

**Аграрен университет – Пловдив**

**Ваньо Цветанов Георгиев**

**Създаване на модели и софтуер, при-  
ложими при обработка на данни от  
селскостопански изследвания**

**Дисертация**

**за присъждане образователната и научна степен "доктор"**

**Научна специалност "Информатика", шифър: 01.01.12**

Научни ръководители: ст.н.с. I ст. д.с.н. Никола Вичев Колев

проф. д.с.н. Антония Матеева

**Пловдив, 2004 г.**

## Съдържание

<b>УВОД</b> .....	<b>4</b>
<b>ГЛАВА 1 ОБЗОРАН АНАЛИЗ НА СЪСТОЯНИЕТО НА ПРОБЛЕМА</b> .....	<b>7</b>
1.1. АНАЛИЗ НА ИЗПОЛЗВАНИ МЕТОДИ ЗА ОБРАБОТКА НА ДАННИ, ПОЛУЧЕНИ ПРИ ИЗСЛЕДВАНИЯ ОТ ОБЛАСТТА НА СЕЛСКОТО СТОПАНСТВО. ....	7
1.2. МОДЕЛИРАНЕ НА СИСТЕМИ, ПРЕДМЕТ НА СЕЛСКОСТОПАНСКИ ИЗСЛЕДВАНИЯ.....	8
1.3. ИЗВЕСТНИ СОФТУЕРНИ ПРОДУКТИ ЗА ОБРАБОТКА НА ДАННИ ОТ СЕЛСКОСТОПАНСКИ ИЗСЛЕДВАНИЯ	10
1.3.1 <i>Комерсиален софтуер без достъпен изходен код</i> .....	11
1.3.2 <i>Безплатен софтуер</i> .....	13
1.3.3 <i>Софтуер с отворен код</i> .....	13
1.4. ИЗВОДИ: .....	15
1.5. ЦЕЛ И ЗАДАЧИ НА ДИСЕРТАЦИЯТА.....	16
1.5.1 <i>Цел</i> : .....	16
1.5.2 <i>Задачи</i> :.....	16
<b>ГЛАВА 2 МЕТОДИЧЕСКИ ПОДХОД ПРИ МОДЕЛИРАНЕТО НА ОБЕКТИ И ПРОЦЕСИ В СЕЛСКОСТОПАНСКИТЕ ИЗСЛЕДВАНИЯ И ПРАКТИКА</b> .....	<b>18</b>
2.1. СЪЩНОСТНИ ЧЕРТИ И КЛАСИФИКАЦИЯ НА МОДЕЛИТЕ, ИЗПОЛЗВАНИ В СЕЛСКОСТОПАНСКИТЕ ИЗСЛЕДВАНИЯ И ПРАКТИКА .....	18
2.2. РОЛЯ НА МОДЕЛИРАНЕТО В СЕЛСКОСТОПАНСКИТЕ ИЗСЛЕДВАНИЯ .....	22
2.3. ВРЪЗКА НА МОДЕЛИРАНЕТО С ОБРАБОТКАТА НА ДАННИ.....	25
2.4. ЕТАПИ НА МОДЕЛИРАНЕТО ПРИ СЕЛСКОСТОПАНСКИ ИЗСЛЕДВАНИЯ.....	26
2.5. СПЕЦИФИКА НА СИСТЕМНИЯ АНАЛИЗ, ПРИЛОЖЕН В ДИСЕРТАЦИЯТА.....	27
2.5.1 <i>Описване на системите, параметри, структура</i> .....	27
2.5.2 <i>Математическо представяне на системи, състояния и процеси</i> .....	30
2.6. ХАРАКТЕРНИ ОСОБЕНОСТИ НА ОПИТНИТЕ ДАННИ, ПОЛУЧЕНИ ПРИ СЕЛСКОСТОПАНСКИ ИЗСЛЕДВАНИЯ .....	31
2.6.1 <i>Видове данни</i> .....	31
2.6.2 <i>Цифровизиране</i> .....	32
2.7. ИЗБОР НА МЕТОД ЗА ОБРАБОТКА НА ДАННИ.....	32
2.7.1 <i>Групиране, кластерен анализ</i> .....	32
2.7.2 <i>Генерална съвкупност и извадка. Статистически анализ и статистическо моделиране</i> . .....	33
2.7.3 <i>Динамични системи и модели</i> .....	35
2.8. ОЦЕНКА НА АДЕКВАТНОСТТА НА МОДЕЛИТЕ .....	36
2.9. ПАРАМЕТРИЗИРАНЕ И ПОДОБРЯВАНЕ НА МОДЕЛИТЕ.....	39
2.10. ИЗВОДИ: .....	40

**ГЛАВА 3 РАЗРАБОТКА НА СПЕЦИАЛИЗИРАНИ СОФТУЕРНИ ПРОДУКТИ ЗА ОБРАБОТКА НА ДАННИ ОТ ЛАБОРАТОРНИ И ПОЛСКИ ИЗСЛЕДВАНИЯ И ЗА СЪСТАВЯНЕ НА МОДЕЛИ В ИЗСЛЕДВАНИЯТА И ПРАКТИКАТА .....41**

3.1. Модул за представяне и действия с матрици (MATRUNIT.PAS) .....	42
3.2. ФОРМА ЗА ВИЗУАЛИЗИРАНЕ НА МАТРИЦИ (MATRIXFORM.PAS) .....	45
3.3. Модул за пресмятане на статистики (STATUNIT.PAS) .....	46
3.4. Модул за оптимизация.....	48
3.5. Модул за програмно представяне на математически изрази (MATHEXPRUNIT.PAS) .....	48
3.6. Модул за кластер анализ (CLUSTANAL.PAS) .....	52
3.7. Модул за графични изображения (GRAPHUNIT.PAS).....	54
3.8. Програмата ANALYSIS TOOLS .....	55
3.9. Други модули и програми.....	55

**ГЛАВА 4 ЕКСПЕРИМЕНТАЛНИ ИЗСЛЕДВАНИЯ - СЪСТАВЕНИ МОДЕЛИ .....56**

4.1. Моделиране толерантността към вируси на сортове и селекционни линии краставици и фасул .....	56
4.2. Групиране на сортове и селекционни линии чрез кластер анализ.....	59
4.3. Модел на зависимостта концентрация-ефект на инсектициди.....	59
4.4. Модел на замърсяването на почвите с тежки метали в района на КЦМ .....	59
4.5. Моделиране динамиката на изменение на витамин С в стафиди .....	61
4.6. Моделиране динамиката на цветни характеристики на сушени гъби .....	62
4.7. Моделиране многогодишната тенденция в изменението на добива от пшеница и царевича, при отглеждането им в двуполка и прилагане на постоянен режим на торене.....	63
4.8. Изследване влиянието на някои абиотични фактори върху динамиката на диапаузиране при гъсениците на ябълковия плод червей.....	64
4.9. Изследване на ентомофауната по люцерната .....	64
4.9.1 Изследване на гостоприемниковата способност на люцерната .....	64
4.9.2 Изследване на корелацията между полезни и вредни видове насекоми по люцерната ..	65
4.10. Имитационен модел на популация на люцернов листояд .....	65
4.10.1 Системен анализ.....	65
4.10.2 Съставяне на модела.....	66
4.10.3 Параметризиране на модела .....	67
4.10.4 Валидиране .....	67
4.11. Математическо моделиране на взаимоотношенията между колорадски бръмбар и седемточкова калинка .....	69
4.12. Структурна схема за модел на растежа и развитието на посеви едногодишен лук.....	74
4.12.1 Моделиране на мястото на посева.....	76
4.12.2 Моделиране на сорта .....	76
4.12.3 Моделиране на посева .....	76
4.13. Моделиране положението на Слънцето спрямо посевите .....	77
4.13.1 Изчисляване положението на Слънцето .....	78

4.13.2 Изчисляване времето на изгрева и залеза и продължителността на деня.....	83
4.14. Изводи: .....	84
<b>ЗАКЛЮЧЕНИЕ .....</b>	<b>85</b>
<b>ЛИТЕРАТУРА .....</b>	<b>87</b>
<b>ПРИЛОЖЕНИЯ.....</b>	<b>92</b>
<b>ПУБЛИКАЦИИ НА АВТОРА ПО ТЕМАТА НА ДИСЕРТАЦИЯТА .....</b>	<b>92</b>

## Увод

В българските научно-изследователски учреждения, провеждащи изследвания в областта на селското стопанство, има събрано голямо количество информация, получена при извършените там полски и лабораторни опити. При всяко изследване събраните данни се обработват и използват съобразно с поставените по съответната тема задачи, но много изследователи споделят, че разполагат и с данни, които са оставили неизползвани и непубликувани.

Причините за това могат да бъдат различни. Например:

- това, че тези данни са събрани извън рамките на предварително планираната работа и затова и впоследствие са останали недокладвани;
- липсват общоприети и известни методи за обработка на съответните данни;
- не се познават от страна на провеждащия изследването, други възможни методи за обработка на данни, извън общоприетите за дадения вид изследвания.

Предвид на това, че експерименталната дейност в областта на селското стопанство е скъпа и ресурсоемка, неизползването на каква да е част от събраните данни очевидно е нежелателно. Чрез по-внимателно и целенасочено планиране това може да се избегне в бъдещи изследвания, но за вече проведените проблемът остава. Ето защо съществува необходимост от намиране на начини за ново оползотворяване на вече наличните експериментални данни. Тази необходимост се засилва и от трудните икономически условия в момента и невъзможността да се вложат достатъчно средства в нови експериментални изследвания. Липсата на възможности за провеждане на нови експерименти естествено насочва вниманието на специалистите към натрупаните по-рано и останали недостатъчно оползотворени данни. Без провеждане на специално изследване не може да бъде отхвърлена надеждата, че в тези данни се съдържа още някаква полезна за науката, но пропусната преди информация, което да превърне старите данни в незаслужаващ пренебрежение източник на знания за изследваните обекти, още повече, че понастоящем такива източници на са много. С малко изключения данните, из-

ползвани в настоящата работа имат именно такъв произход и са предоставени от колегите, провели описаните опити и измервания.

Но дори и да не се касае за обработване на стари данни, при изтъкнатите, отнасящи се за настоящия момент обстоятелства, намирането на научнообосновани методи за извличане на знания и достигане до достоверни хипотези и изводи за изследваните обекти, допълващи познатите, известни методи, може да се окаже полезно и продуктивно и при бъдещи изследвания.

Математическото моделиране и компютърното симулиране на различни реални системи сега е добре развит научен отрасъл. Натрупаният в него опит позволява създаване на модели на обекти, представляващи интерес в много научни и стопански дейности и често тези модели се оказват изключително полезни. Ето защо изглежда естествен и логичен опитът за асоцииране на моделирането с обработката на опитни данни, направен в настоящия труд.

