).

Reference

SCS, USDA. 1984. Soil survey laboratory methods and procedures for collecting soil samples. *In* Soil Survey Investigation Report No. 1. U.S. Gov. Print. Office, Washington, D.C.

Growth Stage and Yield Components

Field Measurements for Drained Upper Limit (DUL) and Lower Limit (LL)

Existing DSSAT crop simulation models make use of soil water parameters as a measure of plant-extractable water. The drained upper limit (DUL) is defined as the highest field measured water content of a soil after it has been wetted and allowed to drain until drainage has ceased. (i.e., less than 0.1 to 0.2% change in water content per day). DUL is analogous to the definition of field capacity. Another parameter is called the lower limit (LL) which is the lowest field measured water content of soil when plants have stopped extracting water (i.e., near premature death or become dormant due to water stress); the LL is analogous to the permanent wilting point. Traditionally, these parameters are measured in the laboratory with a pressure plate apparatus to obtain a relationship between the pressure applied and soil moisture retained under this pressure, a moisture release curve.

Procedure

The procedures for field measurement of soil water at DUL and LL are as follows:

1. Grow a crop in the field during a period when rainfall is least likely to occur. Apply sufficient fertilizer and irrigation water to allow the crop to attain full growth and normal development. A minimum plot size of 10 x 10 m is recommended. Recommended crops are maize, sunflower, sorghum, wheat, and barley.

2. At about two weeks before flowering or silking withdraw irrigation and let the plants extract water until they exhibit severe drought stress or are near death. *NOTE: It is important, during this period, to minimize or prevent any addition of water by rainfall. Plastic sheets or shelter should be used to cover the plot and prevent water from infiltrating the soil.*

3. Determine volumetric soil water content every 10- 20 cm to a depth of 2 m, if bedrock is not limiting. Soil water can be determined either by a neutron probe or gravimetrically with auger samples. A minimum of four sample locations for a neutron probe or 10 sampling holes for auger samples is recommended. A separate pit, as near as possible to the experimental plot, must be dug to collect soil samples for bulk density measurements.

4. After determining the lower limit (LL) of extractable water, remove plants and irrigate the plot until the entire profile is saturated. If sprinklers or drip irrigation is used, saturation can be estimated visually by observing ponding or a slowing of water intake in fine textured soils.

5. Cover plot with plastic.

6. Determine the water content at each depth or layer after allowing the soil to drain for several days until the water content becomes stable or changes from 0.1 to 0.2% one day. For coarse-textured soils, DUL is usually attained after 2 days, while in fine-textured soils 10-14 days may be required. This is called the drained upper limit (DUL).

If field measurements are not possible, data from moisture release curves may be used as a first approximation. Another alternative would be to use field measurements made elsewhere on a similar soil. Similar soils mean similarly classified soils that have either the same soil series name or same phase and soil family.

Growth State and Yield Components

Two types of sampling or harvest procedures are described in this section: (1) biomass harvest at recommended growth stages, and (2) final harvest to determine crop yield. Both types of harvest data are used in evaluating simulated crop model outputs.

In most instances, there will be three or four biomass harvests before the final harvest. In Table 2, the recommended physical dimensions of designated MDS plots are provided for each crop. The biomass harvests correspond to selected vegetative and reproductive stages of growth (Table 3).

TABLE 2. RECOMMENDED ROW DIMENSIONS AND HARVEST INFORMATION. (RESEARCHERS MAY MODIFY THESE DIMENSIONS AND NUMBERS AS NEEDED)

crop	no. of rows	row spacing (m)	plants/ m-row	plants/ha ($x10^3$)	no. of biomass harvests	biomass and harvest rows	biomass plot lengths (m)	final harvest plot lengths (m)	no. of plants	
									biomass harvest	final harvest
Maize	4	0.67	4	59.7	3	2	1.0	9.0	8	72
Sorghum & Millet	4	0.67	9	134	3	2	1.0	3.6	18	65
Wheat, Barley & Rice	15	0.2	50	2500	3	5	.50	1.5	125	375
Soybean & Dry Bean	4	0.67	18	269	5	2	1.0	5.0	36	180
Chickpea	4	0.30	10	333	5	2	1.0	5.0	20	100
Peanut	4	0.67	5	74.6	2	3	1.0	5.0	15	75
Cassava	8	1.0	1	10.0	5	6	1.0	1.0	6	6
Aroid	6	0.6	2	33.3	4	4	0.5	3.0	4	24
Potato	6	0.9	4	44.4	3	4	0.5	2.5	8	40

TABLE 3. GROWTH STAGES FOR BIOMASS HARVEST.

Crop	Harvest Number	Growth Code	Stages: Descriptions
Maize	1st	V6	50% of plants with collar of 6th leaf visible.
	2nd	R1	50% plants with some silks visible outside husks.
	3rd	R4	50% plants in "dough" stage—endosperm with pasty consistency—often 24-28 days after silking.
Sorghum, millet, barley, wheat, & rice	1st	13	3 leaves unfolded.
	2nd	65	Flowering half-way.
	3rd	85	Soft dough (fingernail impression not held).
Soybean, dry bean, cowpea, & chickpea	1st	V4	50% of plants with 4 nodes on main stem beginning with unifoliate node.
	2nd	R3	50% of plants with a pod 2.0 cm long at any node
	3rd	R5	50% of plants with beans beginning to develop at any node
	4th	R6	50% of plants containing full size beans at any node
	5th	R7	50% of plants with pods yellowing; 50% of leaves yellow–physiological maturity.
Peanut	1st	V4	50% of plants with 4 developed nodes on the main axis.
	2nd	R4	Full pod. 50% of plants with one fully-expanded pod, to dimensions characteristic of the cultivar.
	3rd	R6†	Full seed. 50% plants have one pod with seeds filling cavity of pod when fresh.
Cassava	1st	VE	50% of plants with at least 1 shoot greater than 1 cm in length on the planting stick.
	2nd	V1	50% of plants with at least 1 apical meristem.
	3rd	V2	50% of plants with at least 2 apical meristems.
	4th	V(n)	50% of plants with at least n apical meristems.
	5th	R1	50% of plants flowering.
Aroid	1st	V3	50% of plants have 3 fully opened leaves.
	2nd	V8	50% of plants have 8 fully opened leaves.
	3rd	V11	50% of plants have 11 fully opened leaves.
	4th	V25	50% of plants have 25 fully opened leaves.
Potato	1st	T1	Tuber initiation. 50% plants have at least one tuber greater or equal to 1 cm in diameter.
	2nd	T2	T1 + 20 days
	3rd	T3	T1 + 40 days

† NOTE: Two biomass harvests must be taken after R4: one taken 3-4 weeks after R4, and one taken 6-8 weeks after R4.

The recommended MDS plot is of sufficient size to accommodate the requirements of both biomass and final harvests. Final harvest procedure flow charts and plot diagrams in this section, provide harvesting guides and recommended plot and subplot size for each crop.

Procedure for Biomass Harvest

To harvest all aboveground biomass, cut the plant with a sharp knife as close to the ground as possible. Weigh harvested material, place material in a paper bag, and oven-dry at 70°C to constant weight. Since respiration can reduce plant weight, oven dry the harvested material immediately after harvesting, or if that is not possible, place in a cold container, or a refrigerator until it can be dried. For crops with thick stems or woody tissues, chop the stems into smaller pieces (2-5 cm) before oven drying. For further details, consult the following section before conducting biomass sampling during the reproductive (R) stages.

Procedure for Final Harvest

Steps for final harvest include harvestable plant identification, harvesting, and plant preparation and drying. Most crops have a unique set of data to be collected. The preparation of plant samples for each crop is described in the flow charts on the proceeding pages. Harvestable plant identification and harvesting procedures are similar for all crops and are also described on the flow charts.

1. Before harvest, the harvestable plants must be identified and harvest areas determined.

2. Harvestable plants adjacent to border plants within the harvest row should be tagged so as to clearly demark the harvestable row. Refer to the plot diagrams on the following pages for identification of border and harvest plants. Multiplying the total length of all harvested rows by the space between rows gives the harvest area of the plot.

3. Once harvestable rows are marked and harvest area determined, harvesting may proceed.

In harvesting, plants between and including tagged plants, should be collected. Cut all plants at ground level. Count the plants as they are being cut, and record the total number of plants harvested for plant population determination. To calculate plant population of the plot, divide the number of plants harvested by the harvest area. When all plants have been removed from the plot, preparation of plant samples may be executed according to the appropriate flow chart for that crop.

TABLE 4. FINAL HARVEST FOR EACH CROP.

Crop	Growth Code	Stages: Descriptions
Maize	After R6	50% of plants with black layer at the base of the seed. Also in the absence of a black layer, grains become shiny and translucent.
Sorghum & millet	92-94†	Soon after 80% of grains are hard and difficult to break with fingernails or teeth.
Rice, barley, & wheat	92-94†	70-80% of grains are brown which may or may not be associated with browning of the leaves.
Soybean, dry bean, cowpea, & chickpea	R8	95% of pods are dry and brown.
Peanut	R8	70-80% of pods have inner pericarp darkening (or show light-brownish color of mesocarp when outside of pod is scratched). Pods are very firm when pressed between thumb & forefinger. Leaves may or may not be senescing.
Aroid	R7	50% of plants have only 2 or 3 active leaves remaining on plant (approximately 10-11 months after planting).
Cassava	R7	50% of plants at harvest maturity.
Potato	T4	Green canopy cover reaches 20% of the maximum achieved (harvest maturity).

† *For the final harvest, determine the growth stage code (92, 93, or 94) which most closely corresponds to the growth stage of the plant at physiological maturity.*

Final Harvest Procedure for MAIZE

For each plot: Determine plants and area to be harvested. Cut plants in harvest area at ground level. Remove ears leaving husk intact on plant.

Ears

Calculate ear number (ears/m^2) by dividing number of ears by harvest area (12 · 06 m^2) and record. Enter total ear weight (undried).

Select a subsample of 10 representative ears for dry weight and seed number determination. Log ear subsample weight (undried).

Remove seed from cob of subsample and log seed subsample weight (undried). Count three sets of 100 seeds and note the average 100 seed weight (undried). Calculate and record seed number.

$$\text{seed no. (seed/m2)} = \frac{\text{ear wt. (undried)} \times \frac{\text{seed subsamp. wt. (undried)}}{\text{ear subsamp. wt. (undried)}} \times \frac{100}{\text{100 seed wt.}}}{\text{harvest area}}$$

Dry seed and cob subsamples to constant weight at 70°C (approximately 7 days). Log seed and cob subsample weights (dry) separately.

Seed

Calculate and record seed weight (dry).

$$\text{seed wt. dry (g/m}^2\text{)} = \frac{\text{ear wt. (undried)} \times \frac{\text{seed subsample wt. (dry)}}{\text{ear subsample wt. (undried)}}}{\text{harvest area}}$$

Cob

Calculate & record cob weight (dry).

$$\text{cob wt. dry (g/m}^2\text{)} = \frac{\text{ear wt. (undried)} \times \frac{\text{cob subsample wt. (dry)}}{\text{ear subsample wt. (undried)}}}{\text{harvest area}}$$

Stover

Divide number of plants by harvest area (12 · 06 m^2) and record under plant population of harvest area (plants/m^2).

Select a well-mixed, 500 g subsample of leaf blade. Log the leaf blade subsample weight (undried).

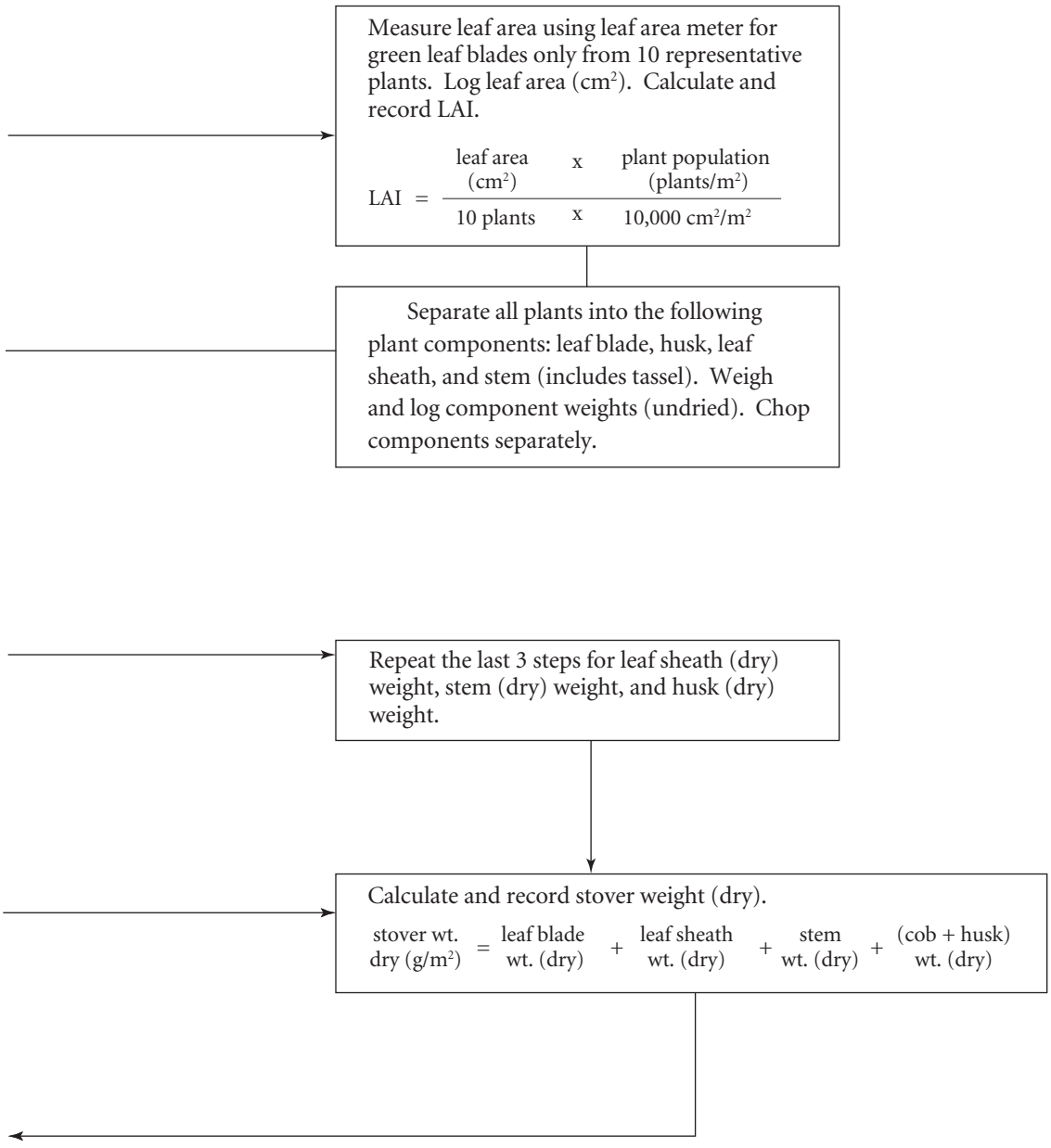
Oven dry the leaf blade to constant weight at 70°C. Log leaf blade subsample weight (dry).

Calculate and record leaf blade weight (dry).