Интересът на автора към компютърното моделиране е продиктуван и от възможностите и условията за провеждане на научни разработки в България в момента. Например, сега съществуват много трудности (най-вече финансови) за снабдяването на научните организации с нова съвременна, специализирана апаратура. От оскъдните бюджети се предпочита вместо да се закупуват лабораторни уреди с по-тясно предназначение, да се отделят средства поне за някой компютър, който се използва както за научна, така и за канцеларска и административна работа, за комуникация през Интернет, достъп до информация и научни публикации и т.н. Не е изключение на много места един или няколко компютъра да се оказват единствените уреди, с които все още може успешно да се върши някаква научна дейност. Ясно е, че при такива условия всеки научен работник, който желае да постигне резултат в изследователската работа, се изправя пред необходимостта да използва максимално единствения уред, с който разполага - компютъра, включително и чрез създаване за целта на собствени компютърни програми.

Поради липса на финансови средства е силно затруднено и снабдяването със софтуер, приложим за моделиране на обектите, предмет на селскостопански изс-

ледвания. Известно е, че цените на качествените софтуерни продукти за научно приложение са високи. Това също прави актуална необходимостта от създаване на собствени компютърни програми, с помощта на които да бъдат осъществени набелязаните идеи и задачи. Създаването на собствен софтуер обаче се налага не само по финансови причини, то в редица случаи се оказва единствено възможен начин за постигане на набелязаната цел и получаване на желанния резултат, защото колкото и да съществуват различни готови програми, никоя програма не е с неограничени възможности и не може да предложи готови решения за всички възникващи в изследователската работа случаи. И нещо повече, винаги съществува възможност някои елементи от създадения софтуер поради уникалния си характер и специфично предназначение да няма аналози между намиращите се на пазара програми и да послужи за създаване на продукти, имащи потенциал да бъдат предложени на пазара и да станат източник на приходи за научното звено, в което са създадени.

Накрая може да се изтъкне и това, че много научни списания приемат за публикуване само резултати от селскостопански изследвания, обработени с коректен математически апарат, изпълнен на компютър. Това показва, че въпреки съществуващите финансови ограничения, изследователите трябва да разполагат със софтуер, с който да могат да извършат обработката.

## Глава 1

### Обзорен анализ на състоянието на проблема

#### *1.1. Анализ на използвани методи за обработка на данни, получени при изследвания от областта на селското стопанство.*

Селскостопанските специалисти обикновено се запознават с методите за обработката на данни, получени при селскостопански изследвания от литературата по опитно дело и биометрия. В прегледаните от автора книги по опитно дело и биометрия (например: Шанин, 1977; Лакин 1973) и особено в учебната литература по този предмет (Запрянов и Маринов, 1978; Димова и Маринков, 1999) прави впечатление, че в тези книги се обръща по-голямо внимание на практическата страна на експерименталната работа – набелязване на целите, избор на място за провеждане на опита, начини за разполагане на опитните парцели върху терена, агротехническите особености на културите и т.н., но в тях е отделено много малко място на теоретичната страна на експерименталната работа и макар че винаги се споменава, слабо е показана връзката на прилаганите начини за обработка на данните с математическата статистика, а още по-малко (дори и никак) не е показана връзка с моделирането, като изследователски метод. В тази литература, макар и да се излагат някои изчислителни прийоми от статистиката, то дълбочината на представяне на тази материя, според нас, е незадоволителна. Разбира се, има и книги, касаещи обработката на данни от полски опити, написани от математици, в които основен обект на изложение са математическите идеи, представени с пълнота и строгост. В такива книги полските опити се посочват само като един примерен, възможен източник на данни, които могат да бъдат обработени чрез статистически методи. Такава например е книга на Фишер (1958), която може да се нарече класическа.

Недостатъчното осветляване на въпроса в литературата по опитно дело може да бъде компенсирано от съществуващия голям обем литература, посветена на обработката на опитни данни в други близки или по-далечни на селското стопанство науки - селскостопанска техника (Митков и Минков, 1989; Митков и Кардашевски, 1977), метеорология (Кабишева, 1971), физика (Тихонов и Уфимцев,



1988) или курсове по математическа статистика, в които коректно и задълбочено се разглеждат статистическите методи, независимо от областта на приложение (Божанов и Вучков, 1979; Хог и Крейг 1982). Такива източници играят съществена роля за формиране на същностните елементи на настоящата работа, защото предоставят информация, идеи и опит, които са пренесени в нашите изследвания по проблеми от областта на селското стопанство.

### ***1.2. Моделиране на системи, предмет на селскостопански изследвания***

Още от началото на масовото навлизане на цифровите електронно-изчислителни машини в научно-изследователската и стопанската дейност са започнали опити да се прилага компютърно симулиране за решаване на редица проблеми в селскостопанското производство. Моделирането е събудило интерес отдавна и като подход при решаване на проблемите на защитата на растенията (Колективен труд, 1983), екологията (Сиротенко, 1981; Петросян и Захаров, 1986) и опазването на околната среда и взаимоотношението общество – природа (Гвишиани и др., 1986). Натрупването на все по-точни знания за биологичните обекти, естествено води до все по-честа необходимост от прилагане на математическо моделиране при изучаването им. Както изтъкват някои специалисти (Садовски, 1986) това е естествен процес в развитието на селскостопанската наука, следващ общата тенденция към преминаване от качествено към количествено описание на предмета на изследване и все по-широко навлизане на математиката, наблюдавано във всички науки.

Волвач (1978) посочва три етапа при разработване на модели, приложими в растителната защита за представяне на динамиката на развитие на насекомите: първи етап - характерен с описателно разкриване на връзките между факторите на средата и вредните насекоми; втори етап - характеризиращ се с приложение на статистически методи за установяване на емпирични зависимости; и трети етап, при който се преминава към математическо моделиране на динамиката на насекомните популации, разглеждани като системи, развиващи се под въздействие на редица вътрешни и външни фактори. Цитираме тези етапи, защото те могат да се открият и при други модели и представят една тенденция за преминава-

не от качествен, описателен към количествен, математически формализъм и накрая към моделиране на компютър.

Трудно е да се направи удовлетворително пълен преглед на моделите на обектите, изследвани експериментално в селскостопанската наука поради голямото им разнообразие и наличието на вече много голям брой такива. Представа за това разнообразие дава прегледа на редица обзорни монографии по темата, например: Франс и Торнли, (1987); Thornley & Johnson, (1990), и др. Известни са няколко бази данни, достъпни през Интернет, за математически модели, като например:

REM - Register of Ecological Models, поддържана на сървъра на университета в Касел, Германия (<http://dino.wiz.uni-kassel.de/ecobas.html>) и описана в публикациите: Knorrenschild и др., 1996; Hoch и др., 1998.

The CAMASE Register of Agro-ecosystems Models - проект към библиотеката на университета във Вагенинген, Холандия (<http://library.wur.nl/camase/srch-cms.html>).

FMA – Forest Model Archive, на университета Гринуич, Англия (<http://cms1.gre.ac.uk/conferences/iufro/fma/>).

Огромното разнообразие от модели е довело до необходимост от формулиране на препоръки за оформление на документацията, описваща тези модели. Нещо повече, по проекта REM е предложен специален език за описване на модели (Model Interchange Format – MIF) и е създаден софтуерен продукт (ECOBAS Modelling Assistant - EMA), чрез който може да се облекчи цялостно процеса на моделиране – от оформянето на идеята, през съставянето на математическите формули, превеждането им на подходящ език за програмиране и същевременно изготвянето на документацията, както и накрая - регистрирането на модела в базата данни ECOBAS (Benz и др., 2001).

В българската селскостопанска наука прилагането на методите на математическо моделиране също има дългогодишна традиция и разнообразен обхват. Няколко примера: статистическо моделиране на зависимостта почва-добив (Садов-

ски, 1986), моделиране продуктивността на царица (Тончева, 2001), моделиране при агротехнологичен трансфер (Клевцов, 2002).

Ако се разгледат внимателно предпоставките, довели до създаване на всеки от съществуващите сега модели на селскостопански или биологични обекти, ще се забележи, че те са създадени тогава, когато вече са натрупани достатъчно знания или данни за моделираните обекти – данни, публикувани от други изследователи, провели опити или данни получени в специално планирани във връзка с моделирането опити. От това следва, че за създаване на модел винаги се тръгва от конкретен набор от данни и че моделирането представлява естествено продължение на процеса на изследване, в който се използват тези данни. Но тази тясна връзка между данни и модел много пъти става причина един модел да не може да се използва в други ситуации, в които не е в наличност подобен по обем набор данни. Това означава още, че всеки модел почти не може да се използва за други цели, освен за целите, за които е създаден и са изключително редки случаите, в които за решаване на дадена задача, може да се използва готов вече модел, без той да се подлага на съществени промени. Ето защо при много задачи може да се окаже по-удачно, вместо да се изразходва време за търсене на готов модел, да се прави нов модел, съобразен с наличните данни и цел на поставената задача. По-ценно може да бъде вместо да се използват готови модели, да се изучава и трупа опит в прилагане на принципите и начините за построяване на модели, като успоредно с това се създават предпоставки (във вид на знания, софтуер и др.) тази дейност да се провежда все по-успешно с всеки нов модел, към който се пристъпва. А съществуващите, създадени вече към момента модели следва да бъдат един богат източник за идеи, при изграждане на нови модели, подходящи за конкретните ситуации и изследователски задачи.

### ***1.3. Известни софтуерни продукти за обработка на данни от селскостопански изследвания***

Количеството на софтуера, който се създава в света расте лавинообразно и по същия начин се увеличава и този, предназначен за обработка на експериментални данни в различни клонове на науката. През последните години силно се уве-

личава значението на Интернет, като източник на информация и достъп до много от софтуерните продукти. Голяма част от изнесената тук информация е именно от този източник. Според предназначението, имащо отношение към разглежданата тема ще се спрем на програмите за:

- обработка на опитни данни,
- съставяне на модели.

Освен това ще разделим софтуера, за който беше събрана информация и на следните категории:

- комерсиален софтуер, който се предоставя без изходния код;
- безплатен софтуер, в много случаи предлаган също без изходен код и
- софтуер с отворен код (Open source).

С разрастването на Интернет се появи и една нова категория софтуер, който можем да наречем web-базиран - програми, представляващи JavaScript, Java аплети или ActivX контроли и т.н., които се изтеглят от съответния сървър в момента на използването им и рядко се съхраняват на компютъра на потребителя, както също и сървърни приложения, които се изпълняват от сървъра, а не от компютъра на потребителя. Един сайт, предлагащ услуги от този сорт е: <http://www.statcrunch.com/>.

### 1.3.1 Комерсиален софтуер без достъпен изходен код

Комерсиален ще наричаме софтуера, за ползването на който може да се получи лиценз само срещу заплащане. Известни комерсиални продукти са:

**Mathematica** (към момента във версия 5) – продукт, предназначен за математически пресмятания (<http://www.wolfram.com/products/mathematica/>; Воробъев, 1988). Цена: \$595.

Продуктите на **The MathWorks Inc.** – широка гама продукти (MathLab – описана в Потемкин, 1999; SimuLink и др.) предназначени за математически пресмятания, обработка на данни и моделиране. (<http://www.mathworks.com>).

**SPSS Base for Windows** (сега във версия 12) – продукт за статистически анализ, разработка на SPSS Inc. (<http://www.spss.com/spss/> )

**Minitab** (към момента във версия 14) – продукт за статистически анализи. (<http://www.minitab.com>). Цена: \$1195.

**Statistica** (към момента във версия 6) – мощен продукт с разнообразно предназначение, разработка на StatSoft. (<http://www.statsoft.com>; Kornbrot, 2002). Цена: \$795 + \$120 за доставка до България.

**SYSTAT** (последна версия 10.2) – според рекламата на разработчика Systat Software Inc. (SSI) това е най-мощният и бърз статистически пакет. (<http://www.systat.com/>). Положителните качества на продукта не са само реклама, а се изтъкват и в журнални публикации - Review of the SYSTAT v.10.2 software package 2003. Цена: \$1299. Предлага се и в пробна (trial) версия.

**S-PLUS** – продукт за статистически анализи във версии за Windows и Unix. (<http://www.insightful.com/products/splus/>)

**SAS/STAT** програма за статистически анализ предимно с бизнес насоченост. (<http://www.sas.com/technologies/analytics/statistics/stat/index.html>)

**JMP 5.1** – пакет за статистически анализ от SAS Institute Inc. (<http://jmp.com/>) Има версии за Windows, Macintosh и Linux. Цена \$995. Предлага се безплатна демонстрационна версия.

**Origin 7.5** – професионална система за построяване на графики и диаграми и за статистически анализ. Разработка на OriginLab Corporation. (<http://originlab.com>) Има безплатна демонстрационна версия.

**Stella** (към момента във версия 8) – програма с визуален интерфейс за съставяне модели на системи от различни области – биология, икономика, образование. Разпространител: High Performance Systems, Inc. (<http://www.hps-inc.com/stellavpsr.htm>). Подходяща е за бързо съставяне на сравнително прости модели с цел проверка на хипотези. Цена \$199.

Brian Francis в рубрика на сп. Applied Statistics описва и други продукти: **LOGXACT TURBO**, в момента е във версия 5.0, разработка на CyTel Software Corporation (<http://www.cytel.com/>), специализирани в създаването на статистически програми за обработка на данни чрез т.нар. точни статистики. Цена \$595.

(Francis, 1995 г.) **STAT-ITCF** – програма за статистически анализ, приложима в селското стопанство (Francis, 1994 г.)

### 1.3.2 Безплатен софтуер

Най-често това са програми, писани от отделни ентузиастасти – студенти, докторанти, преподаватели и др., създали съответните продукти паралелно с работата си по друга, по-важна за тях задача. В много от случаите авторите при поискване или при съблюдаване на определени, поставени от тях условия, предоставят и изходния код на продуктите си. Количеството на такъв софтуер е огромно и съставът му много динамично се променя. Много от програмите скоро "изчезват от сцената" поради това, че техните автори изгубват интерес или ентузиазъм да продължават да ги поддържат и на тяхно място се появяват други. Качеството на подобни програми не е особено високо.

Ще посочим няколко продукта от тази категория:

**IRISTAT 4.4** – безплатна програма за статистическа обработка на данни, предлагана от Международния институт за изследване на ориза (IRI - International Rice Research Institute, <http://www.iri.org> ).

**ViSta** – програма, написана от Forrest W. Young (<http://forrest.psych.unc.edu/research/> ), оригинална по замисъл заради визуалния начин на представяне на етапите на извършваната от потребителя статистическа обработка (Dumouchel & Lane 1995; Pregibon, 1995)

**OpenStat** – програма за статистическа обработка на данни с автор Bill Miller (<http://www.statpages.org/miller/openstat/> ). Предоставен е и програмният код.

### 1.3.3 Софтуер с отворен код

Към тази категория софтуер ще причислим програмите, които отговарят на определенията, дадени от организации като Open Source Initiative (OSI) (<http://www.opensource.org/> ) или GNU Project (<http://www.gnu.org/> ). Този тип софтуер добива все по-голямо разпространение благодарение на разрастването на Интернет. Много известна е операционната система Linux и предназначенията за нея програми с отворен код с различно приложение. Университетите се явяват важни центрове, в които се създават и разработват такива програми, защото мно-

го от специалистите в различни области там проявяват интерес и към програмите с отворен код. Разработена е и система, позволяваща лесно пренасяне на програми с отворен код от Unix или Linux платформи към Windows – Sygwin (<http://cygwin.com/>) на RedHat Inc. (<http://www.redhat.com/>) (Linfield, 2002) Това прави възможно и в Windows използването на програми с отворен код, писани за Unix. Но дори и без тази възможност, редица проекти с отворен код се разработват паралелно за няколко операционни системи.

**Scilab** – пакет с разнообразни възможности, предназначен за научни изследвания. (<http://scilabsoft.inria.fr/>)

**R-project** – статистически пакет с отворен код. (<http://www.r-project.org/>) Включва команден интерпретатор, чрез който може да се изпълняват множество подпрограми. Има възможност неограничено да се допълва с нови подпрограми и команди.

**Octave** – пакет за числени пресмятания (<http://www.octave.org/>), намиращ често приложение в изследванията по физика. Построен е също като команден интерпретатор с възможности за допълване с нови подпрограми.

Освен посочените цялостни в известен смисъл софтуерни продукти, като отворен код в Интернет може да се намерят и библиотеки с реализации на отделни алгоритми за обработка на данни на различни езици за програмиране: FORTRAN, C, Pascal. По-известни и богати на алгоритми са: StatCodes (<http://www.astro.psu.edu/statcodes/>) и StatLib (<http://lib.stat.cmu.edu/>).

Прегледът показва, че съществуващият комерсиален софтуер е скъпо струващ и в определена степен специализиран в едно или друго направление. Това означава, че дори и да се намерят средства за закупуване на даден продукт, той не може да реши всички проблеми и може да се наложи закупуване и на друг, че и на трети и после да се окаже, че на практика се използват само много малка част от възможностите на целия закупен софтуер - твърде голямо разточителство за нашите възможности.

Използването на софтуер с отворен код също има редица неудобства. Този софтуер се създава от голям брой доброволци, които обикновено не отделят мно-

го време за подробна документация. Ориентирането в програмния код на някой обемист проект, изисква доста програмистки опит и прекарване на много време в разглеждане на кода, докато се открие кои части от него изпълняват близка на стоящата пред нас задача. После, добавянето или променянето на някакви детайли в такъв проект рядко води веднага до желан резултат, защото всяка промяна трябва да бъде съобразена с други негови части. Ако не се познава добре проекта в цялост, промените по-често водят до провал във функционирането на крайната програма. Когато имаме конкретна задача и срок за намиране на решение, едва ли можем да се впуснем в такава бавна и рискована работа по модифициране на някой проект с отворен код.

Изходът е в създаване на собствен софтуер, предназначен за решаване само на поставените изследователски задачи. Вярно е, че програмирането, макар и изискващо доста знания, се явява рутинна дейност и писането на малки програми за решаване на конкретни задачи трудно може да се окачестви като научна работа, но когато програмирането представлява незаменима част от решаването на по-обща задача за обработката на експериментални данни, неговото значение и роля не може да бъдат подценени.

#### ***1.4. Изводи:***

1. Обработката на опитни данни, получени при селскостопански изследвания, традиционно се прави по познати методи, предписвани в зависимост от схемата на залагане на опита. При запознаване на селскостопанските специалисти с тези методи обаче слабо се изяснява връзката на тези методи с математическата статистика, на основата на която са съставени.

2. Много пъти се провеждат и опити по нетрадиционни схеми, за които няма предварително предписани методи за обработка на данните. Тогава провеждащите опитите селскостопански специалисти, извършват обработката на получените данни, разчитайки по-скоро на собствената си логика и интуиция, отколкото на математически коректни разсъждения. Други данни пък се събират допълнително, поради това, че са събудили интерес у експериментатора; за тях също може да няма готови методи за обработка. В такива случаи се налага да се прилагат



методи, основаващи се на математическата статистика, които може да се заемат от други експериментални науки - технически дисциплини, физика, и т.н. или директно да се разглеждат нещата от гледна точка на математическата статистика и да се съставят подходящи за случая статистически методи.

3. Статистическите методи също могат да имат ограничена приложимост и може да се наложи търсене и на други, по-универсални идеи за добиване на знания за изследвания обект по данните от проведения експеримент.

4. Опитът при съставянето на модели на изследваните обекти показва, че до съставяне на модел се стига след събиране на достатъчно данни за обекта. Това ни навежда на мисълта да направим връзка между моделирането и обработката на данни, т.е. когато трябва да обработим данни да се опитаме да съставим модел на основа на тези данни и в резултат на моделирането да направим изводи за изследваната система. Макар и да има основания да се счита, че такъв подход за прилагане на моделиране при обработка на данни може да се окаже плодотворен, на автора не са известни случаи този подход да е изтъкван и прилаган явно при обработка на данни от селскостопански изследвания.

5. Осъществяването на опит да се приложи моделиране в различни случаи за обработка на данни от селскостопански експерименти, с цел установяване продуктивността на тази идея, изисква подходящ софтуер, който е най-добре да бъде специално съставен за целта, за да може да се модифицира и допълва според необходимостта във всяка конкретна задача.

## ***1.5. Цел и задачи на дисертацията***

### **1.5.1 Цел:**

Да се създадат модели на обекти, представляващи предмет на селскостопански изследвания, както и компютърни програми, с чиято помощ се реализират тези модели и да се използват при интерпретиране на експериментални данни.

### **1.5.2 Задачи:**

1. Да се очертае обща схема за последователността от действия при изграждане на моделите, приложима към всички разгледани в работата обекти.

2. Да се анализират нуждите от програмно осигуряване за реализиране на конкретни модели по тази схема и да се напишат необходимите програми.

3. Да се приложи съставената схема и да се създадат модели на различни системи, във връзка с възникнали в практиката на изследователската работа случаи за обработка на опитни данни.

4. За всеки от реализираните програмно модели да се направи анализ на изследваната система на основата на резултата от моделирането.

5. Да се оцени полезността на създадените модели и програми като допълнение към методите за статистическата обработка на данни.

## Глава 2

### Методически подход при моделирането на обекти и процеси в селскостопанските изследвания и практика

Във връзка с първата от поставените задачи ще бъде направен кратък анализ на редица иначе добре известни от научната литература понятия. Необходимост от такъв анализ има, защото в детайлите на тези понятия може да се забележат различия между отделните литературни източници. Този анализ се прави, за да се посочат точните нюанси на тези понятия, към които авторът има предпочитание и по такъв начин да бъдат вложени по един непротиворечив начин в контекста на настоящото изследване. Това не е литературен преглед, а по-скоро необходимо обобщение с цел да се направи едно теоретично обосноваване на приложния подход за моделиране.

#### *2.1. Същностни черти и класификация на моделите, използвани в селскостопанските изследвания и практика*

Моделът е изкуствена система, която се създава, за да наподобява в някакъв смисъл друга реална система. Всеки модел показва определена, макар и абстрактна прилика със своя прототип - реалната система, на която съответствува. Колкото тази прилика е по-пълна, толкова е по-добър моделът.

По своята природа моделите се разделят на две групи: *материални* модели (когато една материална система се пресъздава чрез друга, също материална система, например умален макет) и *абстрактни* модели - когато на дадена система се съпоставя не материална, а абстрактна система от математически или други понятия. Материалните модели не са предмет на настоящия труд.