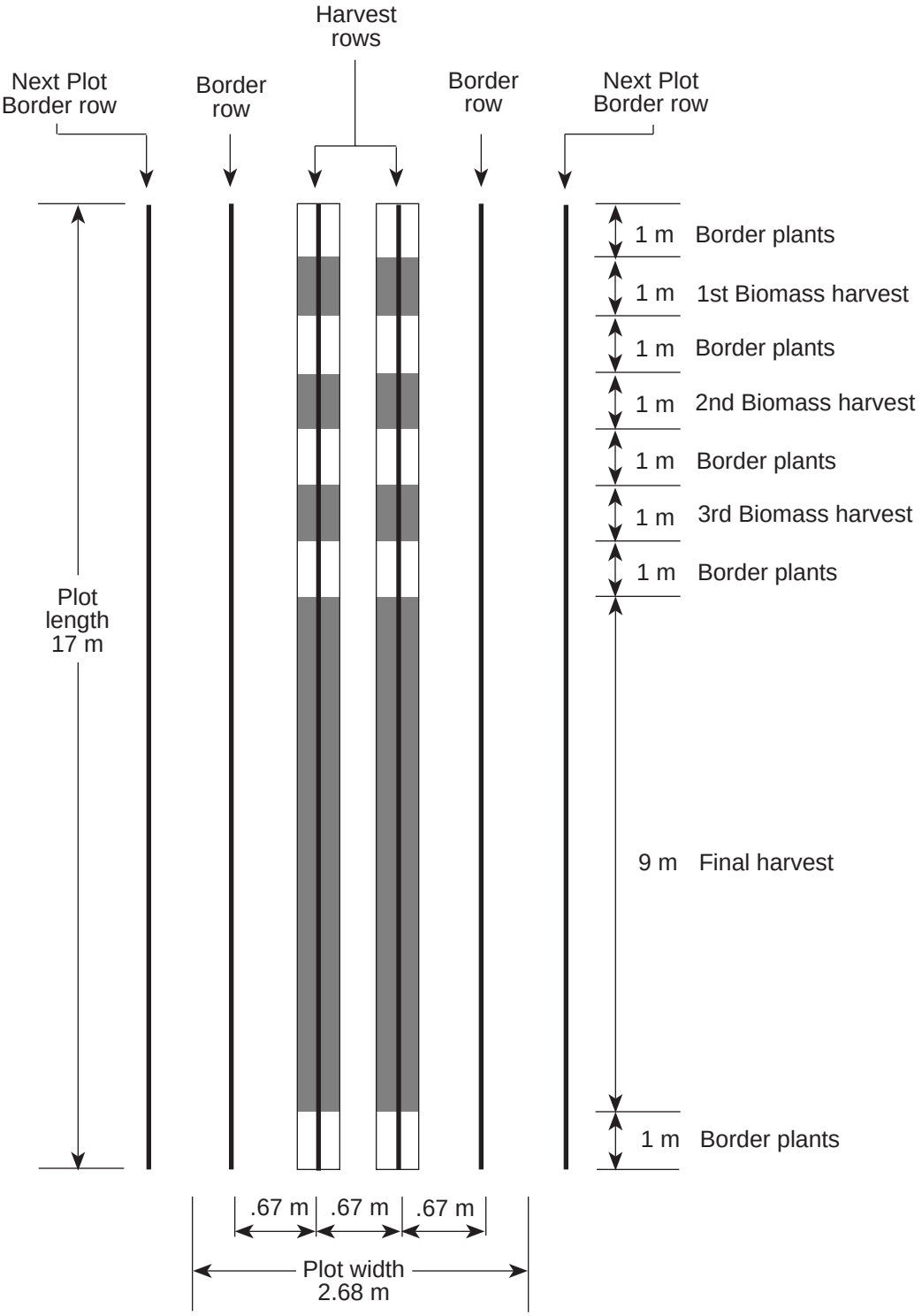
$$\text{leaf blade wt. dry (g/m}^2\text{)} = \frac{\text{leaf blade wt. (undried)} \times \frac{\text{leaf blade subsample wt. (dry)}}{\text{leaf blade subsample wt. (undried)}}}{\text{harvest area}}$$

Calculate and record aboveground biomass (dry).

$$\text{aboveground biomass dry (g/m}^2\text{)} = \text{seed wt. (dry)} + \text{stover wt. (dry)}$$


Measure leaf area using leaf area meter for green leaf blades only from 10 representative plants. Log leaf area (cm²). Calculate and record LAI.
LAI = leaf area (cm²) / 10 plants x plant population (plants/m²) / 10,000 cm²/m²
Separate all plants into the following plant components: leaf blade, husk, leaf sheath, and stem (includes tassel). Weigh and log component weights (undried). Chop components separately.
Repeat the last 3 steps for leaf sheath (dry) weight, stem (dry) weight, and husk (dry) weight.
Calculate and record stover weight (dry).
stover wt. dry (g/m²) = leaf blade wt. (dry) + leaf sheath wt. (dry) + stem wt. (dry) + (cob + husk) wt. (dry)

PLOT DIAGRAM OF HARVEST AREA FOR MAIZE



FINAL HARVEST PROCEDURE FOR SORGHUM & MILLET

Measure area to be harvested. Cut plants at base as near to the ground as possible.

Calculate plant population of harvest area (plants/m^2) and record.

$$\text{plant population } (\text{plants}/m^2) = \frac{\text{no. of plants harvested}}{\text{harvest area}}$$

Weigh and log panicle and stover weights (undried) after separation.

Panicle

Calculate and record panicle number (panicles/m^2) by dividing number of panicles by harvest area.

Select a subsample of 10 representative panicles for dry weight and seed number determination. Log panicle subsample weight (undried).

Remove seed from rachis of subsample and log seed subsample weight (undried). Count 1000 seeds, weigh, and log weight.

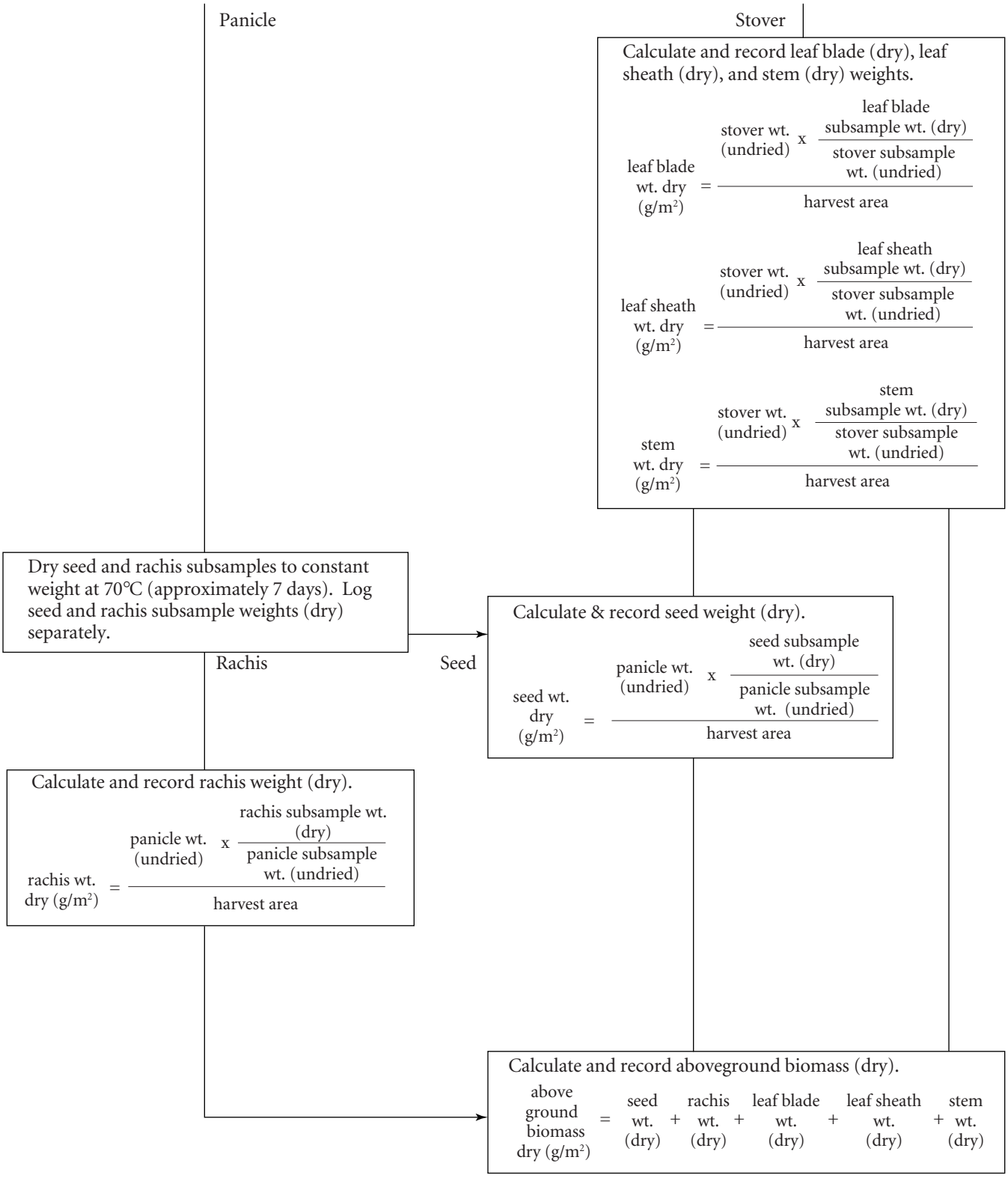
Calculate and record the seed number.

$$\text{seed no. } (\text{seed}/m^2) = \frac{\text{panicle wt. (undried)} \times \frac{\text{seed subsamp. wt. (undried)}}{\text{panicle subsamp. wt. (undried)}} \times \frac{1000 \text{ seeds}}{1000 \text{ seed wt.}}}{\text{harvest area}}$$

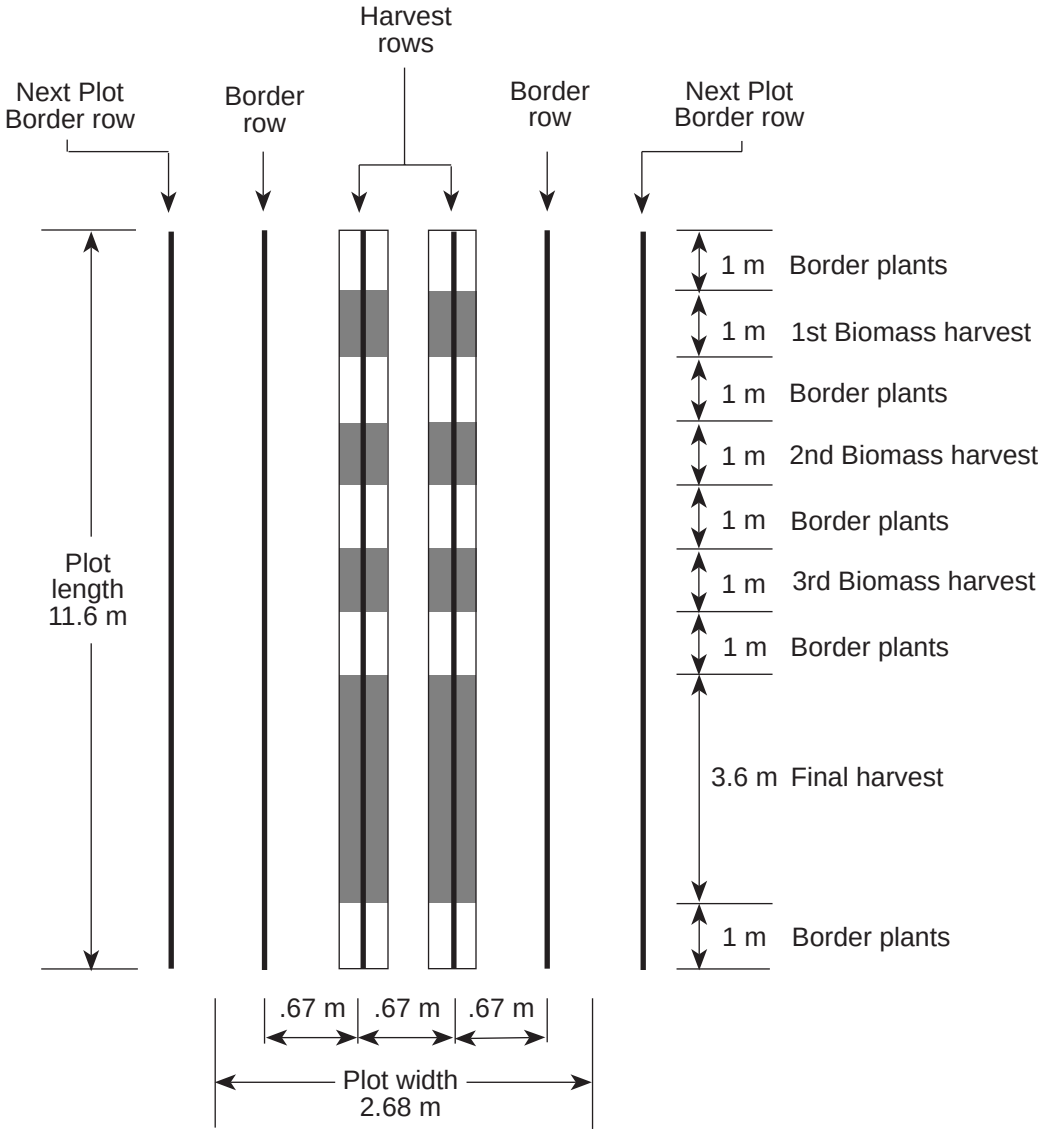
Stover

Select a subsample of 10 representative stalks for dry weight determination. Log subsample weight (undried).

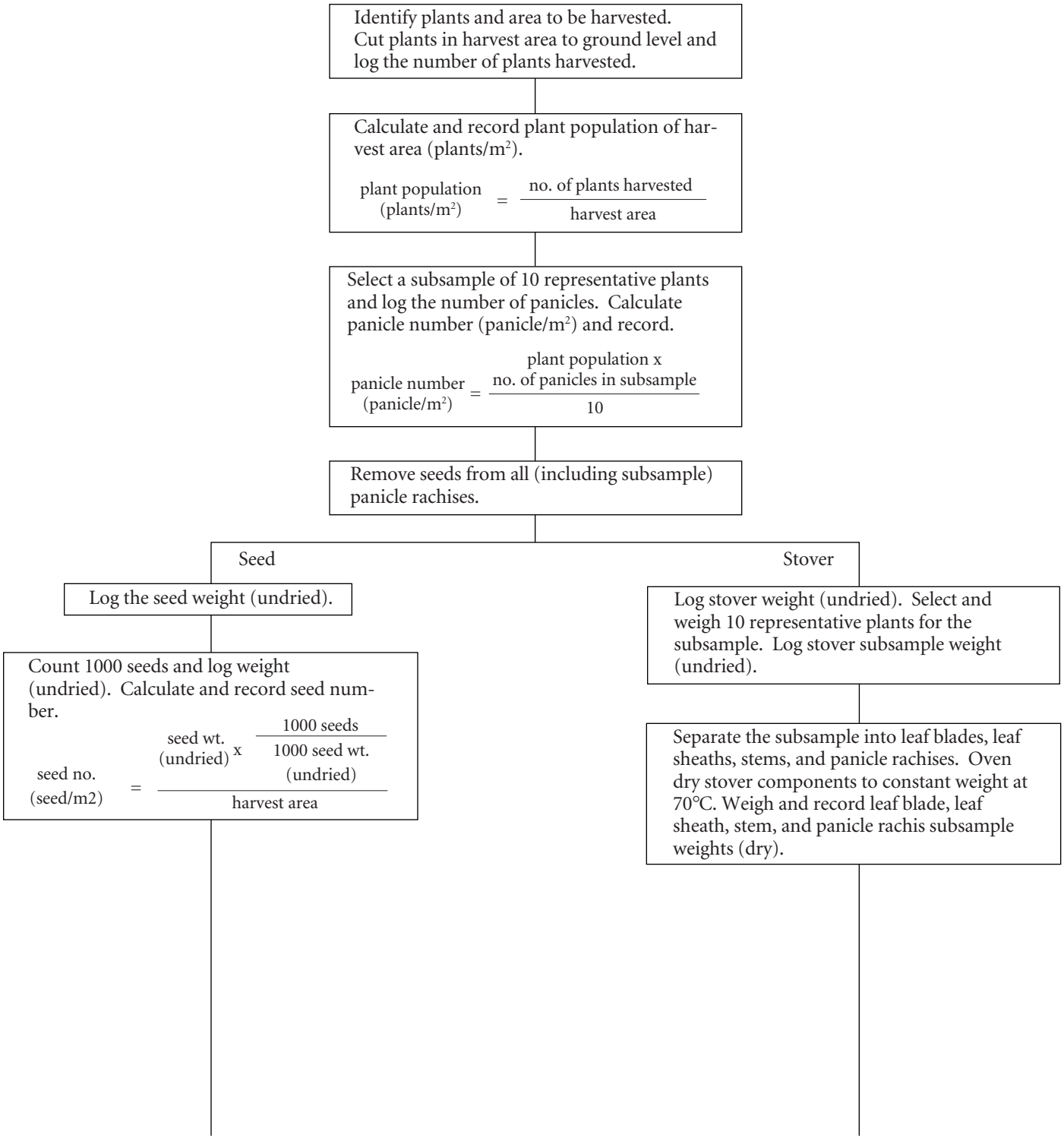
Separate subsample into leaf blades, leaf sheaths, and stems. Oven dry each subsample at 70°C. Log each subsample weight separately.