С цел да покажем универсалността и приложимостта на идеите на моделирането в максимално широк кръг от изследователски задачи ще изходим от схващането, че *модел представлява изобищо и самото знание* за даден обект. Подобна теза застъпват Перегудов и Тарасенко (1989). Основанието за това схващане е, че знанието е система от понятия, която отразява (съответства на) детайлите в структурата, поведението и връзките на обекта с други обекти, и очевидно може да се разглежда като система, имаща прилика с реалния обект - като

модел на този обект. В този смисъл всяко изследване, което цели уточняване, коригиране или установяване на нови знания за даден обект, можем да разглеждаме като дейност за създаване на модел на този обект.

Към *абстрактните* модели спадат математическите и имитационните модели. При математическите модели реалната система се представя чрез понятия от математиката: вероятности, случайни величини, функции, уравнения или системи от уравнения и т. н.. При имитационните модели реалната система се представя чрез понятия от компютърното програмиране: структури от данни и подпрограми и програми за обработка на данни. Очевидно най-тясна връзка с компютърната техника има имитационното моделиране, но и за математическо моделиране също се използват компютри. Всички модели, за съставянето, представянето или използването на които се използват компютри, в настоящата работа ще наричаме *компютърни модели*.

*Математическите модели* може да се класифицират от една страна по математическия апарат който използват, а от друга - по дълбочината на представяне на моделираната система.

Математическият апарат, използван при съставяне на математически модели, може да е различен - теория на вероятностите, математическа статистика, линейна алгебра, математически анализ и т.н.; какъв точно, зависи от конкретната задача, която трябва да бъде решена.

В зависимост от това, дали в даден модел участва като параметър времето, различаваме два вида модели: статични и динамични. *Статичните* модели представят състоянието на моделираната система в даден момент от време, без да отразяват нейното развитие във времето. В тях времето не участва като параметър. *Динамичните* модели описват измененията, настъпващи в системата с течение на времето и в тях времето участва като параметър. Динамичните модели от своя страна могат да се отнасят за по-къс или по-дълъг период от време. *Късопериодните* модели показват развитието на системата за обособен интервал от време (час, денонощие или друг), много по-кратък от времето на живот на системата, докато *дългопериодните* модели представят развитието на системата

през период, съизмерим с времето на живот на системата, като възпроизвеждат цял етап от развитието ѝ или цялото ѝ време на живот.

От друга страна се разграничават детерминистични и стохастични модели. *Детерминистичните* модели при еднакви входни данни водят до еднозначно определен, един и същ за тези данни резултат. Докато *стохастичните* модели при еднакви входни данни водят до различни изходни резултати, т. е. при тях, както и в реалните системи, е в наличност някаква случайна компонента.

Thornley и Johnson (1990) посочват, че всяка система попада в някаква йерархия на системите, т. е. от една страна представлява подсистема на други по-големи системи, а от друга - самата тя се състои от елементи, представляващи по-малки системи. Това естествено намира отражение при моделирането. Когато моделите отразяват единствено йерархичното ниво на изследваната система, без да се влияят нито от наличието на по-ниски, нито на по-високи нива, споменатите автори говорят за *емпирични* модели (empirical models). Може да се каже още, че емпиричните модели само описват поведението на системите, но не го обясняват. Моделите, които преминават и към по-ниските йерархични нива, т.е. разглеждат системата като съставена от по-прости елементи и обясняват поведението ѝ чрез взаимодействието и взаимовръзките между тези елементи и по този начин проникват в механизма на явленията, посочените автори наричат *механистични* модели (mechanistic models). И накрая, моделите, които отразяват и влиянието на по-горните нива на йерархия, т. е. обясняват поведението на системата и чрез ролята ѝ като елемент в по-широко обхватни системи, те наричат *телеономични* модели (teleonomic models).

Нека обърнем малко повече внимание на емпиричните модели. В тази група модели поставяме моделите, в които по опитните данни се установяват представими аналитично връзки между параметрите на изучаваната система или връзки между външни параметри и параметри на системата. Зависимостите се представят аналитично с подходящи функции на една променлива, или на повече променливи.

Често пъти емпиричните модели се смесват със статистическите регресионни модели, но по наше мнение е редно да се прави разликата между тези два типа модели. При статистическия регресионен модел влизащите в модела параметри на системата се разглеждат като многомерна случайна величина, която има някакво разпределение и задачата по съставяне на модела е задача от математическата статистика - задача за установяване параметрите на това разпределение. При емпиричните модели не се използват понятия от математическата статистика (макар, че използваната терминология е много сходна), използват се понятия от математическия анализ и от линейното или нелинейното оптимизиране.

Задачата по съставяне на един емпиричен модел са свежда до два етапа:

1. Избор на функция от подходящ вид, която да играе ролята на модел.
2. Определяне на параметрите на функцията от избрания вид, които дават най-точно приближение на експерименталните данни.

Първият етап се провежда с помощта на средства за визуализиране на експерименталните данни (компютърни програми за построяване на графики) и по разположението на експерименталните точки се избира подходяща математическа функция.

При вторият етап се формулира подходящ критерий за точност на модела и се решава задача за намиране на стойности на параметрите, които осигуряват максимална точност. По същество това е оптимизационна задача и начинът за решаване на тази задача зависи както от избрания критерий за точност така и от вида на модела.

Най-често като критерий за точност на модела се използва сумата от квадратите на разликите между изчислените с модела  $\hat{y}_i$  и получените в експеримента  $y_i$  стойности:

$$(1) \quad S = \sum_{i=1}^N (\hat{y}_i - y_i)^2 .$$

Търсят се такива параметри, които осигуряват минимум на тази сума. Това е известният метод на най-малките квадрати.

Когато емпиричният модел е линейна функция или дори и в по-общ случай – полином, задачата за оценка на коефициентите, осигуряващи минимум на  $S$ , се свежда до решаване на система линейни уравнения. Съществуват и някои функции, които чрез подходящо преобразуване на данните могат да се сведат до линейни зависимости. В още по-общи случаи, при използване в качеството на модели на други функции, задачата може да се реши чрез числени методи за намиране на минимум на  $S$  като функция на търсените параметри.

## ***2.2. Роля на моделирането в селскостопанските изследвания***

Смисълът на моделирането се състои в това, че се заменя една реална система с друга система (модел), която по много причини е по-удобна за боравене и вместо да се извършват манипулации върху реалната система, такива се извършват върху модела. Поради очакваната прилика на модела с истинската система се предполага, че реакцията на модела спрямо приложените манипулации е същата, каквато би имала реалната система спрямо същите или подобни манипулации.

Значението на моделите и моделирането се изразява в два аспекта: познавателен и практически (Thornley, 1990).

Връзка с познавателното (гносеологично) значение на моделирането има това, което вече беше изтъкнато, че дори самото изучаване (изследване) на даден обект може да се разглежда като вид моделиране на този обект, като в този случай създаваният модел е знанието за реалния обект. Но дори и когато моделирането не се разглежда в такъв абстрактен вид, а се има предвид моделиране в прекия смисъл (целенасочено създаване на материални или абстрактни модели), може да се каже, че всеки модел въплъщава знанията на своите създатели за реалната система и е своеобразна проверка за тези знания (Бихеле и др., 1980). А именно, ако на основата на наличните знания се състави добър модел, това означава, че и самите знания са пълни и точни.

При построяването на модели почти винаги се правят редица хипотези за поведението на реалната система. Ако построенният модел се окаже адекватен, това може да се счита за потвърждение на направените хипотези. Откриването на причините за разликите в поведението на модела с това на реалната система на-

лага отхвърляне на някои хипотези. По такъв начин се получават нови знания за моделираната система и моделирането се превръща в метод за изследване на тази система.

Представените в настоящата работа модели на специфични обекти са модели, имащи повече познавателно значение, защото са създадени във връзка с обработката на данни от селскостопански експерименти, а експериментите винаги целят събиране на нова информация и знания за изследваните обекти.

Моделирането има и ще има все по-голямо практическо значение за производството. Всеки производствен процес се състои в управляване поведението на дадена система (производствено предприятие, машина, посев, продуктивно животно или друга) с цел привеждането или поддържането ѝ в определено състояние. За направляване развитието на системата към желаното състояние трябва да се познават ефектите от прилагането на различни въздействия върху нея и във всеки момент да се взема решение какво въздействие да се приложи. Използването на модел при вземане на решение, може да има значителен ефект за постигане на крайния резултат. Наборът от възможни варианти на въздействие се изпробва върху модел и се избира онзи, който е най-ефективен. Специалистите, използващи модел, не е задължително да притежават всички знания за управляваната система, защото тези знания стават достъпни за тях чрез модела, вградени са в този модел. Чрез практическото прилагане на моделирането се осъществява пряко проникване на най-новите научни знания там, където те принасят най-забележима полза - в производството. Създаването на модели с пряко практическо приложение в производството обаче не е предмет на настоящия труд.

Опитът на автора в работата по компютърно моделиране на селскостопански обекти го е довел до убеждението, че **моделирането винаги е съподчинена дейност**, част от друга, по-обща задача и всички дейности при моделирането винаги са вплетени с дейностите, свързани с решаването на тази по-обща задача. Например, моделирането на един посев е винаги подчинено на други научни или практически задачи, свързани с този посев. В тази връзка на моделирането следва да се гледа като на изследователски метод, само като на един от възмож-



ните начини за решаване на даден проблем и достигане до поставената цел. Дали ще се предпочете моделиране пред друг начин на работа зависи от критериите и предпочитанията на хората, заети с решаването на конкретния проблем. Само по себе си моделирането не може да бъде цел, то е само средство и ако не се отчита тази негова съподчиненост в рамките на една по-мощна дейност, съществува опасност от неговото надценяване, което може да доведе дори и до обезсмислянето му - до създаване на модели без практическа или познавателна стойност.

Следователно, за ценността на модела трябва да се съди не по-качествата на модела сам за себе си, а по това дали той изпълнява изискванията, произтичащи от конкретната задача - дали той е полезен и ценен в конкретният случай. Много често се сравняват различни модели, като само се съпоставят по сложност, пълнота, точност и т.н., а не се вземат предвид целите, заради които е създаден всеки модел и условията в които се налага да работи. Това води до незаслужено подценяване на едни модели и надценяване на други, или до опити да се прилагат много "висококачествени" (според определени критерии) модели в ситуации, в които не могат да се задоволят изискванията им към входни данни.

Моделирането е **многодисциплинарна дейност** и изисква съгласувана съвместна работа на селскостопанските специалисти, поставящи задачите и целите на експериментите и специалистите по компютърно моделиране, които съставят модели, приложими за постигане на тези цели. Но специалистите по моделиране трябва да вземат участие не само в крайната обработка на резултатите от вече проведени експерименти, а трябва да имат възможност да дадат своето мнение и да участвуват още в най-ранния етап, още при поставянето на задачата и планирането на експериментите.