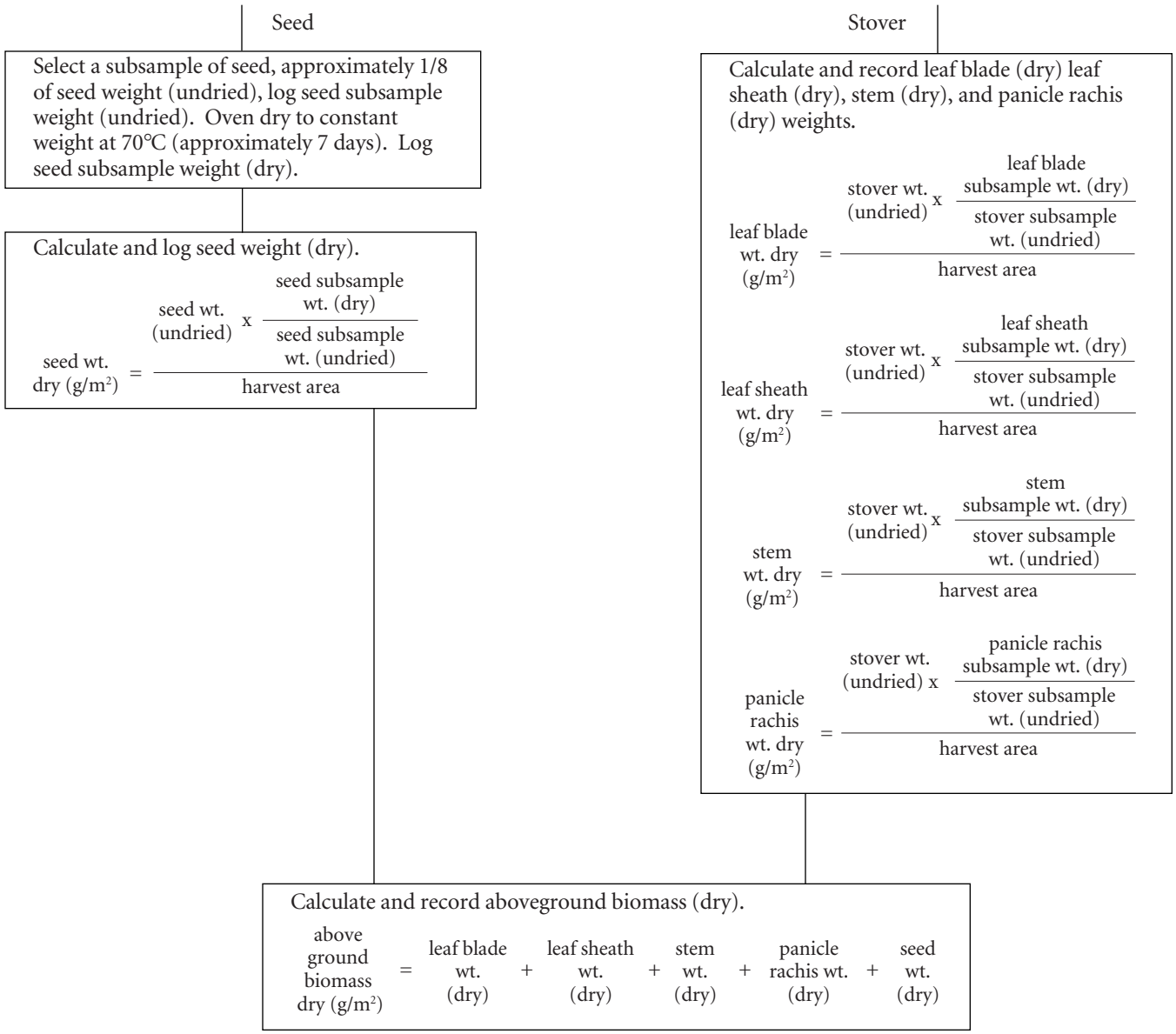

Panicle
Stover
Calculate and record leaf blade (dry), leaf sheath (dry), and stem (dry) weights.
leaf blade wt. dry (g/m²) = stover wt. (undried) x leaf blade subsample wt. (dry) / stover subsample wt. (undried) / harvest area
leaf sheath wt. dry (g/m²) = stover wt. (undried) x leaf sheath subsample wt. (dry) / stover subsample wt. (undried) / harvest area
stem wt. dry (g/m²) = stover wt. (undried) x stem subsample wt. (dry) / stover subsample wt. (undried) / harvest area
Dry seed and rachis subsamples to constant weight at 70°C (approximately 7 days). Log seed and rachis subsample weights (dry) separately.
Rachis
Seed
Calculate & record seed weight (dry).
seed wt. dry (g/m²) = panicle wt. (undried) x seed subsample wt. (dry) / panicle subsample wt. (undried) / harvest area
Calculate and record rachis weight (dry).
rachis wt. dry (g/m²) = panicle wt. (undried) x rachis subsample wt. (dry) / panicle subsample wt. (undried) / harvest area
Calculate and record aboveground biomass (dry).
above ground biomass dry (g/m²) = seed wt. (dry) + rachis wt. (dry) + leaf blade wt. (dry) + leaf sheath wt. (dry) + stem wt. (dry)

PLOT DIAGRAM OF HARVEST AREA FOR SORGHUM & MILLET

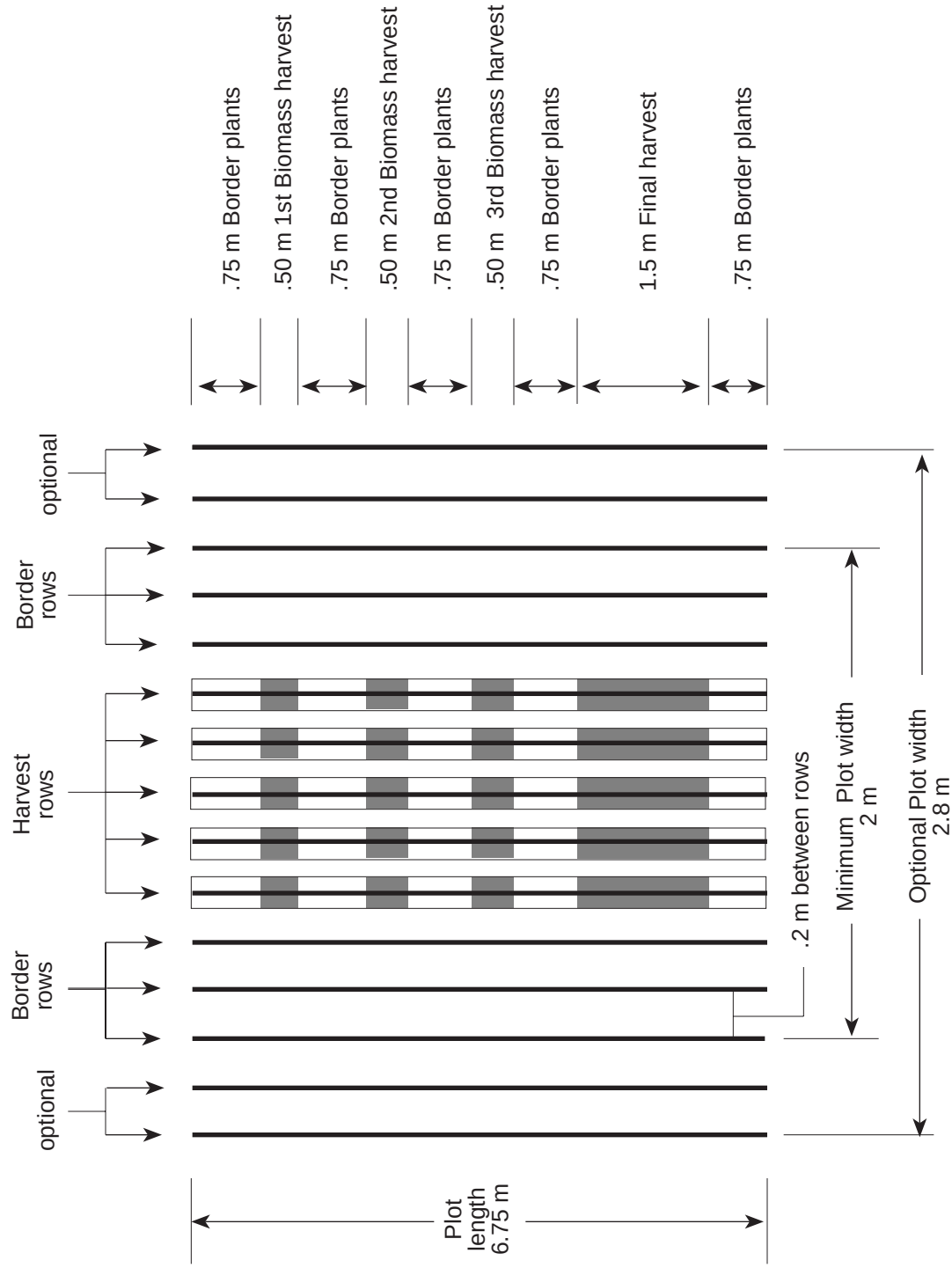


FINAL HARVEST PROCEDURE FOR RICE/WHEAT/BARLEY




Seed
Select a subsample of seed, approximately 1/8 of seed weight (undried), log seed subsample weight (undried). Oven dry to constant weight at 70°C (approximately 7 days). Log seed subsample weight (dry).
Calculate and log seed weight (dry).
seed wt. dry (g/m²) = seed wt. (undried) x seed subsample wt. (dry) / seed subsample wt. (undried) / harvest area
Stover
Calculate and record leaf blade (dry) leaf sheath (dry), stem (dry), and panicle rachis (dry) weights.
leaf blade wt. dry (g/m²) = stover wt. (undried) x leaf blade subsample wt. (dry) / stover subsample wt. (undried) / harvest area
leaf sheath wt. dry (g/m²) = stover wt. (undried) x leaf sheath subsample wt. (dry) / stover subsample wt. (undried) / harvest area
stem wt. dry (g/m²) = stover wt. (undried) x stem subsample wt. (dry) / stover subsample wt. (undried) / harvest area
panicle rachis wt. dry (g/m²) = stover wt. (undried) x panicle rachis subsample wt. (dry) / stover subsample wt. (undried) / harvest area
Calculate and record aboveground biomass (dry).
above ground biomass dry (g/m²) = leaf blade wt. (dry) + leaf sheath wt. (dry) + stem wt. (dry) + panicle rachis wt. (dry) + seed wt. (dry)

PLOT DIAGRAM* OF HARVEST AREA FOR RICE, WHEAT, & BARLEY



Optional border rows can be utilized for destructive sampling to determine panicle initiation data, grain filling rate, etc.

*This is a sample plot layout. The experimenter may change the dimensions, row spacing, etc. as appropriate to the specific experiment.

Final Harvest Procedure for CASSAVA

Identify and determine the number of plants and area to be harvested. Harvest tubers and aboveground biomass from the designated area. Separate above ground biomass from the tubers.

Tubers

Aboveground Biomass

Wash soil off tubers. Count tubers. Calculate tuber number by dividing harvest area by total tuber number (tubers/m^2).

$$\text{tuber number } (\text{tubers}/m^2) = \frac{\text{no. of tubers harvested}}{\text{harvest area}}$$

Separate "usable" tubers (undried, weighing over 200 g) from smaller ones.

Usable Tubers

Count number of usable tubers, divide by harvest area and record usable tuber number (tubers/m^2).

Log usable tuber weight (undried).

Chop tubers into 2 cm pieces and mix well.

Select a usable tuber subsample (undried) of approximately 2 kg and log weight.

Oven dry the subsample at 70°C to constant weight. Record usable tuber subsample weight (dry).

Small Tubers

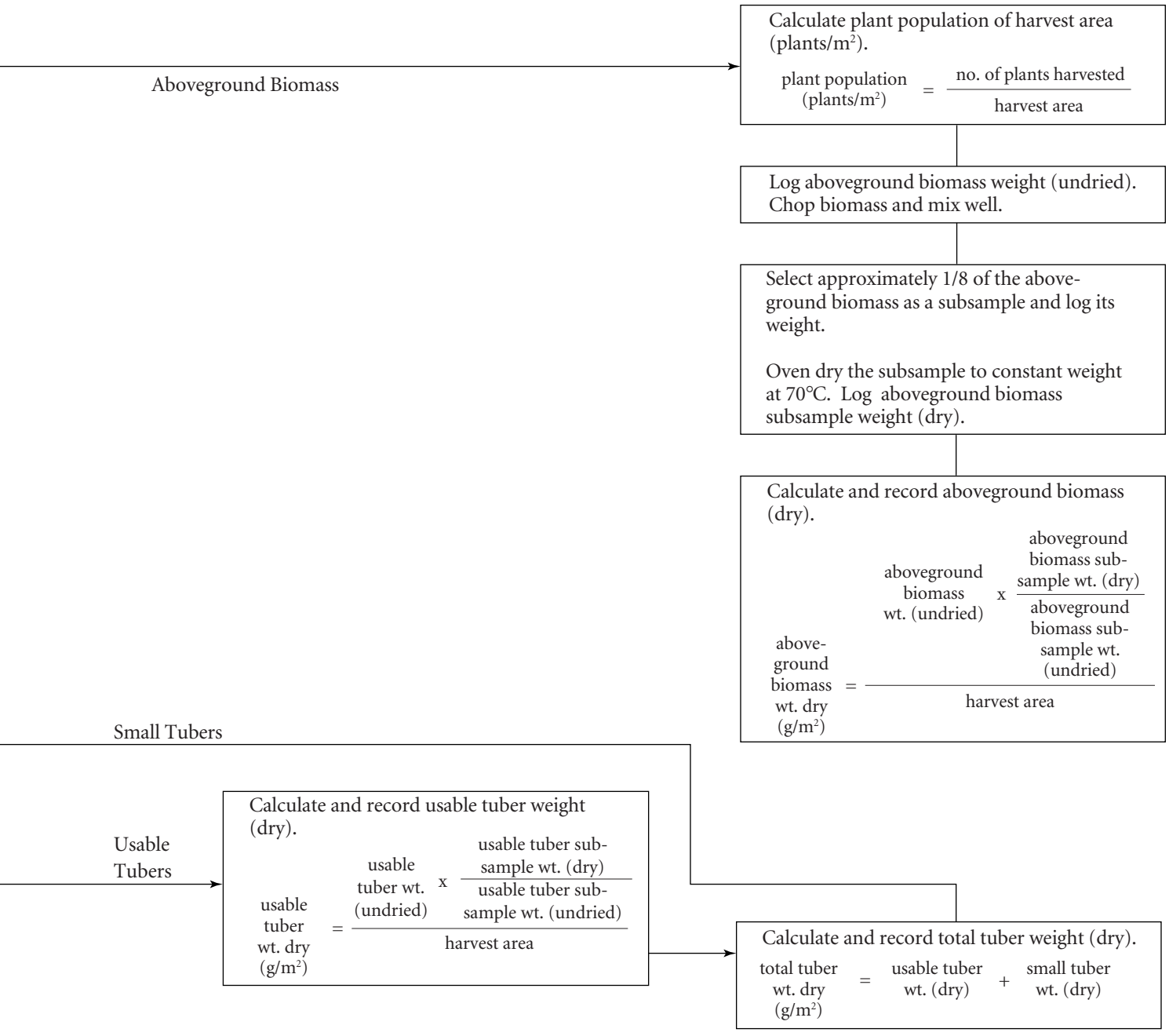
Log small tuber weight (undried). Chop tubers into small pieces and mix well.

Select a small tuber subsample (undried) of approximately 2 kg and log weight.

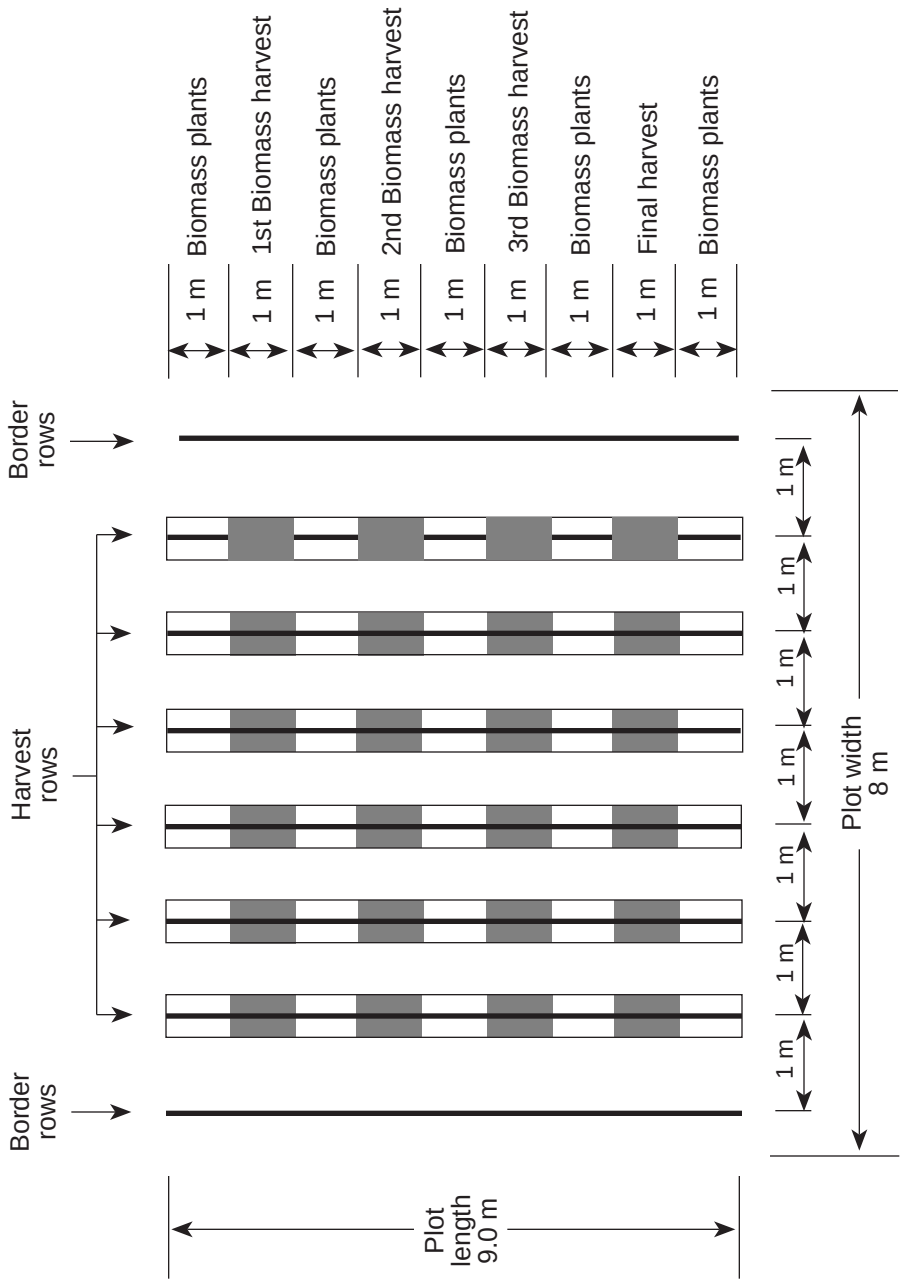
Oven dry the subsample at 70°C to constant weight. Log small tuber subsample (dry) weight.

Calculate and log small tuber weight (dry) per unit area.

$$\text{small tuber wt. dry } (g/m^2) = \frac{\text{small tuber wt. (undried)} \times \frac{\text{small tuber subsample wt. (dry)}}{\text{small tuber subsample (undried)}}}{\text{harvest area}}$$


Aboveground Biomass
Calculate plant population of harvest area (plants/m²).
plant population (plants/m²) = no. of plants harvested / harvest area
Log aboveground biomass weight (undried). Chop biomass and mix well.
Select approximately 1/8 of the aboveground biomass as a subsample and log its weight.
Oven dry the subsample to constant weight at 70°C. Log aboveground biomass subsample weight (dry).
Calculate and record aboveground biomass (dry).
aboveground biomass wt. dry (g/m²) = aboveground biomass wt. (undried) x aboveground biomass sub-sample wt. (dry) / aboveground biomass sub-sample wt. (undried) / harvest area
Small Tubers
Usable Tubers
Calculate and record usable tuber weight (dry).
usable tuber wt. dry (g/m²) = usable tuber wt. (undried) x usable tuber sub-sample wt. (dry) / usable tuber sub-sample wt. (undried) / harvest area
Calculate and record total tuber weight (dry).
total tuber wt. dry (g/m²) = usable tuber wt. (dry) + small tuber wt. (dry)

PLOT DIAGRAM OF HARVEST AREA FOR CASSAVA



Final Harvest Procedure for AROID

Identify plants and area to be harvested. Count plants and calculate plant population of area harvested (plants/m^2). Record plant population of harvest area.

$$\text{plant population} = \frac{\text{no. of plants harvested}}{\text{harvest area}}$$

Remove plants from harvest area. Wash soil from corm and cormels and drain until water completely evaporates from the plants.

Separate corms, cormels, petioles, and leaves.

Petioles & Leaves

Corms

Count the corms. Calculate and record corm number (corms/m^2).

Weigh the corms and log the corm weight (undried).

Chop the corms and mix well. Select a corm subsample (undried) of approximately 2000 g and log its weight.

Oven dry the corm subsample to constant weight at 70°C. Log the corm subsample weight (dry).

Calculate and record corm weight (dry).

$$\text{corm wt. dry } (g/m^2) = \frac{\text{corm wt. (undried)} \times \frac{\text{corm subsample wt. (dry)}}{\text{corm subsample wt. (undried)}}}{\text{harvest area}}$$

Cormels

Count the cormels. Calculate and record cormel number (cormels/m^2).

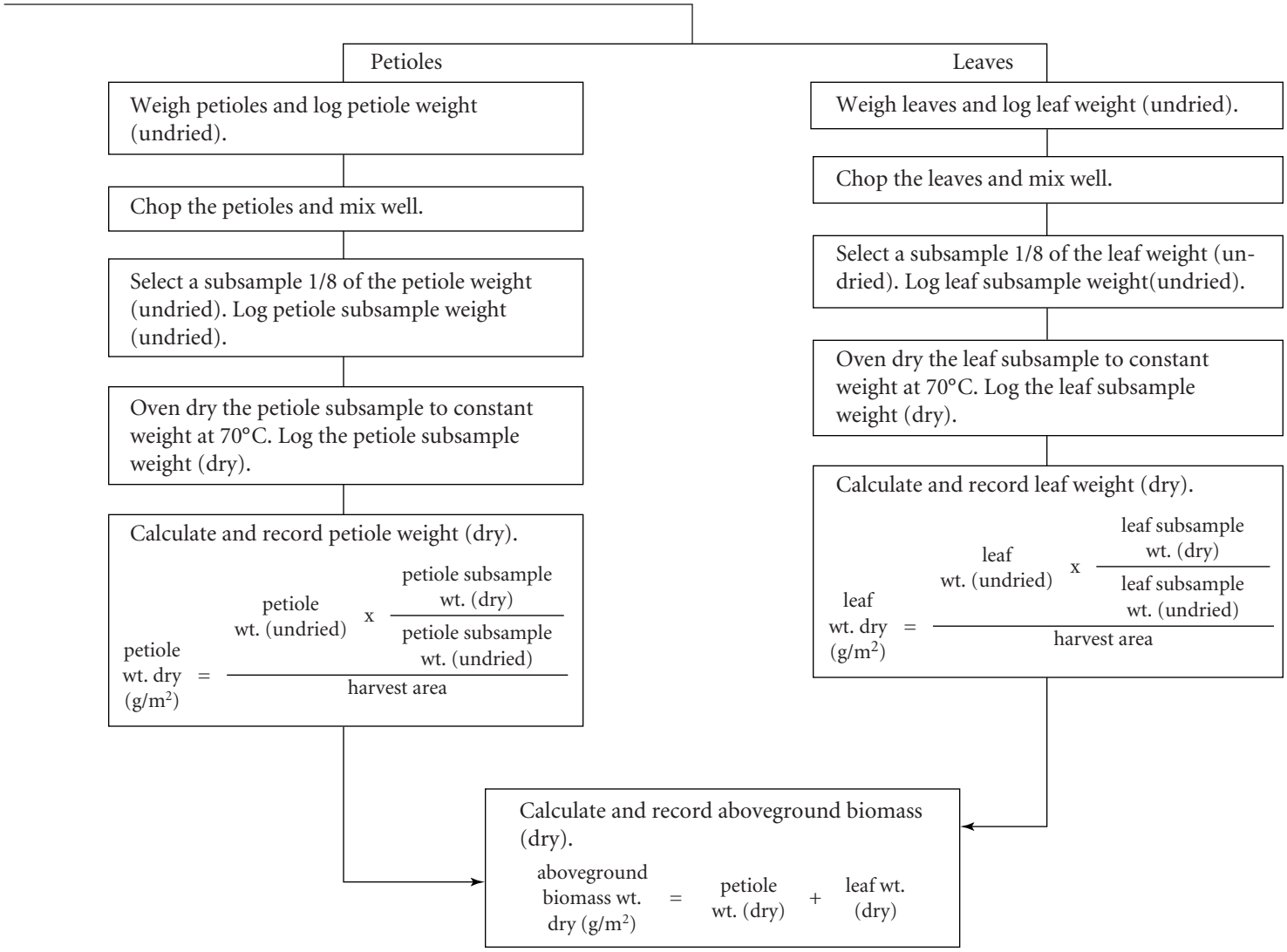
Weigh the cormels and log the cormel weight (undried).

Chop the cormels and mix well. Select a cormel subsample (undried) of approximately 2000 g and log its weight.

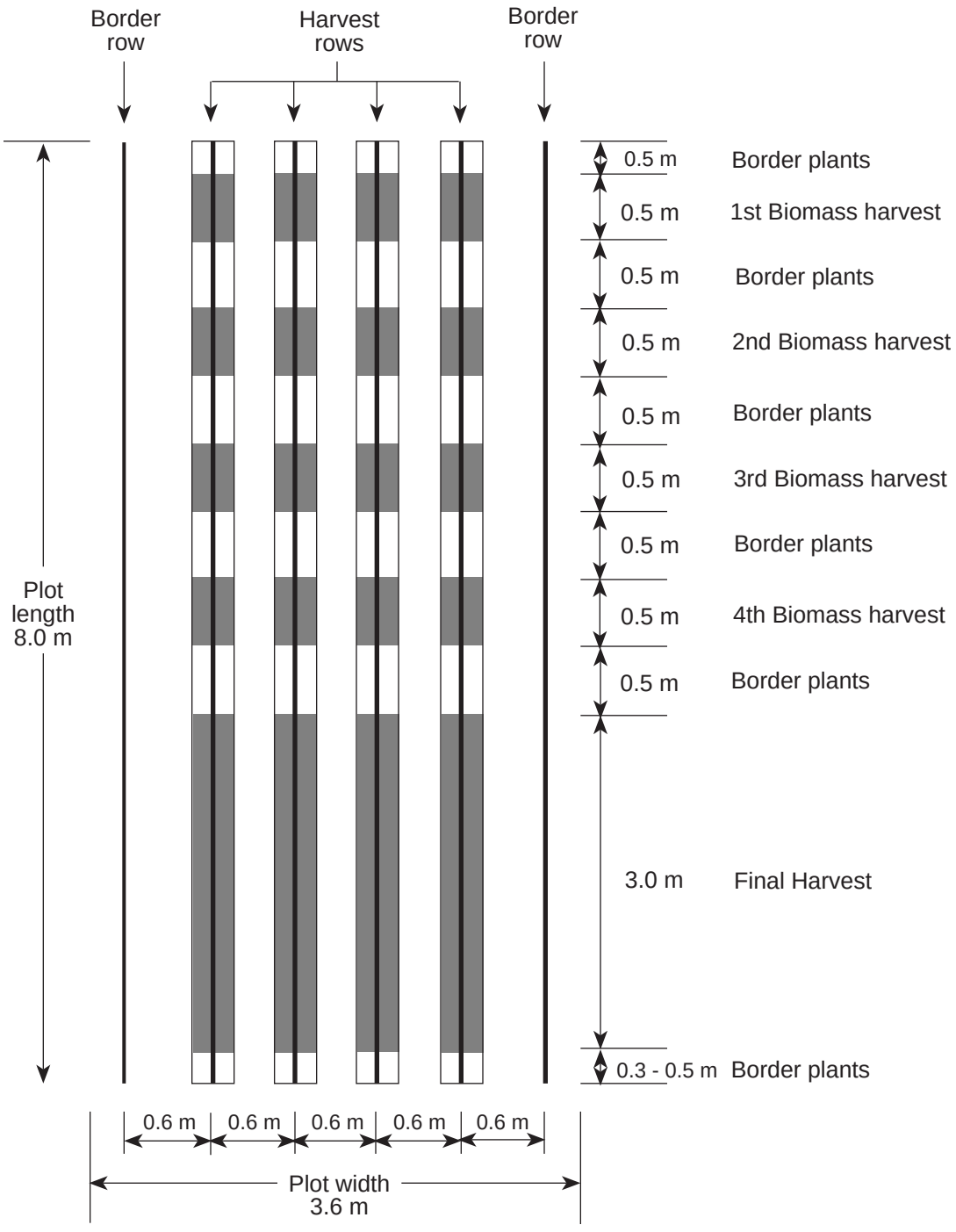
Oven dry the cormel subsample to constant weight at 70°C. Log the cormel subsample weight (dry).

Calculate and record cormel weight (dry).

$$\text{cormel wt. dry } (g/m^2) = \frac{\text{cormel wt. (undried)} \times \frac{\text{cormel subsample wt. (dry)}}{\text{cormel subsample wt. (undried)}}}{\text{harvest area}}$$


Petioles
Weigh petioles and log petiole weight (undried).
Chop the petioles and mix well.
Select a subsample 1/8 of the petiole weight (undried). Log petiole subsample weight (undried).
Oven dry the petiole subsample to constant weight at 70°C. Log the petiole subsample weight (dry).
Calculate and record petiole weight (dry).
petiole wt. dry (g/m2) = petiole wt. (undried) x petiole subsample wt. (dry) / petiole subsample wt. (undried) / harvest area
Leaves
Weigh leaves and log leaf weight (undried).
Chop the leaves and mix well.
Select a subsample 1/8 of the leaf weight (undried). Log leaf subsample weight(undried).
Oven dry the leaf subsample to constant weight at 70°C. Log the leaf subsample weight (dry).
Calculate and record leaf weight (dry).
leaf wt. dry (g/m2) = leaf wt. (undried) x leaf subsample wt. (dry) / leaf subsample wt. (undried) / harvest area
Calculate and record aboveground biomass (dry).
aboveground biomass wt. dry (g/m2) = petiole wt. (dry) + leaf wt. (dry)

PLOT DIAGRAM OF HARVEST AREA FOR AROID



Final Harvest Procedure For POTATO

Identify plants and area to be harvested. Cut plants in harvest area at ground level. Log the total number of plants harvested and how many of those were main stems. Calculate the plant population and the number of main stems in the harvest area.