Моделите, описани в този труд, са съставяни във връзка с изследвания на различни въпроси от областта на селскостопанската наука. Повечето са създадени на един късен етап от работата по всяко изследване. Първо е проведено изследването, с него са решени дадени задачи, постигнати са определени цели, без да се прави или дори мисли за моделиране и едва след това, в някои случаи след години, се е появила идеята, че на основата на получените данни може да се със-

тавят и някакви модели и чрез тях да се достигне до нови изводи или да се извлече информация, пропусната при предишния анализ. Това, разбира се, е много порочна практика и не може да се очаква, че при подобни обстоятелства ще се получат особено ценни резултати, но за съжаление, поради инертността на системата тази практика така и не можа да се преодолее нито веднъж през годините на работа. Но това е причината в настоящия труд да се поставя акцент върху моделирането само при обработката на получени вече данни, а не върху приложение на моделирането в цялостното изследване, което според автора би било много по-плодотворно.

### ***2.3. Връзка на моделирането с обработката на данни***

Необходимостта от моделиране (създаване на модел или използване на вече съществуващ модел на изследваната система) при обработката на данни произтича от обстоятелството, че данните са източник на нова информация за коригиране или допълване на нашите знания за изследваната система и обработката на данните се състои именно в това, да се установи как данните се съгласуват с наличните вече знания за системата и как те биха коригирали или допълнили тези знания.

Както беше посочено, самото знание е един вид модел на системата и от тук следва, че обработката на данни, в същност, е една процедура за валидиране на този модел - установяване дали данните отговарят на модела (знанието) и ако не отговарят в какво точно не отговарят и как следва от това да се промени модела (знанието), за да се получи съгласие с данните.

Гледайки по този начин на обработката на данни следва, че може да се очаква, че една обработка на опитни данни може да е още по-пълноценна, ако знанието, което имаме за изследваната система е представено чрез подходящ неин модел, който отразява не всички, а само част от знанията ни и то тази част, която непосредствено се отнася за данните, които се обработват.

#### **2.4. Етапи на моделирането при селскостопански изследвания**

Успешното моделиране на един обект изисква преди всичко ясно формулиране на поставената задача, определяне на системата, която подлежи на моделиране и провеждане на подходящ анализ на тази система (Митков и Минков, 1989). Такъв анализ се нарича системен анализ. При него се борави с понятията като: система, околна среда, подсистеми, параметри, вход, изход, прави и обратни връзки, управления и т. н. Системният анализ се явява **първи етап** на моделирането и при него се работи с понятията от научната област, имаща за предмет на изучаване системи от разглеждания вид.

При **втория етап** на базата на проведения системен анализ се намира подходящо съответствие между елементите на моделираната система и понятията от научната област, която разглежда системи от рода на съставяния модел. Когато се съставя математически модел - от математиката, когато се съставя имитационен модел - понятията от компютърното програмиране и т. н. На този етап се борави с понятията и средства и от двете научни дисциплини. Това е етапът, в който фактически се съставя моделът и той започва да се появява като такъв на бял свят.

**Третият етап** представлява анализиране на съставения модел сам по себе си от гледна точка на научната дисциплина, която изучава такива модели. Ако това е математически модел, се анализира като математически обект – установява се дали е построен коректно, търсят се начини за решаване и се намират неговите решения. Ако е имитационен модел, се съставят и настройват съответните програмни модули и се регистрира резултатът от тяхната работа.

**Четвъртият етап** се състои в изучаване на поведението на модела и сравняването му с това на моделираната система. Установява се неговата адекватност, точност, надеждност и т. н. Това е т. нар. валидиране на модела.

Трябва да подчертаем, че изтъкването на тези етапи не означава, че те съществуват наистина диференцирано и следват последователно един след друг във времето. Описаната етапност по-скоро отразява типовете дейности, които се изпълняват в процеса на моделиране и е напълно възможно във всеки момент едновременно да се провеждат дейности от няколко от посочените етапи.

## **2.5. Специфика на системния анализ, приложен в дисертацията**

Когато пред нас стои масив от данни, получени при някакво експериментално изследване първото, което следва да се направи преди тяхната обработка е да се проследи как са получени и какво отразяват, да се уточни поведението на каква система и в какви условия е отразено в тях, т.е. да се извърши системен анализ.

Да направим системен анализ означава:

1. Да определим ясно системата, която изучаваме - какво включва тя и какво остава извън нея, да очертаем точните ѝ граници;
2. Да определим частите, от които се състои тази система - подсистемите ѝ;
3. Да определим изучаваната система представлява ли част от по-голяма система и каква роля играе в тази по-голяма система;
4. Да опишем всяка от подсистемите;
5. Да определим връзките между подсистемите и между подсистемите и външните системи.

Докато първите три от изброените елементи не се нуждаят от повече пояснения четвъртият, изглежда трябва да се разгледа по-обстойно.

### **2.5.1 Описване на системите, параметри, структура.**

Един възможен начин за описване на дадена система е *словесното описание*. Той може да се счита за универсален и очевидно това е най-разбираемият за хората начин. Така се описва обекта на дадено изследване в учебниците, монографиите, научните трудове и т.н. - във всички случаи, предназначени за ползване от хора. Предимството на словесното описание е неговата не обвързаност с някакъв формализъм, което му позволява да се прилага и за системи, които не попадат в никакви формални схеми и строги рамки. Словесно може да се опише всяка нова система. Именно чрез словесни описания хората представят и си предават знанията за даден обект. В областта на изследванията по изкуствен интелект се разработват методи за представяне на знания и в други форми, подходящи за компютърна обработка (Стойчев и др., 1989), но това вече е извън предмета на настоящия труд.

Най-подходящ за целите на моделирането при обработка на данни от селскостопански експерименти е описанието на системите *чрез набори от параметри*, представящи качествените и количествените характеристики на системите.

Всеки параметър описва една характеристика на разглежданата система, има определено име (обозначава се с някакъв термин) и притежава някакво пространство, в което може да търпи изменения. (Например: добив, маса на растение, форма на плода и т.н.) Пространствата, в които се изменят параметрите могат да бъдат крайни или безкрайни множества. В тези множества може да съществува естествено произтичаща от природата на дадената характеристика подредба, но може и да не съществува такава. В повечето случаи обаче може да се установи взаимно еднозначно съответствие между пространството на изменение на даден параметър и множество от числа - крайно или безкрайно множество от цели, рационални или реални числа. Установяването на такова съответствие е много полезно и удобно, защото то отваря пътя за лесно компютърно представяне на всеки параметър. Това съответствие ще наричаме *числено представяне*. Но такова съответствие не винаги може да бъде установено по един естествен и заложен в природата на разглеждания параметър начин. Например: Ако се разглежда характеристиката маса на плодовете, численото представяне е съвсем очевидно - измерената стойност на масата на всеки плод е число. Но ако става дума за характеристика, като форма на плодовете, численото представяне вече става малко изкуствено и не така директно. Формата на плода се описва словесно: кръгла, продълговата, пирамидална, и т.н. Специалистите по съответната селскостопанска култура разполагат със списък на възможните форми. Този списък съвсем формално може да бъде подреден, например по азбучен ред или по друг начин, при което всяка възможна форма придобива номер, отговарящ на мястото на формата в така подредения списък. Полученият номер можем да използваме за числено представяне на параметъра форма, но очевидно, това представяне е твърде изкуствено и формално и може да се направи по различни начини. Тези параметри чиито числени представяния се получават по естествен, произтичащ

от природата на самите параметри начини ще наричаме *количествени параметри*, а признаците на системата, които те описват - количествени признаци на системата. Тези параметри чиито числени представяния се получават чрез формална, привнесена изкуствено процедура, ще наричаме *качествени параметри*. Но независимо дали се касае за количествени или качествени параметри за целите на компютърната обработка на данни, която ще прилагаме, ще приемем, че винаги ще правим числено представяне на параметрите и ще извършваме обработката само на числено представени параметри.

Някои от параметрите на изучаваната система се установяват чрез непосредствено наблюдение върху системата, или чрез измерване. Тези параметри отразяват свойства на системата, проявяващи се непосредствено в процеса на съществуването ѝ. Всеки изследовател обаче притежава и представи за изучаваната система, които не подлежат веднага на явна формулировка, още по-малко пък на числено представяне. Такива понятия могат да имат дори интуитивен характер. Тези представи се изграждат и доизясняват в резултат на натрупването на опит и също представляват част от знанието на човека за изследвания обект. Например: специалистът по дадена селскостопанска култура има представа за пригодността на даден сорт към условията на определен район или за резистентността на този сорт срещу дадена болест или неприятел. Очевидно на споменатите тук характеристики "пригодност" и "резистентност", много трудно може да се припише множество от възможни стойности и после да се даде числено представяне. Но трудно не означава невъзможно. При един добър последователен анализ на произхода на представите за въведените понятия, последователно може да се въвеждат нови параметри, които се получават от параметрите, участващи при формиране на представата за тези интуитивни понятия. Така възникват и нови параметри, чието въвеждане и определяне произтича не от непосредственото наблюдение върху поведението на системата, а от анализа и обобщаването на това поведение, от натрупващите се знания на изследователя за системата и от поставената за решаване задача.

И така системата, която подлагаме на моделиране се характеризира с набор от параметри, като всеки от параметрите може да има стойност принадлежаща на характерно за този параметър множество от възможни стойности. Всички характеристики, дори и тези, за които отначало представата ни е неясна и малко интуитивна, могат да бъдат представени с подходящи параметри и да добият съвсем определено количествено изражение. Така че изследваната система за целите които си поставяме, можем да отъждествим с нейния набор от параметри, т.е. да приемем, че всяка система за нас не е нищо друго освен (нищо повече от) един набор от числени параметри. Още повече, че представянето на една система чрез нейните параметри в случая разглеждаме във връзка с обработката на данни, а данните са измерените по време на проведения експеримент стойности на параметрите.

Съществува определена връзка между *понятието структура* на системата и набора ѝ от параметри. Отхвърлянето на някои параметри, означава опростяване на системата и на нейната структура, а добавянето на нови параметри - усложняване. Доколко сложна е структурата на дадена система, с други думи колко на брой параметри са необходими за описанието ѝ, зависи от поставената задача и от природата на системата. Системите с малък брой параметри естествено наричаме прости, а тези с голям брой - сложни.

Когато извършваме системен анализ и описваме изследваната в даден експеримент система във връзка с обработката на получените данни, винаги ще се ограничаваме с разглеждане на възможно най-проста система - такава система, която може да се опише само с параметрите измервани в опита и за които стойности имаме данни. Включването на други параметри ще бъде ограничено и ще се допуска само, ако има принципна възможност чрез оптимизационни методи да бъдат установени техните стойности.

### **2.5.2 Математическо представяне на системи, състояния и процеси.**

Обектите, които се изследват в селскостопанската наука са динамични (променливи). Промените могат да бъдат, промени в структурата на системата (в броя на описващите я параметри) или само промени в стойностите на параметри-

те. Промените, изразяващи се в промяна само на стойностите на постоянен по брой параметри, свързваме с понятието *състояние* на системата. Казваме, че една съвкупност от стойности на всички параметри определя едно състояние на системата. Когато се променя стойността на поне един от параметрите, казваме че се променя състоянието на тази система, или че в системата протича процес. *Процес* е всяко изменение в състоянието на системата.