Collect all tubers >5 mm in diameter and discard balance of underground plant parts. Separate tubers into two size categories, those 5-30 mm in diameter and those >30 mm. Count and weigh both categories and log data.

5-30 mm | >30 mm

Calculate tuber number for both size categories by dividing numbers of tubers in each category by the harvest area. Calculate tuber weight (undried) for both categories by dividing tuber weight in each category by the harvest area.

Determine dry tuber weight of each category. Either oven dry a 2 kg representative sample of fresh tubers to constant weight at 70°C or determine the specific gravity of about 4 kg of fresh tubers.

Determine specific gravity of tubers. To do this, take 4 kilograms of tubers. Clean and air dry them. Weigh the tubers. Hang basket to hold the tubers in water.
Tare the basket. Put the potatoes in the basket and weigh underwater. The formula for specific gravity is:

$$\text{specific gravity} = \frac{\text{air dry wt.}}{\text{air dry wt.} - \text{underwater wt.}}$$

(Note: Specific gravity is usually 1.09)

5-30 mm | >30 mm

Calculate tuber weight (dry) for the two categories using the appropriate formula shown below.

$$\text{tuber wt dry } (g/m^2) = \frac{\text{tuber wt (undried)} \times \text{tuber subsample wt (dry)}}{\text{harvest area} \times \text{tuber subsample wt (undried)}}$$

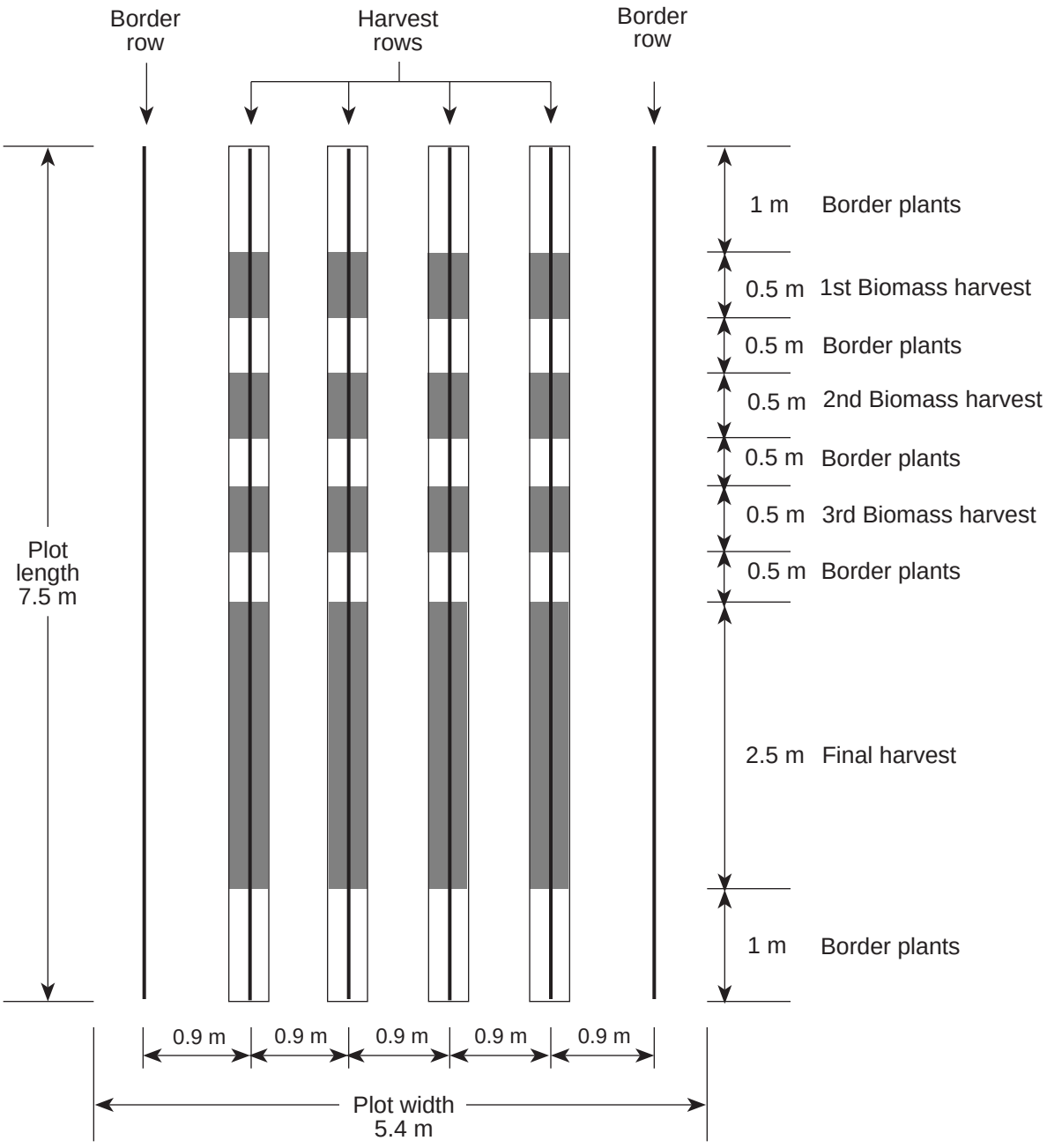
$$\text{tuber dry wt. } (g/m^2) = \frac{\frac{\%\ \text{solids}}{100} \times \text{tuber wt. (undried)}}{\text{harvest area}}$$

$$\%\ \text{solids} = -214.9206 + 218.1852 \times \text{specific gravity}$$

Calculate and record the tuber dry matter content.

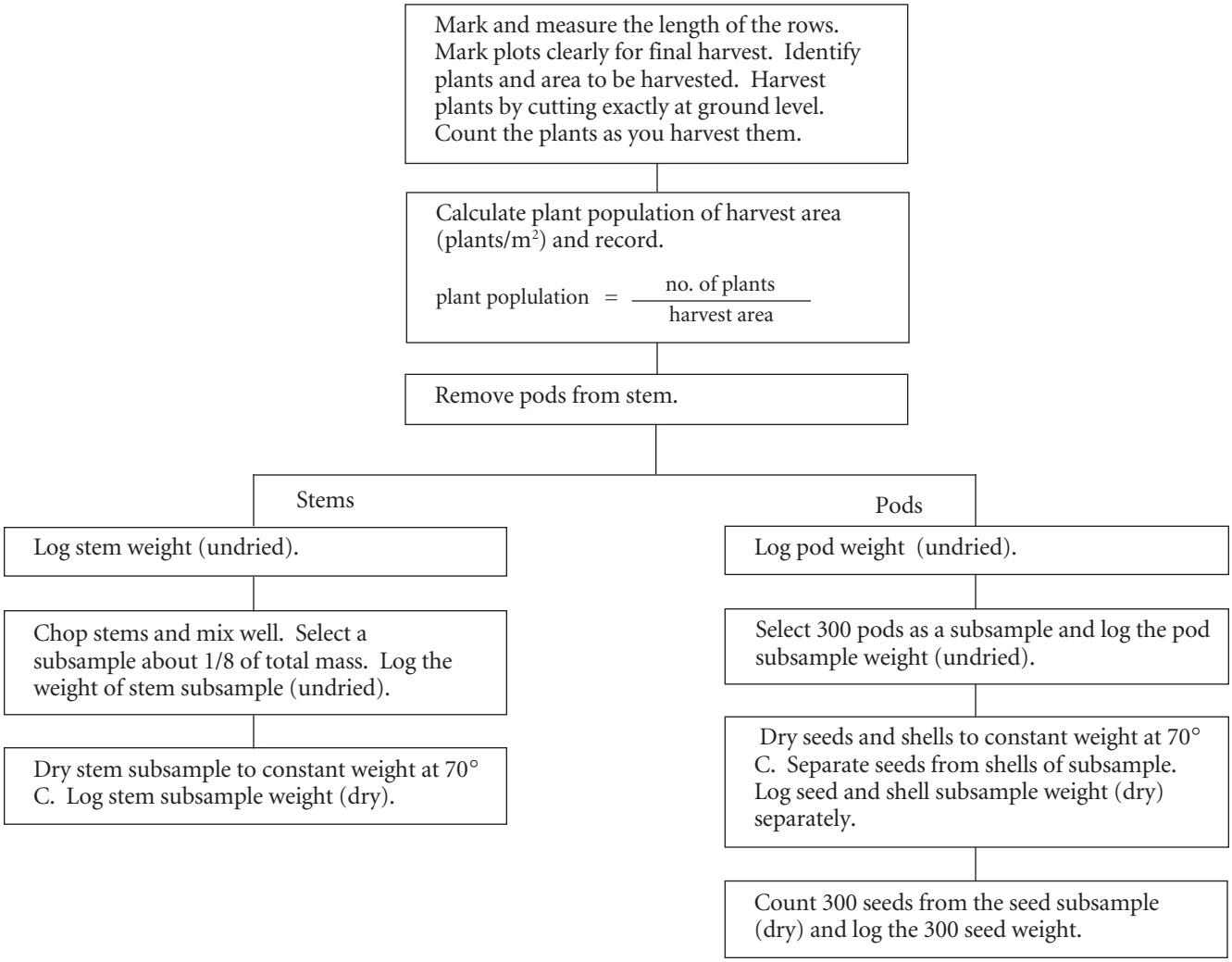
$$\text{tuber dry matter content } (\%) = \frac{\text{tuber wt. dry} \times 100}{\text{tuber wt. undried}}$$

PLOT DIAGRAM OF HARVEST AREA FOR POTATO



Alternatively, plot length can be shortened and plot dimensions made more square by using 6 or 8 row plots. Biomass sampling would then be carried in two outer rows (i.e., rows 2 and 7) and final harvesting in the center rows (i.e., rows 4 and 5).

Final Harvest Procedure for SOYBEAN, DRY BEAN, COWPEA, and CHICKPEA



Formulas for soybean and dry bean calculations:

$$\text{stem weight (dry) (g/m}^2\text{)} = \frac{\text{stem weight (undried)} \times \frac{\text{stem subsample weight (dry)}}{\text{stem subsample weight (undried)}}}{\text{harvest area}}$$

$$\text{pod weight (dry) (g/m}^2\text{)} = \frac{\text{pod weight (undried)} \times \frac{\text{seed subsample weight (dry)} + \text{shell subsample weight (dry)}}{\text{pod subsample weight (undried)}}}{\text{harvest area}}$$

$$\text{shelling percentage} = \frac{\text{seed subsample weight (dry)}}{\text{seed subsample weight (dry)} + \text{shell subsample weight (dry)}} \times 100$$

$$\text{seed weight (dry) (g/m}^2\text{)} = \text{pod weight (dry)} \times \frac{\text{shelling percentage}}{100}$$

$$\text{empty shell weight (dry) (g/m}^2\text{)} = \text{pod weight (dry)} \times \frac{\text{shell subsample weight (dry)}}{\text{seed subsample weight (dry)} + \text{shell subsample weight (dry)}}$$

$$\text{seed number (seed/m}^2\text{)} = \text{seed weight (dry)} \times \frac{300 \text{ seed}}{300 \text{ seed weight (dry)}}$$

$$\text{pod number (pod/m}^2\text{)} = \text{pod weight (dry)} \times \frac{300 \text{ pods}}{\text{seed subsample weight (dry)} + \text{pod subsample weight (dry)}}$$

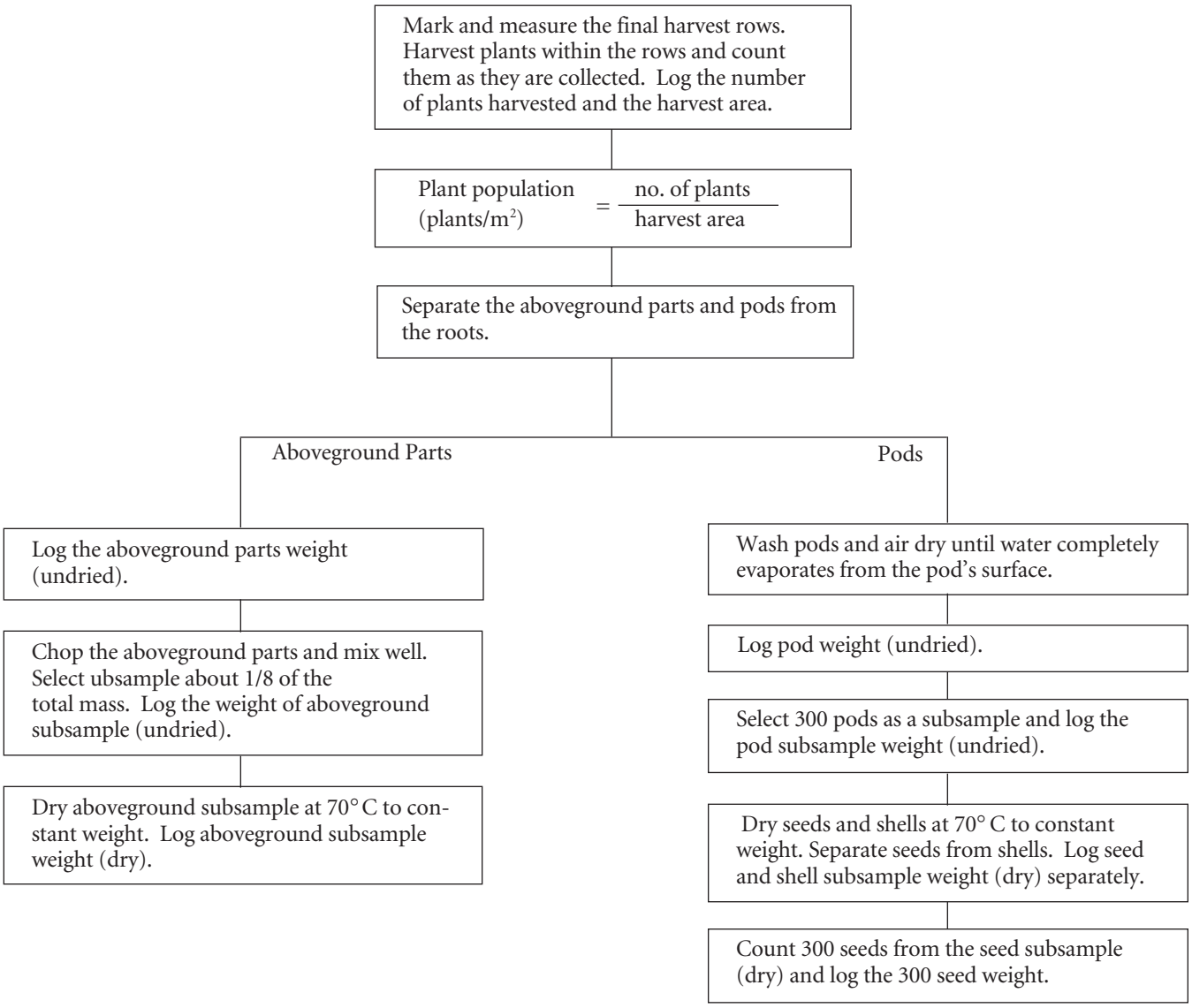
$$\text{seeds per pod} = \frac{\text{seed number}}{\text{pod number}}$$

$$\text{seed size} = \frac{300 \text{ seed weight (dry)}}{300 \text{ seeds}}$$

$$\text{aboveground biomass (dry)} = \text{stem weight (dry)} + \text{pod weight (dry)}$$

$$\text{harvest index} = \frac{\text{seed weight (dry)}}{\text{aboveground biomass (dry)}}$$