Математически състоянието на една система се представя с понятието за точка в n-мерно пространство - на всяко състояние на системата се съпоставя точка в n-мерно пространство, което наричаме пространство на състоянията. Множеството от точки, съответстващи на възможните състояния на системата, образуват подмножество в пространството на състоянията, което може и да не съвпада с цялото пространство. Координати на всяка точка се явяват стойностите на параметрите на системата в дадено състояние.

На преминаването на системата от едно състояние в друго съответства преминаване от една точка в пространството на състоянията към друга точка от същото, т.е. на движение. Това движение може да бъде непрекъснато или скокообразно.

## ***2.6. Характерни особености на опитните данни, получени при селскостопански изследвания***

### **2.6.1 Видове данни**

Данните получени в резултат на експериментални наблюдения на селскостопански обекти ще класифицираме в три групи: логически, качествени и количествени.

Логически данни са тези които отразяват достоверността или не достоверността на дадено твърдение относно изследвания обект. Качествен характер имат данни отразяващи състоянието на качествени признаци на изследвания обект, а количествени са данните отразяващи наблюдаваните в опита стойности на количествените параметри на този обект. В раздела посветен на системния анализ, беше отбелязано, че както количествените така и качествените параметри могат да имат числено представяне. Сега ще отбележим още, че и данните от логически



тип могат да бъдат представени числено. Може да се приеме, че истинността на дадено твърдение е параметър от качествен тип с две възможни стойности: истина и неистина, и може да се премине към числено представяне, като, например, се отбелязва неистината с 0, а истината с 1.

### **2.6.2 Цифровизиране**

За целта на компютърната обработка всички събрани в опита данни първо трябва да бъдат представени числено - да бъдат, така да се каже, цифровизирани. В числен вид данните от всеки експеримент могат да се представят във вид на матрици. Всеки ред на една матрица с данни съответства на едно, наблюдавано в опита, състояние на системата (на координатите на една точка в пространството на състоянията). Елементите от даден ред на матрица с данни съответстват на различните, измерени параметри на системата, включвайки при това като параметри и истинността на различни твърдения, относно система. В една матрица може да се обединяват или данните за различни състояния на една и съща система, или данни за съпоставими състояния на различни системи.

## ***2.7. Избор на метод за обработка на данни***

### **2.7.1 Групиране, кластерен анализ**

Когато един изследовател работи с много обекти една от обичайните дейности при изследване на тези обекти е сравняването им един с друг, откриване на това кои обекти си приличат най-силно, кои се различават, разделяне на обектите на групи със сходни свойства и т.н. (Айвазян и др., 1974) Известно е че задачата за класифициране на обекти в групи със сходни свойства може да се автоматизира чрез средствата на т.нар. кластерен анализ (Мандель, 1988). Описаният вече формализъм за описание на системите може да се приложи и при кластерен анализ.

В едно и също пространство на състоянията може да се разглеждат (изобразяват) не само различните състояния на една система, а и състоянията на множество системи с еднаква структура, тогава ако всяка от системите има само едно състояние, множеството разглеждани точки изобразяват множество от системи.

Състоянията на една система (или системите в единствени състояния) в пространството на състоянията може да се сравняват на базата на математическото понятие за разстояние (метрика) в  $n$ -мерно пространство. На базата на разстоянието системите може да се групират в групи (кластери), образувани от системи с близки свойства, изобразяващи се с точки, между които разстоянията са относително малки.

### **2.7.2 Генерална съвкупност и извадка. Статистически анализ и статистическо моделиране.**

В селскостопанските изследвания, обектът на изследване не е отделен, конкретен обект, а множество от обекти от определен клас. Например: изследва се даден сорт растения. Сортът не е едно конкретно растение, а група от много (дори огромен брой) растения, притежаващи дадено съчетание на качества и свойства. Броят растения, принадлежащи към даден сорт очевидно е толкова голям, че практически не е възможно всички да бъдат подложени на изследване. При всяко изследване се взема една малка група растения, тя се подлага на наблюдение, измерване, отчитане и т.н., получените данни се обобщават, правят се изводи и се приема, че изводите се отнасят не само за малкото на брой изследвани растения, а и за всички растения от този сорт. Очевидно тук трябва да се прави разлика между обекта, на изследването - сорта и конкретната малка група растения от този сорт, непосредствено наблюдавани в опитите.

Известно е че множеството от всички обекти, които представляват интерес в дадено изследване се нарича *генерална съвкупност*, а тази малка част от обекти, които непосредствено се наблюдават и вземат участие в изследването, се нарича *извадка*.

Всеки параметър на обект от генералната съвкупност може да има различни стойности при различните обекти. Например, масите на различните растения от даден сорт са различни. Такива величини, които при отделни наблюдения (измервания), проведени при еднакви условия показват различни стойности в математиката са известни като *случайни величини*. Случайните величини се изследват в *теорията на вероятностите*, където се описват общите свойства и

характеристики на тези величини - вид (дискретни или непрекъснати), разпределение, числови характеристики (средна стойност, дисперсия, мода, медиана и др.) и т.н.

Когато едно изследване се провежда върху извадка величините, които описват извадката, са също случайни величини (поради това, че ако се повтори опита и в извадката се вземат други обекти, може да се получат други стойности). Връзките между величините, характеризиращи извадката и величините, характеризиращи генералната съвкупност се изследват от *математическата статистика*. На базата на методите на математическата статистика, от данните получени при наблюдение върху извадка, могат да се направят най-достоверни и точни изводи за генералната съвкупност.

Основни типове задачи в математическата статистика са: определяне вида на разпределението на една едномерна или многомерна случайна величина, определяне на числовите характеристики на случайна величина (едномерна или многомерна) и задачите за проверка на статистически хипотези. Изучаването на изследвания обект чрез средствата на математическата статистика ще наричаме *статистическо моделиране*.

В действителност при всяко изследване с методите на математическата статистика се прилага понятието случайна величина, която можем да разглеждаме като математически модел на измерването на величините, извършвано в експерименталната работа. Ето защо си позволяваме да причислим към моделирането и статистическата обработка на данни.

За съжаление, много статистически процедури за обработка на данни в селскостопанската изследователска практика са станали някак си доста "обичайни", "традиционни", "общоприети" и рутинни, и много рядко някой се замисля все още какви точно са били логическите (математически) основания, за да започнат да се прилагат именно тези процедури. Това довежда до едно "механично", безкритично прилагане на станали нещо като мода методи във всички случаи, без достатъчна проверка дали наистина са подходящи за тези случаи. При това ма-

тематическите (вероятностни) модели, на които се базират статистическите методи остават на много заден план и дори се забравя за тях.

Статистически модели ще наричаме модели, в които изследваната система представлява генерална съвкупност (в смисъла на това понятие вложен в математическата статистика) и обработката на експериментални данни, получени при наблюдения върху тази система, се извършва с явно позоваване на този факт и посочване на цялата поредица от логически основания, за извършване на обработката именно по избрания във всеки от случаите начин.

### 2.7.3 Динамични системи и модели

Динамични системи наричаме системите, които променят състоянието си с течение на времето. При описанието на такива системи присъства като параметър и времето. Обработката на данни от опити, в които е наблюдавано поведението на динамични системи може да се посочи, като типичен пример за обработка на данни, която трудно може да бъде направена само чрез популярните методи на математическата статистика. Според автора обработката на данни за такива системи трябва да включва моделиране. Само след съставяне на подходящ модел на такава система и съпоставяне на експерименталните данни с данните, получени на изхода на модела може да се правят адекватни изводи за процесите в системата.

Състоянието на една динамична система в даден момент време се представя с точка в пространството на състоянията, а при изменение на състоянието на системата, точката описва *траектория* в пространството на състоянията. Същността на моделирането на такава система се състои в съставяне на *уравненията на движение*, които позволяват да се определи траекторията на системата. При динамичните системи сред параметрите, описващи състоянието има и такива, които описват бързината на промяна на състоянието ѝ във времето - параметри аналози на величината скорост от физиката. Уравненията на движение на една динамична система обикновено са уравнения, представящи зависимостта на параметрите, описващи бързината на изменение на състоянието от останалите параметри в даден момент време. Математическият апарат за представяне на урав-

ненията на движение може да е различен: уравнения с крайни разлики, обикновени диференциални уравнения или дори частни диференциални уравнения. При по-проста система, уравнението на движение може да е едно, при по-сложна - може да са няколко и да образуват система уравнения.

В някои системи съществена част от процесите може да се състоят в придвижване на вещества, индивиди или др. субстанции. При моделиране на такива системи се съставят уравнения на баланса, които описват количествено преноса на всяка субстанция.

### ***2.8. Оценка на адекватността на моделите***

Целта на моделирането е създаване на система (модел), която да наподобява наблюдаваното поведение на определена реална система, т. е. моделът трябва да предсказва възможно най-точно наблюдаваните в опита параметри на реалната система при зададени параметри на условията, при които функционира тя. Да означим набора от наблюдавани параметри на реалната системата  $y_1, y_2, \dots, y_n$  с  $Y=(y_1, y_2, \dots, y_n)$ , а наборът от параметри на условията  $x_1, x_2, \dots, x_m$  с  $X=(x_1, x_2, \dots, x_m)$ . Моделът притежава набор от свои собствени параметри  $p_1, p_2, \dots, p_k$ , които ще означим с  $P=(p_1, p_2, \dots, p_k)$ .  $X$ ,  $Y$  и  $P$  разглеждаме като точки от в пространства с размерности съответно  $m$ ,  $n$  и  $k$ . Символично модела можем да представим като функция, която на зададени  $X$  и  $P$  съпоставя някакви  $Y'=(y'_1, y'_2, \dots, y'_n)$ , от които очакваме да бъдат близки до  $Y$ :

$$(2) \quad Y'=M(P,X).$$

Най-прост случай имаме когато  $Y$  и  $Y'$  са едно компонентни  $n=1$  (скаларни) величини. Тогава разликата между тях представлява абсолютната грешка на модела:

$$(3) \quad \Delta Y = Y' - Y.$$

Модулът на отношението на абсолютната грешка към наблюдаваната стойност представлява относителната грешка на модела:

$$(4) \quad \varepsilon = \left| \frac{\Delta Y}{Y} \right|.$$

Когато  $n > 1$  може да се дефинира абсолютна и относителна грешка за всеки от параметрите  $y_i$ ,  $i=1, \dots, n$ :

$$(5) \quad \Delta y_i = y'_i - y_i \quad \text{и} \quad \varepsilon_i = \left| \frac{\Delta y_i}{y} \right|.$$

Тези грешки характеризират точността на модела по отношение на отделните моделирани величини, но за да се характеризира качеството на модела като цяло може да се използва подходяща реалнозначна функция от вида  $f(Y', Y)$ , която има свойства на разстояние в  $n$ -мерно пространство. Например:

$$(6) \quad d_E = \sqrt{\sum_{i=1}^n C_i (y_i - y'_i)^2},$$

$$(7) \quad d_M = \sum_{i=1}^n M_i |y_i - y'_i|, \quad \text{или друга.}$$

$C_i$  и  $M_i$  в тези формули са неотрицателни тегловни коефициенти, чиито смисъл е в това да придадат различна тежест на отделните параметри. На параметрите с по-големи тегловни коефициенти им се придава, така да се каже, по-голямо значение, а когато всички тегловни коефициенти са еднакви (най-често равни на 1) всички параметри стават равнопоставени. Въвеждането на тегловни коефициенти има и този смисъл, че чрез тях може да се приведат към еднаква размерност всички параметри и да се избегне ефекта от това, върху оценката за точността да оказват най-голямо влияние само параметрите, чиито стойности се изразяват просто с по-големи числа. Ползването на тегловни коефициенти е необходимо и в случай, че параметрите на системата са измервани опитно с различна точност - чрез тегловни коефициенти на по-точните данни може да се отдаде по-голямо значение.

Оценката на точността на модела по представените формули (3) - (7) се отнася за определен набор от параметри на околната среда  $X$ , установени при един експеримент и зададен набор вътрешни параметри на модела  $P$ . Пълна оценка за точността на модела се получава при сравняване на резултата, който той дава, с наблюдаваното поведение на реалната система не само в един, а в множество

опити. Нека множеството от комбинации на параметри на средата  $X_1, X_2, \dots$ , наблюдавани при различни опити, да означим с  $\mathfrak{X}=(X_1, X_2, \dots)$ .  $\mathfrak{X}$  може да се разглежда като пространство от случайни събития, а грешката на модела (някоя от стойностите (3) - (7)), като случайна величина в това пространство. Чрез известни от статистиката формули може да се намерят оценки за числените характеристики на грешката: средна стойност, дисперсия и т. н. Колкото средната стойност на грешката на модела е по-близка до нула, толкова той е по-точен, а колкото е по-малка дисперсията  $\sigma$  - толкова модела е по-надежден. Ако средната стойност на абсолютната грешка на модела е различна от нула, това означава, че той дава систематично отклонение от истинските стойности на моделираните величини: надвишава ги, когато средната грешка е положителна или предсказва по-малки стойности, когато е отрицателна.

Друга оценка за точността на един модел по отношение на една от моделираните величини  $y_i$  представлява и отношението на сумата от квадратите на абсолютните грешки на модела към сумата от квадратите на отклонението на моделираната величина  $y_i$  от средната  $\bar{y}$  стойност  $\bar{y}$ :

$$\frac{\sum_{\mathfrak{X}} (y'_i - y_i)^2}{\sum_{\mathfrak{X}} (y_i - \bar{y})^2}.$$

Сумирането и осредняването става по множеството на наблюденията  $\mathfrak{X}$ , извършени над реалната система, за които е приложен модела.

Близка до нула стойност на това отношение означава, че съставеният модел отчита добре факторите от които зависи моделираната величина и предсказаната от модела стойност е по-близка до истинската стойност в сравнение със средната стойност на моделираната величина. Това отношение, при по-свободно тълкуване, може да се приеме, че показва каква част от варирането на моделираната величина се дължи на неотчетени от модела фактори. От друга страна пък, остатъкът до единица:

$$(8) \quad R^2 = 1 - \frac{\sum_{\mathfrak{R}} (y'_i - y_i)^2}{\sum_{\mathfrak{R}} (y_i - \bar{y})^2}$$

може да се приеме, че показва до каква част всички фактори оказващи влияние върху измененията на моделираната величина са отчетени в модела. Ако разглежданият модел е статистически линеен регресионен модел, (8) съвпада с квадрата на коефициента на корелация.

### ***2.9. Параметризиране и подобряване на моделите***

Както беше посочено всеки модел се характеризира с едно множество от собствени, вътрешни параметри:  $P=(p_1, p_2, \dots, p_k)$ . Намирането на подходящи стойности за тези параметри наричаме параметризиране на модела. Много от параметрите имат съвсем конкретен реален смисъл и представляват величини достъпни за пряко или косвено измерване. Параметризирането в този случай се свежда до провеждане на подходящи измервания или използване на резултатите от вече проведени такива, публикувани или не в литературата. Възможно е обаче в модела да се наложи използване и на параметри, които в момента не могат да бъдат измерени и даже няма информация да са били измервани. Това налага търсенето на подходящи стойности за тях да се прави чрез методи за оптимизиране. При оптимизация първо се дефинира подходящ критерий за оптимизация.

Критерият за оптимизация представлява функция на параметрите на модела, която има добре изразен абсолютен екстремум. Параметризирането на модела се свежда до намиране на стойностите на параметрите отговарящи на този екстремум.

При избора на критерий за оптимизация първо трябва да се вземе предвид по отношение на какво качество моделът трябва да бъде подобрен. Съставя се математическа функция, която отразява количествено това качество и има максимум или минимум при онези стойности на параметрите на модела, при които посоченото качество е най-силно изразено. Например, ако трябва да се повиши точността на модела като критерий за оптимизация може да се използва някоя от



описаните оценки за точност (3) - (8), които имат изразен минимум, когато моделът е най-точен.

Към критерия за оптимизация има изискване да е сравнително прост и пресмятането му да изисква малко машинно време, а от друга страна да има такива качества, че да съществува достатъчно надежден и икономичен метод за намиране на глобалния му екстремум. Съобразяването с такива изисквания налага да се избират като критерии за оптимизация функции, които нямат нагледен и лесен за интерпретиране обективен смисъл.

### ***2.10. Изводи:***

1. От направения анализ се вижда, че съществува връзка между обработката на опитни данни и моделирането, затова се налага проверка на изложените идеи - да се приложи моделирането при обработка на данни получени при различни селскостопански експерименти.

2. В зависимост от целта, преследвана с провеждането на даден селскостопански експеримент и характера на получените данни в различните случаи може да се наложи използването на различни видове математически модели и методи - емпирични модели, съставяни по метода на най-малките квадрати, статистически модели, кластерен анализ, моделиране на динамични системи и др. Това налага съставяне на широк набор от програмни модули, реализиращи съответните алгоритми.

## Глава 3

### **Разработка на специализирани софтуерни продукти за обработка на данни от лабораторни и полски изследвания и за съставяне на модели в изследванията и практиката**

Всяка програма, предназначена за обработка на данни може да се раздели на модули, предназначени за: 1. компютърно представяне на данните, 2. въвеждане и извеждане на данните, 3. обработка и преобразуване на данните.

Необходимата за решаване на дадена задача програма може да бъде организирана по различен начин в зависимост от това какви са перспективите за по-нататъшно използване на тази програма и езика за програмиране, на който се пише. Стремешът при написване на програмите, представени тук, беше да се търси винаги най-икономично решение на всяка конкретна задача. Това означава, че е обръщано много по-малко внимание на такива черти на програмите като удобство за ползване, универсалност и др. и е акцентирано единствено на това програмата да реши конкретната задача, като представи по подходящ начин данните и направи изискваното им преобразуване. Входът и изходът винаги е оставал на много по-заден план и се осъществява според случая. При всяка нова задача е променян и реорганизиран вече написаният преди това за друга задача код, така че с негова помощ да се реши и новата задача. Така постепенно се оформяха отделни модули, които след всяка нова задача се обединяваха във все по-добре организирани и решаващи повече задачи програми.

Езикът за програмиране е обектно ориентиран Pascal, а за създаване на програмите са използвани развойни среди като Turbo Pascal 4.0, Turbo Pascal for Windows и Delphi 4. Само в няколко от най-ранните научни публикации авторът е използвал Basic (Ангелова, Желепов, Георгиев 1988). Тук се описва само програмният код, написан с Delphi 4. По-старите програми, направени с Turbo Pascal 4.0 и Turbo Pascal for Windows, няма да се разглеждат, защото са предназначени за остарелите вече операционни системи MS DOS и Windows 3x, но може да се видят на компакт диска приложение, който съдържа и други компютърни файлове, създадени във връзка с настоящия труд. На диска се намира програмата

M1.exe, с първичния код на всички модули, от които е изградена. Тя представлява MS DOS програма за статистическа обработка на данни и моделиране. Може да се компилира с Turbo Pascal 4.0 или по-късна версия до 7.0.

Стилът на програмиране е обектно ориентиран, следващ, когато е възможно, парадигмата "Модел - интерфейс - управление" (MVC - Model View Control). Пълно описание на целия написан код би било много обемисто и едва ли е оправдано, затова тук се описва само общият принцип на организацията и само тези модули, които имат най-непосредствена връзка с обработката на данни и съставяне на модели. Много рутинни елементи от програмирането няма да бъдат засегнати въпреки, че без тези елементи съставените програми едва ли биха имали необходимите за удобно и пълноценно използване качества.

### ***3.1. Модул за представяне и действия с матрици (MatrUnit.pas)***

Както беше изтъкнато в предишната глава, за да бъдат обработени на компютър, опитните данни се цифровизират и представят във вид на матрици. Ето защо важна част от програмния код, който ни е необходим, е тази част, която служи за представяне и извършване на действия с матрици. През различните периоди от своята работа автора е използвал различни начини за реализиране на този код, но тук е представена само реализацията използвана в най-новите работи. Това е един модул (unit), на който е дадено името MatrUnit.pas. (Вижте приложения компакт диск.)

Реалните числа, представляващи елементи на матрици, се представят в компютъра с числа с плаваща запетая, като за различни цели може да се използва различна точност. В Delphi 4 числата с плаваща запетая могат да бъдат от тип: single, real, double или Extended. За да може за различни цели да се използва всеки от тези типове се обявява един базов тип: TNumValue, който се използва навсякъде по-нататък, където трябва да се обявяват променливи, представлящи числа с плаваща запетая. Според необходимостта този тип се прави идентичен с някой от типовете single, real, double или Extended. Като начало е обявен за идентичен с типа double. Освен това се обявява и един тип указател PNumValue, към променливи от тип TNumValue.

За представяне на динамичен масив от стойности от тип TNumValue се обявява тип TVector, а за представяне на елементите на матрица се обявява тип TVecArray, който е динамичен масив от променливи от тип TVector.

Освен елементите си една матрица има и редица други характеристики: например брой редове и стълбове, а и върху матрицата следва да може да се извършват различни операции. За обединяване на кода представящ, както елементите, така и характеристиките, и някои от операциите с матрици използваме тип TMatrix, който представлява обект.