Final Harvest Procedure for PEANUT



Formulas for peanut calculations:

$$\text{aboveground biomass (dry) (g/m}^2\text{)} = \frac{\text{aboveground part weight (undried)} \times \frac{\text{aboveground subsample (dry)}}{\text{abovegound subsample (undried)}}}{\text{harvest area}} + \text{pod weight}$$

$$\text{pod weight (dry) (g/m}^2\text{)} = \frac{\text{pod weight (undried)} \times \frac{\text{seed subsample weight (dry)} + \text{shell subsample weight (dry)}}{\text{pod subsample weight (undried)}}}{\text{harvest area}}$$

$$\text{shelling percentage} = \frac{\text{seed subsample weight (dry)}}{\text{seed subsample weight (dry)} + \text{shell subsample weight (dry)}} \times 100$$

$$\text{seed weight (dry) (g/m}^2\text{)} = \text{pod weight (dry)} \times \frac{\text{shelling percentage}}{100}$$

$$\text{empty shell weight (dry) (g/m}^2\text{)} = \text{pod weight (dry)} \times \frac{\text{shell subsample weight (dry)}}{\text{seed subsample weight (dry)} + \text{shell subsample weight (dry)}}$$

$$\text{seed number} = \text{seed weight (dry)} \times \frac{\text{300 seed}}{\text{300 seed weight (dry)}}$$

$$\text{pod number} = \text{pod weight (dry)} \times \frac{\text{300 pods}}{\text{seed subsample weight (dry)} + \text{shell subsample weight (dry)}}$$

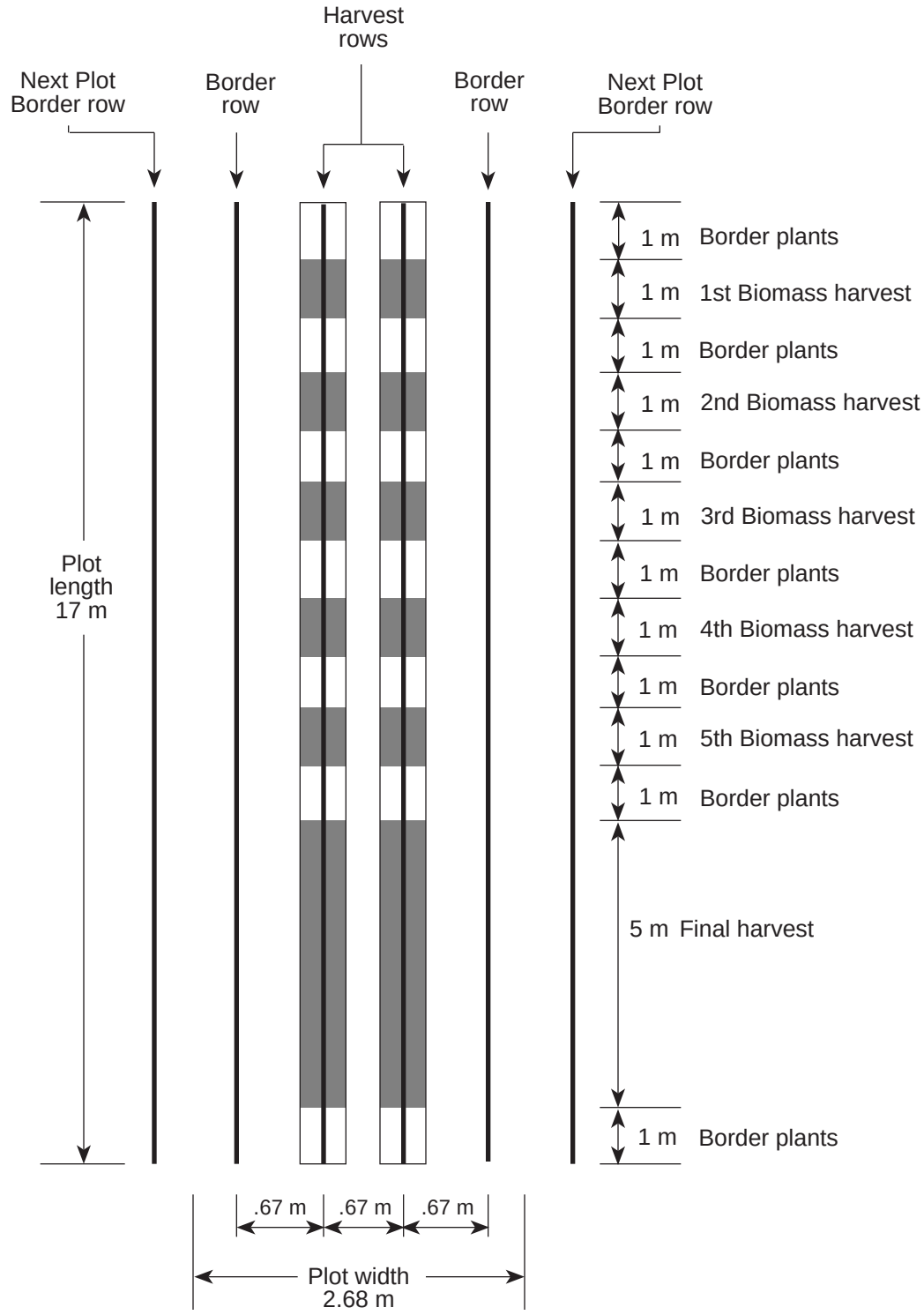
$$\text{seeds per pod} = \frac{\text{seed number}}{\text{pod number}}$$

$$\text{seed size} = \frac{\text{300 seed weight (dry)}}{\text{300 seeds}}$$

$$\text{pod harvest index} = \frac{\text{pod weight (dry)}}{\text{aboveground biomass (dry)}}$$

NOTE: For peanut, "aboveground biomass" refers to everything except the roots. This term does refer to the peanut pods even though those are not literally above the ground.

PLOT DIAGRAM OF HARVEST AREA FOR SOYBEAN, PEANUT, CHICKPEA, AND DRY BEAN



CHAPTER THREE. LABORATORY METHODS

Soil Fertility Measurements

Preplant Soil Sampling

1. For each experimental plot, a representative number of samples from the surface layer (0–15 cm) and at least one set of subsurface samples from each layer should be taken. These samples should be taken from a cross-section of the experimental plot at least two weeks before planting. For example, for a 4 m x 8 m plot, eight surface samples should be collected, mixed throughly, and a representative subsample should be retained for analysis. Use an auger to collect the subsurface samples at depths of 15–30, 30–50, 50–75, and 75–100 cm.

2. Samples from each layer should be mixed thoroughly and subsampled (200 g) for soil chemical analysis.

3. Samples from soils classified as Andepts should not be air dried. Dehydration results in irreversible changes in soil chemical and physical properties.

4. Field moist soil samples for NO_3-N and NH_4-N determinations should be stored in a cooler and refrigerated. Samples should be analyzed within a week.

Post-Harvest Soil Sampling

For experiments where soil fertility or applied fertilizer are variables, post-harvest surface soil samples are required to obtain residual fertilizer levels. These soil samples need to be collected from the experimental plot by the following procedure.

1. Collect the samples within a two week period after the final harvest.

2. For experiments with four rows (two harvest andtwo border rows), collect the samples within the plot harvest area, four subsamples between rows and another four subsamples within the row. These subsamples should be evenly distributed within the final harvest area for each plot.

3. For each plot, mix all eight subsamples and retain about 500 g as the final sample for that plot.

4. Place the samples in a plastic bag, seal, label the bag according to date, plot and replicate number, and place in a second plastic bag.

5. Samples from *soils classified as Andepts should not be dried.*

6. Soil samples for NO_3-N and NH_4-N determination should be refrigerated after collection and analyzed as soon as possible.

7. Samples should then be air-dried and passed through a 2 mm sieve.

Soils Moisture Factor Determination Procedure

All soil data are reported on oven-dry basis; consequently, the moisture factor should be determined.

Procedure

1. Weigh 10 g soil into a weighed Aluminum (Al) dish.
2. Place in an oven at 105°C for 24 hours.
3. Weigh dried sample and Aluminum dish.

CALCULATION:

Sample + Al dish - Al dish = weight of dry sample.

$$\text{Moisture Factor} = \frac{\text{moist sample weight}}{\text{dry sample weight}}$$

Repeat the above for air-dried soil samples as well.

Soil pH Determination

1:1 Soil to H_2O and 1:1 Soil to 1N Potassium Chloride (KCl) Slurry

Apparatus

pH meter; 50-ml beakers.

Procedure

1. Weigh 20 g of air-dried composite soil sample into each of two 50-ml beakers.
2. Add 20 ml distilled water to one beaker, stir vigorously, let stand one hour, then read pH with a pH meter. Record this as pH in a 1:1 soil:water.
3. Add 20 ml 1N KCl to the remaining beaker, stir vigorously, let stand one hour, then read pH with a pH meter. Record this as pH in 1:1 soil:1N KCl.

KCl-Extractable Nitrate and Ammonium Nitrogen

Apparatus

Kemmerer Hallet nitrogen distillation unit; 100-ml sample flask with side arm; 50-ml Erlenmeyer flasks; 50-ml plastic tube; Whatman No. 5 filter paper.

Reagents

1. Potassium Chloride (KCl), 2N.
2. Magnesium Oxide (MgO-Heavy)—ignited at 650˚C for one hour.
3. Devarda alloy—ground to a fine powder.
4. Mineral oil.
5. 2% boric acid (H_2BO_3).
6. Sulfamic acid (NH_2SO_3H)—dissolve 2 g in 100 ml water. Store the solution in a refrigerator or at 7°C.
7. Mixed indicator—3 ml of 0.1% methyl red and 15 ml of 0.1% brom cresol green diluted to 1 liter with ethanol.
8. Standard sulfuric acid (H_2SO_4), 0.005N.

Preparation

1. Weigh 3 g of soil sample into 50-ml plastic tube.
2. Add 30 ml of 2N KCl.
3. Cap and shake for one hour.
4. Filter with Whatman No. 5 filter paper.
5. Pipette 25 ml aliquot into sample flask.

Determination of Nitrogen

1. Add 1 ml NH_2SO_3H into the aliquot and swirl a few seconds to destroy nitrite in extract.

2. Add 1 ml mineral oil to sample to prevent foaming during distillation.

3. Add 0. 2 g of MgO to sample.

4. Place 50-ml Erlenmeyer flask, which contains 5 ml of 2% H_2BO_3 and 3 drops of mixed indicator, at the end of distillation unit.

5. Cap sample flask with side arm with ground glass cover.

6. Attach sample flask to distillation unit.

7. Distill 30 ml into boric acid solution. Set aside for NH_4-N determination. Titrate NH_4-N with standardized 0.005N H_2SO_4.

8. Stop distillation by removing side arm clamp of unit (do not turn off heating unit).

9. Remove ground glass cover of sample flask.

10. Add 0.2 g Devarda alloy.

11. Immediately replace ground glass cover and distill 30 ml into 50 ml flask containing 2% boric acid with 3 drops of mixed indicator. Set aside for NO_3-N determination. Titrate with 0.005N H_2SO_4.

12. Correct for blank.

Calculations

$$mgN = 1\ ml\ of\ 0.005N\ H_2SO_4 \times 70$$

Reference

Page, A.L., R.H. Miller, and D.R. Keeney. 1982. Nitrogen-inorganic forms. p. 643-698. *In* Methods of soil analysis. No. 9, part 2. American Society of Agronomy Inc., and Soil Science Society of America Inc., Madison, Wisconsin.

Extractable Potassium Analysis

The following procedure applies not only for the determination of ammonium acetate extractable K but also for ammonium acetate extractable Ca, Mg, and Na.

Apparatus

Atomic Absorption 7 or flame photometer; Buchner funnel (Coors No.1); 250-ml suction flask; 100-ml volumetric flask; 125-ml Erlenmeyer flask; 500-ml volumetric flask; Whatman No. 42 filter paper.

Reagents

1. Ammonium acetate (CH_3COONH_4), 1N, pH 7.0.
   - Mix 68 ml ammonium hydroxide (NH_4OH), specific gravity 0.90, and 57 ml 99.5% acetic acid (CH_3COOH) per liter of solution desired.
   - Cool, dilute to volume with water, and adjust to pH 7.0 with CH_3COOH or NH_4OH.
   - Optionally, prepare from CCH_3COONH_4 reagent salt and adjust pH.

2. Standards
   - Stock standard = 100 ppm
   - Working standard = 10, 20, 30 ppm.

Procedure

1. Weigh 10 g air-dried <2-mm soil into a 125-ml Erlenmeyer flask and add 15 ml CH_3COONH_4 solution.

2. Cover with stopper, shake the flask for several minutes, and allow to stand 18 hours or more.
3. Transfer contents of the flask to a Buchner funnel attached to suction flask (Coors No.1) fitted with moist Whatman No. 42 filter paper.
4. Filter, using suction.
5. Leach with additional 75 ml CH_3COONH_4, adding small amounts. Minimize leaching time to less than one hour.
6. Transfer leachate to a 100-ml volumetric flask and dilute volume with distilled water.
7. Retain CH_3COONH_4 extract for analysis of exchangeable cations—Ca, Mg, Na, K.

Determination of K

•Flame Photometer
•Atomic Absorption

8. Compare absorption of samples at 766.5 nm on flame photometer, or atomic absorption at 383 (visual), with that of the standard solution to determine the amount of exchangable K.

Adapted From

SCS, USDA. 1972. Soil survey laboratory methods and procedures for collecting soil samples. *In* Soil Survey Investigation Report No. 1. U.S. Gov. Print. Office, Washington, D.C.

Phosphorus Determination

P Extraction: Modified Truog

Apparatus

Spectrophotometer; 4-liter volumetric flask; 2-liter volumetric flask; 1-liter volumetric flask; 200-ml volumetric flask; 25-ml volumetric flask; dark pyrex bottle; 125-ml plastic bottle; 50-ml graduated test tubes; Whatman No. 42 filter paper.

Reagents

1. Extracting solution, 0.02N sulfuric acid (H_2SO_4) and 0.03% ammonium sulfate ($NH_4)_2SO_4$. Mix 2.22 ml concentrated H_2SO_4 and 12 g ($NH_4)_2SO_4$. Dilute to 4 liters.
2. Ascorbic acid reagent (see following page, Murphy and Riley, Reagent B, for method of preparation).
3. Standard should be made with Extraction solution.

Procedure

1. Weigh 1 g soil sample (oven-dried) into 125-ml plastic bottle.

2. Add 100 ml of extracting solution.
3. Shake for 1/2 hour.
4. Filter with Whatman's No. 42 filter paper.
5. Save leachate for P determination. Prepare a blank with extracting solution.

P Determination

Proceed as in Murphy and Riley (see below for procedure).

Reference

Ayres, A.S. and Hagihara, H.H. 1952. Available phosphorus in Hawaiian soil profiles. Hawaiian Planters' Record 54:81-99.

P Determination: Murphy and Riley

Reagents

1. Standard P solution: 5 to 10 ppm solution in extracting solution.
2. Reagent A:
   - Dissolve 12 g of ammonium molybdate [$(HN_4)_6Mo_7O_{21}$: $4H_2O$] in 250 ml of distilled water.
   - In 100 ml distilled water dissolve 0.2908 g of antimony potassium tartrate.
   - Pour both of the dissolved reagents into a 2000-liter volumetric flask, then add 1000 ml of 5N sulfuric acid (H_2SO_4), mix throughly, and fill to volume. Store in a dark pyrex bottle in a cool and dark compartment.
3. Reagent B:
   - Dissolve 1.056 g ascorbic acid in 200 ml of Reagent A and mix.
     NOTE: Reagent B should be prepared daily.

Procedure

1. Pipette suitable aliquot into 25-ml or 50-ml volumetric flask.

2. Dilute to 20 ml with H_2O.

3. Add 4 ml of Reagent B to develop color.

4. Fill to volume and mix throughly.

5. After 20 minutes read optical density. Use any suitable spectrophotometer and read at 850-nm wavelength.

Preparation of Standard Curve

1. Pipette aliquots of P solution containing from 5–40 μg of P into 50-ml graduated tubes.
2. Add 4 ml of Reagent B, dilute to 50 ml, and read absorbance as stated above.

Calculations

$$\text{ppmP} = \frac{\text{micrograms of standard}}{\text{absorbance of standard}} \times \frac{\text{total volume of extractant}}{\text{weight of sample x Aliquot of extract}} \times \frac{\text{moisture}}{\text{factor}}$$

Reference

Murphy, J. and J.P. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta 27: 31–36.

KCl-extractable Acidity and Aluminum

For soils with pH <5.5, aluminum saturation is reported on Form I as a percent of effective cation exchange capacity.

Apparatus

7-1/2 cm. Buchner funnels; 250-ml suction flask; 125-ml Erlenmeyer flasks; Whatman No. 42 filter paper.

Reagent

1. Potassium chloride (KCl), 1N
   - Standard Sodium hydroxide (NaOH): 0.05 N NaOH
   - Standard Sulfuric acid (H_2SO_4) 0.05N
2. Potassium fluoride (KF), 1N

Procedure

1. Weigh 10 g soil samples into 125-ml Erlenmeyer flasks.

2. Add 50 ml 1N KCl to each flask, mix several times, and let stand for 30 minutes.

3. Filter through moistened Whatman No. 42 filter paper in Buchner funnel, with suction flask.

4. Leach each sample as rapidly as possible with about five 9 ml portions of KCl, using the first to help transfer the remaining soil in the Erlenmeyer flasks to the Buchner funnels.

5. Add 3 drops phenolphthalein to the leachate in the suction flask.

6. Titrate with standard NaOH to a pink color that persists for 30 seconds or more.

7. Correct for a KCl blank to obtain KCl-extractable acidity.

8. To obtain KCl extractable aluminum, add 10 ml KF to filtrate and titrate with standard H_2SO_4 until the pink color disappears.

9. Set aside while other samples are titrated and then complete to a lasting colorless end point.

10. Correct for blank.

Calculations

$$\text{Acidity (meq/100 g)} = \frac{\text{ml NaOH}}{\text{g sample}} \times \text{N of NaOH} \times 100 \times \text{moisture factor}$$

$$\text{Al (meq/100 g)} = \frac{\text{ml } H_2SO_2}{\text{g sample}} \times \text{N of } H_2SO_4 \times 100 \times \text{moisture factor}$$

Adapted From

SCS, USDA. 1984. Soil survey laboratory methods and procedures for collecting soil samples. *In* Soil Survey Investigation Report No. 1. U.S. Govt. Printing Office, Washington, D.C.

Plant Nutrient Concentrations

Tissue Sampling and Preparation

Procedure

1. Finely chop (1-2 cm) approximately 1 to 10 kg of aboveground biomass except seeds.

2. Mix chopped plant material thoroughly (approximately 400 g is required for plant analysis).

3. Oven dry plant material at 70°C for 24 hours or more. A constant weight must be attained.

4. Set aside 50 to 100 g for nutrient analysis.

5. Grind plant samples separately in a stainless steel mill, if possible.

6. Store samples in screw-cap bottles and analyze within two months.

7. Procedure for determination of Moisture Factor.
   •Weigh 5 gm of plant tissue into a weighed, Al dish.
   •Place in oven 105°C for 5 hours.
   •Weigh the oven dry sample and the Al dish.

Calculation

$$\text{Sample} + \text{Al dish} - \text{Al Dish} = \text{weight of dry sample}$$

$$\text{Moisture Factor} = \frac{\text{moist sample wt.}}{\text{dry sample wt.}}$$

Procedures for Nutrient Analysis of Plant Samples

The following procedures are recommended to obtain the plant nutrient concentrations. If other methods are used, please send a copy with your data set.

Total Nitrogen (Micro-Kjeldahl)

Apparatus

100-ml Kjeldahl flask; distillation flask; 40-mesh screen.

Reagents

1. Concentrated sulfuric acid (H_2SO_4) 36N.

2. Potassium sulfate mixture: 10 parts powdered potassium sulfate (K_2SO_4) + 1 part fine crystal ferrous sulfate (Fe_2SO_4) + 1/2 part cupric sulfate ($CuSO_4 \cdot 5H_20$).
3. Sodium hydroxide (NaOH) 50%: 500 g of NaOH/1 liter of distilled water.
4. Boric acid (H_2BO_3), 2% aqueous solution: 20 g/liter of distilled water.
5. Standard sulfuric acid (H_2SO_4), 0.05N: 1.39 ml/liter of distilled water.
6. Brom cresol green-methyl red indicator: Dissolve 0.20 g each of brom cresol green and methyl red indicators in 95% ethyl alcohol and dilute to 200 ml with ethyl alcohol.

Procedure

1. Weigh 0.2 g ground aboveground biomass except for seeds (or 0.2 g of seed).

2. Transfer the tissue sample into an 100-ml Kjeldahl flask.

3. Add 5 g potassium sulfate mixture, 2 ml distilled H_2O and 3 ml concentrated sulfuric acid (H_2SO_4). Swirl until the plant material is fully in contact with the reagent. Let stand 18 hours.

4. Heat gently for 5 minutes, watch the mixture carefully to avoid frothing, and then digest at full heat.

5. Continue digestion for one hour after the solution has cleared.

6. Cool and transfer solution into distillation flask with about 40 ml of distilled water.

7. Add 10 ml sodium hydroxide (NaOH), then distill off 30 ml into 5 ml of 2% boric acid (H_2BO_3) with 3 drops of mixed indicator.

8. Titrate with standard 0.05N sulfuric acid (H_2SO_4) until the first faint pink appears.

9. Run a blank and titrate to the same end point.

Calculation

$$\text{N (pct.)} = \frac{\text{ml } H_2SO_4 \text{ sample - ml } H_2SO_4 \text{ of blank}}{\text{g sample}} \times \text{normality of acid} \times 1.4 \times \text{moisture factor}$$

Reference

Page, A.L., R.H. Miller, and D.R. Keeney. 1982. Regular Kjeldahl method. p. 612-616. *In* Methods of soil analysis. Part 2, number 9. American Society of Agronomy Inc. and Soil Science Society of America Inc. Madison, Wisconsin.

P and K Determination

Apparatus
Spectrophotometer; atomic absorption or flame photometer; 100-ml Kjeldahl flask; 100 volumetric flask.

Reagents
HNO_3: $HClO_4$ as a 2:1 mixture.
conc. Nitric acid (HNO_3) : Perchloric acid ($HClO_4$) as a 2:1 mixture.

Procedure
1. Weigh 0.5 g ground aboveground biomass except for seeds and place in a 100-ml Kjeldahl flask.

2. Add 15 ml of HNO_3: $HClO_4$ solution. Let stand 18 hours.

3. Digest sample until the solution is white. Caution: do not evaporate until dry—$HClO_4$ fumes and organic matter can cause an explosion!

4. Cool, transfer solution to a 100-ml volumetric flask, and fill to volume.

Phosphrous Determination

Continue with P Determination Murphy and Riley: Procedure and Preparation of Standard Curve (page 44).

Potassium Determination

Reagents
1. Stock standard potassium: 1.907 gm of KCl diluted to 1 liter or 1000 ppm.
2. Working potassium standard: 1–100 ppm.
3. Dilute sample to 1:10.
4. Determine potassium with flame photometer or atomic absorption.

Calculation
1. Prepare standard curves for mg/ml K against % transmission or absorbance. For a 1:100 plant:solution ratio

$$\frac{\text{mg/ml in final solution x dilution factor}}{100} = \%K$$

For example: For 10 x dilution: $\frac{\text{mg/ml}}{10} = \%K$

Reference

Chu, A.E. 1990. Personal communication. Dept. of Agronomy and Soil Science, University of Hawaii.

GLOSSARY

aboveground biomass: Biomass that is above the ground, not buried. (NOTE: For peanut, aboveground biomass includes pod weight even though pods are not above the ground.)

anthesis: The action or period of opening of a flower. The start of flowering.

apical meristem: The growing point, composed of meristematic tissue, at the tip of the root or shoot in a vascular bundle.

biomass: Total dry weight of all plants in a particular area or plot.

black layer: A dark, suberized membrane that develops around the base of corn kernels at physiological maturity.

canopy: A formation of leaves and branches affording a cover of foliage.

coleoptile: The sheath enclosing the apical meristem and leaf primordia of the grass embryo; often interpreted as the first leaf.

corm: An enlarged, lower stem where photosynthate is stored; usually underground.

cormel: The corm of a sucker arising from the main corm.

dry weight: The weight of material after water is completely removed.

dough stage: A developmental stage that begins when the kernel's endosperm becomes of pasty consistency.

ear leaf: The leaf directly beneath an ear.

endosperm: A tissue containing stored food, that develops from the union of a male nucleus and the polar nuclei of the central cell; it is digested by the growing sporophyte either before or after the maturation of the seed; found only in angiosperms.

haulm (or halm)**:** The stems or tops of cultivated plants such as peas, beans, potatoes, and cereals. Especially after the crop has been gathered.

internode: The region of a stem between two successive nodes.

leaf blade: The lamina-like portion of the leaf bounded by the leaf collar and tip.

leaf sheath: The portion of a leaf below the collar that forms part of the stalk.

Log vs Record: To "log" means to temporarily note, as opposed to "record" which means to set down in writing.

mesocarp: The middle layer of the mature ovary wall, or pericarp, between the exocarp and the endocarp.

node: The part of a stem where one or more leaves are attached

panicle: An inflorescence, the main axis of which is branched, and whose branches bear loose flower clusters.

petiole: The stalk of the leaf.

pericarp: Fruit wall which develops from the mature ovary wall.

pod: A dry dehiscent seed vessel or fruit that is either monocarpellary (as a legume) or composed of two or more carpels.

root: The usually descending axis of a plant, normally below ground, and serving to anchor the plant and to absorb and conduct water and minerals.

rachis: Main axis of spike; in compound leaves, the extension of the petiole corresponding to the midrib of an entire leaf.

senesce: To age and wither.

senescense: The phase of plant growth that extends from full maturity to actual death and is characterized by an accumulation of metabolic products, increase in respiratory rate, and a loss in dry weight, especially in fruits.

stover: The refuse of a field crop such as the stalks and leaves of corn after the ears are harvested. Used as a food for cattle. Also, hay stubble.

tare: The weight of a container that is deducted from the gross weight to obtain the net weight.

testa: The hard external coating or integument of a seed. Also called episperm.

tiller leaf blade: Leaf blade of a branch originating from the main stem's basal node.

tiller leaf sheath: Leaf sheath of a branch originating from the main stem's basal node.

tiller stem: Stem of a branch originating from the main stem's basal node.

tuber: An enlarged, short, fleshy underground stem, such as that of the potato.

wet weight: The weight of a material before it is dried.

APPENDIX A.
OPTIONAL WEATHER DATA COLLECTION

You may record weather data not required by the minimum data set. The additional weather data collected permit estimations of potential evapotranspiration in the crop models with the Penman method. These estimations may be especially useful in arid regions where advection can occur.

Parameter	Instrument
Dry bulb temperature	Dry-bulb thermometer of a psychrometer
Wet bulb temperature	Wet-bulb thermometer of a psychrometer
Windrun	Anemometer

Dry Bulb Temperature

"Dry bulb" temperature is the temperature of the air taken from a dry temperature-sensing element, such as the dry-bulb thermometer of a psychrometer.

Wet Bulb Temperature

"Wet bulb" temperature is the lowest temperature to be measured by evaporating water from a wick-covered temperature-sensing element, such as the wet-bulb thermometer of a psychrometer.

Windrun

"Windrun" is the measurement of wind movement during a 24-hour period, and is usually recorded with an anemometer.

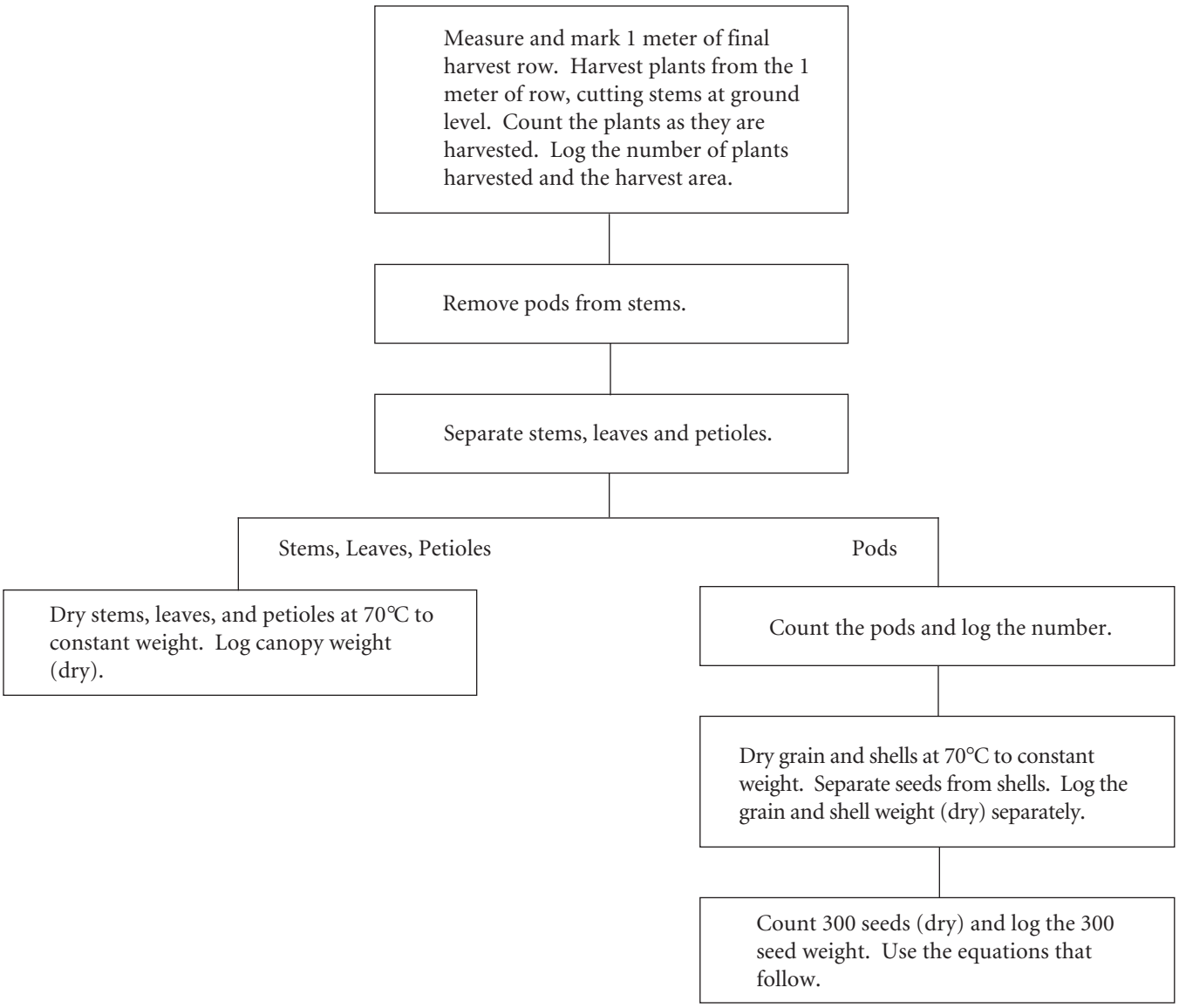
APPENDIX B.

The following flow charts refer to harvest procedures for large plots of soybean, dry bean, chickpea, cowpea, and peanut. A "large plot" may be defined as a plot that is so large as to require the assistance of a harvest combine or other such equipment to harvest the field.

Each of the flow charts is divided into areas of harvest. The first is a) the one meter area of the field and the other, b) is the rest of the field excluding the one meter area. Researchers who follow this large plot format need to gather data from both of these areas for input into the minimum data set.

Please note that for peanut only, "aboveground biomass" refers to not just that part of the plant that is above the ground, but to the pods below the ground as well. However, the roots of the peanut are not considered as part of aboveground biomass.

SOYBEAN, DRY BEAN, COWPEA, AND CHICKPEA:
FINAL HARVEST FOR 1 METER AREA OF LARGE PLOT



Formulas for soybean, cowpea, chickpea, and dry bean calculations:

$$\text{Plant population of harvest area} = \frac{\text{number of plants}}{\text{harvest area}}$$

$$\text{seed weight (dry) (g/m}^2\text{)} = \frac{\text{grain weight (dry)}}{\text{harvest area}}$$

$$\text{Empty shell weight (dry) (g/m}^2\text{)} = \frac{\text{shell weight (dry)}}{\text{harvest area}}$$

$$\text{Pod weight (dry) (g/m}^2\text{)} = \text{seed weight (dry)} + \text{empty shell weight (dry)}$$

$$\text{Shelling percentage} = \frac{\text{seed weight (dry)}}{\text{pod weight (dry)}} \times 100$$

$$\text{Seed Number (seed/m}^2\text{)} = \text{seed weight (dry)} \times \frac{300 \text{ seeds}}{300 \text{ seed weight}}$$

$$\text{Pod number (pods/m}^2\text{)} = \frac{\text{number of pods}}{\text{harvest area}}$$

$$\text{Seeds per pod} = \frac{\text{seed number}}{\text{pod number}}$$

$$\text{Seed size} = \frac{\text{300 seed weight}}{\text{300 seeds}}$$

$$\text{Aboveground biomass (dry) (g/m}^2\text{)} = \text{pod weight (dry)} + \frac{\text{stem, leaf, petiole weight (dry)}}{\text{harvest area}}$$

$$\text{Harvest index} = \frac{\text{seed weight (dry)}}{\text{aboveground biomass (dry)}}$$

SOYBEAN, DRY BEAN, COWPEA, AND CHICKPEA: FINAL HARVEST FLOW CHART FOR LARGE PLOT (EXCLUDING 1 METER AREA)

From the remaining final harvest rows, harvest plants and thresh the pods. Log the harvested area. Dry the grain at 70°C to constant weight. Log the grain weight (dry). Recalculate seed, aboveground biomass, pod, empty shell dry weight, seed number, and pod number.

Use the following equations in the given order.

$$\text{Seed weight (dry) (g/m}^2\text{)} = \frac{\text{seed weight (dry)}}{\text{harvest area}}$$

$$\text{Aboveground biomass (dry) (g/m}^2\text{)} = \frac{\text{seed weight (dry)}}{\text{harvest index*}}$$

$$\text{Pod weight (dry) (g/m}^2\text{)} = \frac{\text{seed weight (dry)}}{\text{shelling percentage*} \div 100}$$

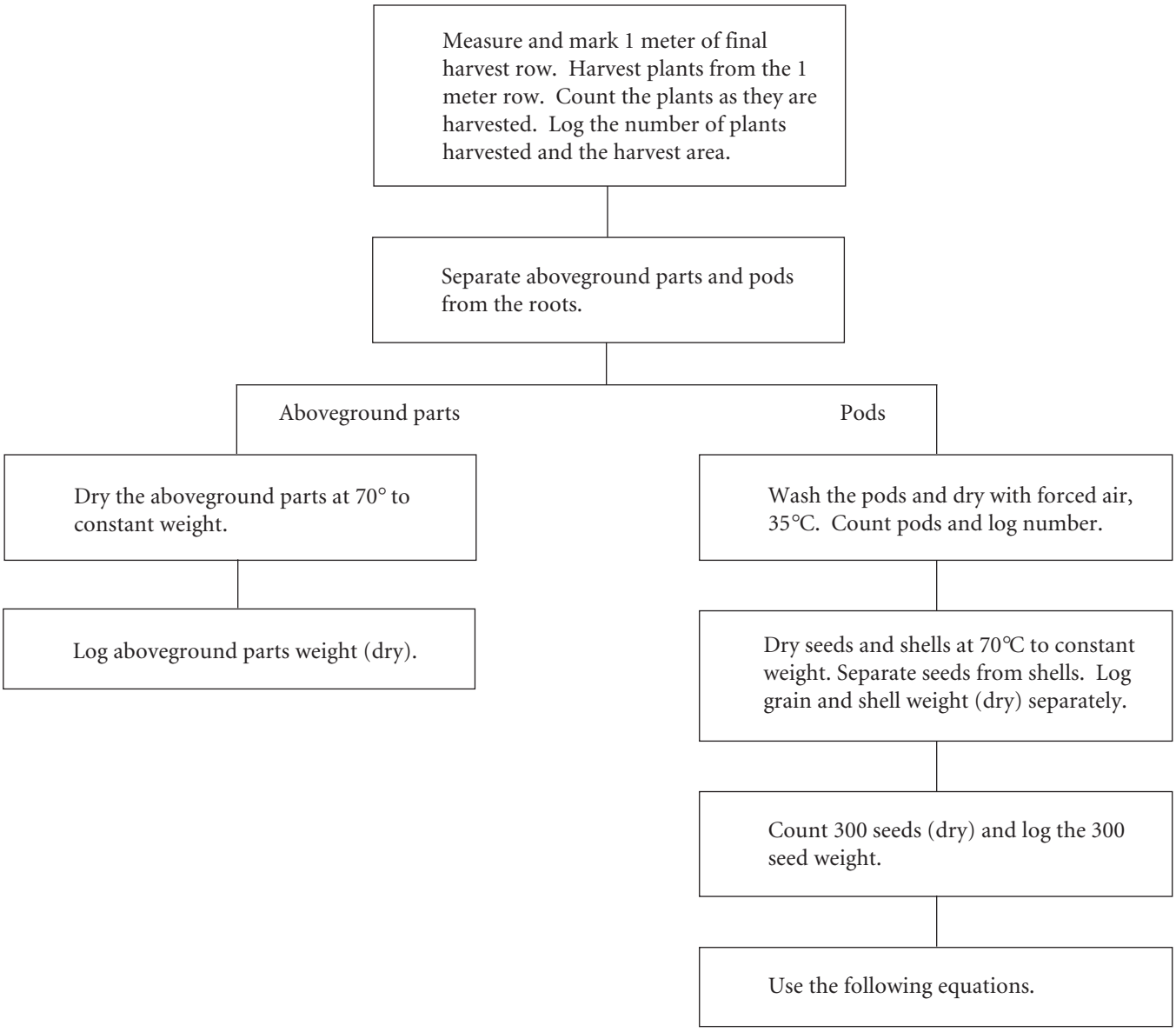
$$\text{Empty shell weight (dry) (g/m}^2\text{)} = \text{pod weight (dry)} - \text{seed weight (dry)}$$

$$\text{Seed number} = \frac{\text{seed weight (dry)}}{\text{seed size*}}$$

$$\text{Pod number} = \frac{\text{seed number}}{\text{seeds per pod*}}$$

* These four values are taken from the one meter area.

PEANUT: FINAL HARVEST FOR 1 METER AREA OF LARGE PLOT



Formulas for peanut calculations:

$$\text{Plant population of harvest area} = \frac{\text{number of plants}}{\text{harvest area}}$$

$$\text{Seed weight (dry) (g/m}^2\text{)} = \frac{\text{grain weight (dry)}}{\text{harvest area}}$$

$$\text{Empty shell weight (dry) (g/m}^2\text{)} = \frac{\text{shell weight (dry)}}{\text{harvest area}}$$

$$\text{Pod weight (dry) (g/m}^2\text{)} = \text{seed weight (dry)} + \text{empty shell weight (dry)}$$

$$\text{Aboveground biomass (dry) (g/m}^2\text{)} = \frac{\text{aboveground parts weight (dry)}}{\text{harvest area}} + \text{pod weight (dry)}$$

$$\text{Shelling percentage} = \frac{\text{seed weight (dry)}}{\text{pod weight (dry)}} \times 100$$

$$\text{Seed Number (seeds/m}^2\text{)} = \text{seed weight (dry)} \times \frac{300 \text{ seeds}}{300 \text{ seed weight}}$$

$$\text{Pod number (pods/m}^2\text{)} = \frac{\text{number of pods}}{\text{harvest area}}$$

$$\text{Seeds per pod} = \frac{\text{seed number}}{\text{pod number}}$$

$$\text{Seed size} = \frac{\text{300 seed weight}}{\text{300 seeds}}$$

$$\text{Pod Harvest Index} = \frac{\text{pod weight (dry)}}{\text{aboveground biomass (dry)}}$$

NOTE: for peanut,"aboveground biomass" refers to everything except the roots. This term does refer to the peanut pods even though those are not literally above the ground.

PEANUT: FINAL HARVEST FLOW CHART FOR LARGE PLOT (EXCLUDING 1 METER AREA)

Harvest and thresh the pods from the remaining harvest rows. Log the harvested area. Dry the grain at 70°C to constant weight. Log grain weight (dry) Recalculate seed, aboveground biomass, pod, and empty shell weights, using the following equations.

$$\text{Seed weight (dry) (g/m}^2\text{)} = \frac{\text{grain weight (dry)}}{\text{harvest area}}$$

$$\text{Aboveground biomass (dry) (g/m}^2\text{)} = \frac{\text{pod weight (dry)}}{\text{pod harvest index*}}$$

$$\text{Pod weight (dry) (g/m}^2\text{)} = \frac{\text{seed weight (dry)}}{\text{shelling percentage*} \div 100}$$

$$\text{Empty shell weight (dry) (g/m}^2\text{)} = \text{pod weight (dry)} - \text{seed weight (dry)}$$

* These two values are taken from the one meter area.

APPENDIX C. DETERMINATION OF LEAF AREA INDEX

A method is outlined for collaborators who do not have a leaf area meter with which to measure leaf area index. Leaf area index is required for maize and for potato at the T1, T2, and T3 stages. For all other cultivars, leaf area index is only desired data, not required by the minimum data set.

Determination of Leaf Area Index

The leaf area index (LAI) is the leaf area subtended per unit of land area (Chang 1968). LAI is used as an indicator of plant growth and for evaluating assimilation and transpiration rates in plant physiological studies. LAI measurements provide indices of plant growth with time and are customarily used as inputs for crop simulation models. Because accurate measurements of LAI are required to develop and/or validate growth simulation models, it is important to select a method that can provide measurements of leaf are with reasonable confidence. This section briefly describes some of the methods currently available for measuring LAI.

Methods for Measuring LAI

Leaf area measurements are both laborious and time consuming. The selection of a particular method to measure leaf area depends largely on the characteristics of leaves to be measured, the availability of time and equipment, and the degree of accuracy (Daughtry and Hollinger 1984). As a general rule, methods of measuring leaf area are classified into two categories: direct and indirect. In the direct methods, the area of leaves is measured using an area meter, whereas the indirect methods employ leaf area and leaf weight relationships to estimate leaf area.

Direct Methods

Leaf area meters are especially designed for precise and automatic area measurement. Area meters utilize an electronic method of rectangular approximation to measure leaf area of either attached or detached leaves (Hatfield et al. 1976). In the LI-3100 model, capabilities for either 0.1 or 1 mm2 area resolution are provided on the same instrument. As a leaf sample travels under the fluorescent light source, the projected image is reflected by a system of three mirrors to a solid-state scanning camera. Object width is sensed by the scanning camera; object length is determined by an attached encoding cord linked to the scanning beam. As the leaf sample passes under the light source, the accumulated area in cm^2 is digitally displayed. The errors of measurement

with this type of area meter is probably less than 2% (Hatfield et al. 1976). In some cases, leaves tend to fold and wrinkle as they pass under the light source, causing some differences in the total area measured. However, these errors may be small compared to other sources of variation (Daughtry and Hollinger 1984).

Indirect Methods

In the absence of leaf area meters, leaf area measurements can be made by measuring the length and width of leaves. Leaf area per plant is calculated by multiplying the average number of leaves per plant (n) by the average length (L) and width (W) of a leaf by an area constant K:

$$LA = L \times W \times K \times n \quad [1]$$

K is an empirically derived constant which is obtained by comparing leaf area measured directly with a planimeter vs. leaf area obtained from length and width measurements. Equation [1] works well with short-statured crops.

Indirect methods also use relationships between leaf area and dry weight of leaves (LDW) to estimate leaf area:

$$LA = b_0 + b_1 \times LDW \quad [2]$$

Equation [2] is obtained from experimental data using regression analysis between leaf area vs. leaf dry weight. Theoretically b_0 should be equal to zero and Equation [2] can be rewritten as follows:

$$LA = b_1 \times LDW \quad [3]$$

A b_1 value of 223 cm^2/g is reported for taro (*Colocasia esculenta*), (Shih and Snyder 1984), whereas a relationship LA = 0.396 + 0.680 x LDW, where b_1 is expressed in m^2/g is reported for corn (Daughtry and Hollinger 1984). The coefficient b_1 in Equation [3] varies with plant age, cultivars, and locations, thus some field testing to determine b_1 for a specific site and cultivar is required.

A variation of the above method employs the relationship between leaf area and leaf weight of a subsample to convert the weight of a large sample of leaves into leaf area. LA is calculated as follows (Daughtry and Hollinger 1984):

$$LA = (LA(S)/LDW(S)) \times LDW/n \quad [4]$$

where LA(S) and LDW(S) are leaf area and leaf dry weight on a subsample of leaves, and LDW/n is the average leaf dry weight of n plants. This method uses a small number of plants to estimate leaf area per LDW, and a larger number of plants to estimate LDW.

Direct measurement of leaf area requires few plants, has the lowest coefficient of variation but requires approximately the same amount of time as indirect methods. Despite these limitations, indirect methods are the only choice where area meters are unavailable.

If you have access to a photocopy machine, here are two more ways to determine leaf area.

1. Photocopy the leaf onto graph paper and then determine the area by determining the area that the leaf covers on the graph paper.
2. Photocopy the leaf onto graph paper and send the photocopy to someone who does have a leaf area meter and they can measure it for you.

These methods may be helpful, but keep in mind that some photocopiers distort their images slightly. You may need to determine what the distortion is (if any) and accommodate your figures for this.

REFERENCES

Ayers, A.S. and H.H. Hagihara. 1952. Available phosphorus in Hawaiian soil profiles. Hawaiian Planters' Record 54:81-99.

Chang, Jen-Hu. 1968. Climate and agriculture: An ecological survey. Aldine Publishing Co., Chicago.

Daughtry, C.S.T. and S.E. Hollinger. 1984. Costs of measuring leaf area index of corn. Agron. J. 76:836-841.

Fitzpatrick, T.J., W.C. Porter, and G.V.C. Houghland. 1969. Continued studies of the relationship of specific gravity to total solids of potatoes. Am. Potato J. 46:120-127.

Hatfield, J.L., C.D. Stanley, and R.E. Carlson. 1976. Evaluation of an electronic foliometer to measure leaf area in corn and soybeans. Agron. J. 68:434-436.

IBSNAT. 1988. Technical report 1, experimental design and data collection procedures for IBSNAT. 3rd edition, revised. Department of Agronomy and Soil Science. College of Tropical Agriculture and Human Resources, University of Hawaii, Honolulu, Hawaii.

Kleinkopf, G.E., D.T. Westermann, M.J. Willie, and G.D. Kleinschmidt. 1987. Specific gravity of Russet Burbank potatoes. Am. Potato J. 64:579-587.

Kushman, L.J. A common regression for estimating dry-matter content of carbohydrate-containing commodities from specific gravity. Am. Potato J. 46:234-238.

Murphy, J. and J.P. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta. 27:31-36.

Page, A. L., R.H. Miller, and D.R. Keeney. 1982. Methods of Soil Analysis, Part 2: Chemical and Microbiological Properties. 2nd edition. American Society of Agronomy, Inc., and Soil Science Society of America, Inc., Madison, Wisconsin.

Ratliff, L.F., J.T. Ritchie, and D.K. Cassel. 1983. Field measured limits of soil water availability as related to laboratory-measured properties. J. Soil Sci. Am. 47:770-75.

Ritchie, J.T. 1981. Soil water availability. Plant and Soil 58:327-38.

SCS, USDA. 1972. Soil Survey laboratory methods and procedures for collecting soil samples. *In* Soil survey Investigation Report No. 1. U.S. Gov. Print. Office, Washington D.C.

Shih, S.F. and G.H. Snyder. 1984. Leaf area index and dry biomass of taro. Agron. J. 76:750-753.

Wilson, J.H. and A.M. Lindsay. 1969. The relation between specific gravity and dry matter content of potato tubers. Am. Potato J. 46:323-328.

World Meteorological Organization. 1981. Typhoon operation experiment: Automatic weather stations for tropical cyclone areas. WMO No. 570, Geneva, Switzerland.

Zaehringer, M.V., R.M. Reeve, E.A. Talley, H.H. Dinkly, and R.B. Hyde. 1965. The estimation of total solids from specific gravity measurements. Potato Handbook.