**Свойствата** (properties) на обекта TMatrix са:

Modified:boolean - помощен параметър, който става истина след промяна или е неистина след създаване на матрицата или след записването ѝ във файл;

Origin:TOrigin; - помощен параметър от изброим тип TOrigin, който показва какъв е произхода на матрицата, т.е дали е въведена ръчно, дали е прочетена от файл или е съставена в резултат от някой от анализите на данни;

m:integer - брой на редовете;

n:integer - брой на стълбовете;

X[i,j:integer]:TNumValue - елементът от i-ти ред, j-ти стълб;

Pt[i,j:integer]:PNumValue – указател към елемента от i-ти ред, j-ти стълб;

Col[i:integer]:TVector - i-ти стълб на матрицата;

Основните операции (**методи**) са:

constructor Create(m1,n1:integer); - създава структура от тип TMatrix;

destructor Destroy; override; - освобождава паметта заемана от матрицата;

procedure ReDim(m1,n1:integer); - променя дименсията на матрицата представена от дадения обект;

procedure InsertCol(i:integer; after:boolean); - вмъква стълб на мястото (когато after е false) или след (когато after е true) i-тия стълб;

procedure InsertRow(i:integer; after:boolean); - аналогично вмъква ред;

procedure DeleteCol(i:integer); - изтрива стълб с номер i;

procedure DeleteRow(i:integer); - изтрива ред с номер i

procedure SetLine(i,k:integer; s:string); - променя стойностите на елементите от ред i, започвайки от k-тия, с редица стойности, зададени в променливата s.

procedure LoadFromFile(const s:string); - чете матрица от файл;

procedure SaveToFile(const s:string); - записва матрица във файл;

function NoIndex(k:TIntArray):boolean; - връща истина, ако в масива от цели числа k има поне едно, което не може да е номер на стълб на матрицата;

function SubMatr(k:TIntArray):TMatrix; - връща нова матрица, която се състои от стълбовете с номера, посочени в масива от цели числа k;

procedure Mult(a:TNumValue); - умножава матрицата с числото a;

function Distance(p,q:integer):TNumValue; - пресмята евклидовото разстояние между редовете с номера p и q;

function Gaus:boolean; - привежда матрицата в диагонален вид, по метода на Гаус с избор на главен елемент (Сендов, В., В. Попов, I част 1976);

procedure GausBack; - извършва обратният ход на метода на Гаус;

procedure Swap(i1,i2:integer); - разменя стълбовете с номера i1 и i2;

procedure Sort(i1,i2,i:integer); - сортира редовете на матрицата с номера от i1 до i2, по нарастване на елементите от стълб с номер i.

Освен обектния тип TMatrix в MatrUnit.pas се дефинират и някои **функции**, които приемат като входни параметри или връщат стойности от тип TMatrix. Такива са функциите:

function IntArray(const s:string):TIntArray; - преобразува, представените в текстов вид цели числа в динамичен масив от цели числа;

function Vector(const s:string):TVector; - преобразува, представените в текстов вид реални числа в динамичен масив от числа с плаваща запетая;

function AsText(V:TVector):string; - обратно, преобразува масив от числа с плаваща запетая във вид на текст;

procedure SetAs(X,Y:TVector); - присвоява на елементите на X съответните стойности на елементите на Y;

procedure DeleteValue(i:integer; var X:TVector); - изтрива i-ия елемент на масива X и намалява с единица неговата размерност;

function Matr1(n:integer):TMatrix; - връща нова единична матрица от ред n;  
function MatrCol(V:TVector):TMatrix; - от масив от реални числа, обявен чрез променлива от тип TVector връща матрица стълб.

function CopyMatr(Mt:TMatrix):TMatrix; - създава копие на матрицата Mt;  
function Transpose(Mt:TMatrix):TMatrix; - създава нова матрица, транспонирана на Mt;

function Inverse(Mt:TMatrix):TMatrix; - връща обратната матрица на Mt, ако тази матрица е квадратна и не изродена или в противен случай връща nil;

function JoinMatr(A,B:TMatrix):TMatrix; - от матриците A и B се съставя нова матрица, която включва стълбовете на A и B;

function Mult(A,B:TMatrix):TMatrix; - връща нова матрица, представляваща произведение на матриците A и B.

### **3.2. Форма за визуализиране на матрици (MatrixForm.pas)**

Този модул представя един интерфейс (обект от тип TMatrixForm1) за показване и работа с матрици. Интерфейсът се състои от прозорец, в който се показва визуален елемент във вид на таблица и меню.

Визуалният елемент е специално програмиран за целта и е наследник на типа TStringGrid от библиотеката с визуални елементи на Delphi 4. Обявен му е тип TMatrixGrid. В допълнение към свойствата на TStringGrid са му програмирани и свойства, които го приспособяват за показване на матрици:

Matrix:TMatrix - структура съдържаща представяната матрица,  
Също така и процедури за актуализиране и модифициране на матрицата:  
CopyToClipboard - копиране на маркираната част от матрицата в Clipboard,  
PasteFromClipboard - извличане на матрица от Clipboard,  
ReShow; - опреснява показването на матрицата. Трябва да се извиква след действия, предизвикващи промени в матрицата.

Формата TMatrixForm1 съдържа още списък с действия (TActionList), меню и контекстно меню. Списъкът действия съдържа описание на методи, които могат да бъдат предизвикани от потребителя посредством посочване на точките от ме-

ното или от контекстното меню и служат за извикване на различните процедури и функции за действия върху матрицата.

### **3.3. Модул за пресмятане на статистики (*StatUnit.pas*)**

Статистическите редове представляващи стойностите на дадена величина, получени при  $n$  наблюдения върху извадка се представят чрез променливи от тип `TVector`. Програмният код, който представя функциите за изчисляване на статистики върху тези статистически редове са обособени в модул, означен с името `StatUnit.pas`, в който са реализирани следните функции:

`function Sum(X:TVector):TNumValue;` - връща сумата на стойностите от статистическия ред;

`function Sum2(X:TVector):TNumValue;` - връща сумата от квадратите на стойностите;

`function Modul(X:TVector):TNumValue;` - пресмята евклидовата дължина на вектора  $X$ ;

`function Average(X:TVector):TNumValue;` - изчислява средната стойност;

`function Variance(X:TVector):TNumValue;` - неизместима оценка на дисперсията;

`function VarianceP(X:TVector):TNumValue;` - изместима оценка на дисперсията;

`function StDev(X:TVector):TNumValue;` - неизместима оценка на стандартно отклонение;

`function StDevP(X:TVector):TNumValue;` - изместима оценка на стандартно отклонение;

`function Covar(X,Y:TVector):TNumValue;` - коефициент на ковариация;

`function Correl(X,Y:TVector):TNumValue;` - коефициент на корелация;

`function CorrelMatr(Mt:TMatrix):TMatrix;` - връща корелационна матрица от коефициенти на корелация между стълбовете на  $Mt$ ;

`procedure Wilcoxon(X1,X2:TVector; var U2,U1:integer);` - процедура, използвана при  $U$ -критерия на Уилкоксон, пресмята броя на инверсиите  $U2$  на стойности-

те на X2 спрямо стойностите на X1 и броя на инверсиите U1 на стойностите на X1 спрямо стойностите на X2;

procedure LinReg(Mt:TMatrix; var Co,Er,Re:TMatrix); - процедура, която извършва линейна регресия върху данните от матрицата Mt. Последният стълб на матрицата се приема за зависима променлива, а останалите стълбове – за независими. Коефициентите на регресия се записват в матрицата стълб Co, а техните грешки в матрицата стълб Er. В матрицата Re в три стълба се записват: дадените стойности на зависимата променлива, изчислените по регресионния модел и разликите между едните и другите.

procedure Regression(const s:string; T:TRegType; Mt:TMatrix; var Co,Er,Re:TMatrix); - Това е друга процедура за извършване на регресионен анализ. s:string трябва да съдържа номера на стълбове, представляващи съответно зависимата променлива и една или няколко независими променливи. T е променлива от избран тип TRegType, показваща вида на регресията: линейна, логаритмична, полиномиална, степенна или експоненциална. Опитните данни трябва да се съдържат в матрицата Mt. Променливите трябва да са разположени в стълбове, а наблюденията в редове. Резултатът от регресията се записва в матриците Co, Er и Re, както при процедурата LinReg.

function RegRaport(T:TRegType; Co,Er,Re:TMatrix):string; - показва в текстов вид резултата от регресионния анализ, получен чрез процедурата Regression.

procedure Standardize(X:TVector); и procedure Standardize(M:TMatrix); - стандартизират елементите на вектора X или на матрицата M, като извършва трансформацията:

$$(9) \quad x_i \rightarrow \frac{x_i - \bar{x}}{s},$$

където  $x_i$  е  $i$ -тия елемент на вектора или стълба на матрицата,  $\bar{x}$  - средната стойност, а  $s$  - средноквадратичното отклонение на елементите от този вектор или стълб.



### **3.4. Модул за оптимизация**

В работата си авторът е използвал два метода за намиране на екстремум на функция на много променливи: метод на случайното търсене с обратна стъпка и метод на деформируемия симплекс на Недлер - Мид (Стоянов, С, 1990). Тези методи бяха реализирани програмно в няколко модула, написани с Turbo Pascal а по-късно и с Delphi. Модулът, направен с Delphi OptimUnit.pas използва някои от дефинираните в MatrUnit.pas и StatUnit.pas типове, процедури и функции. Централно място в този модул заема функцията: `function MaxOfFunction(F:TFunctionOfVector; X:TVector):TNumValue;` която приема като входни параметри, F:TFunctionOfVector – функцията, чиито максимум ще бъде определен и X:TVector – началната точка от която започва търсенето на максимума. Функцията, чиито максимум се търси е от тип: `TFunctionOfVector=function(V:TVector):TNumValue;`. Търсенето на максимума се извършва по метода на деформируемия симплекс на Недлер – Мид.

### **3.5. Модул за програмно представяне на математически изрази (MathExprUnit.pas)**

Във връзка с честото възникване на необходимост при пресмятанията да се прилагат въвеждани от потребителя математически изрази беше съставен модулът MathExprUnit.pas, който позволява математическите изрази да се представят програмно като обекти. В този модул е обявен типът: **TMathExpr=class(TObject)**, който именно се използва за програмно представяне на математическите изрази. За характеризиране вида на математическия израз се използва изброим тип: TExprType, различните допустими стойности от който съответстват на:

oprSum – математически израз представляващ сума от други изрази,

oprMult – произведение от други изрази,

oprDiv – частно на два изрази,

oprPower – степен образувана от един израз – основа и друг израз – степенен показател,

oprEqu - равенство,

oprChSign – израз представляващ противоположен по знак израз от друг израз,

oprSin – функция  $\sin(x)$ ,

oprCos – функция  $\cos(x)$ ,

oprAtan – функция  $\arcsin(x)$ ,

oprExp – експоненциална функция  $\exp(x)$ ,

oprLn – логаритмична функция  $\ln(x)$ ,

oprConst – израз представляващ константа,

oprVar – израз представляващ променлива,

oprNoDeff – недефиниран израз.

Скритите (private) полета на един обект, представящ математически израз са:

OT:TExpType; - показва вида на математическия израз;

VI:TNumValue; - съдържа стойността на израза, ако той представя константа (oprConst);

PVI:PNumValue; - съдържа указател към числена стойност от тип TNumValue, когато изразът е променлива (operVar);

Nm:String; - съдържа името на променливата, когато изразът е променлива (operVar);

EL:TList; - структура съдържаща други математически изрази, съставляващи дадения израз.

Създаването на обекти от тип TMathExpr, представящи математически изрази може да се осъществи чрез един от конструкторите:

constructor Create(V:TNumValue); - създава израз представляващ константа със стойност V;

constructor Create(const N:string); - създава променлива с име N;

constructor Create(T:TExpType; E:TMathExpr); - създава израз от вида T, който съдържа като свой подизраз израза E;

constructor Create(T:TExpType; E1,E2:TMathExpr); - създава израз от вида T, съставен от два израза E1 и E2.

Други специфични методи на обектите от тип TMathExpr са:

`procedure Add(E:TMathExpr);` - ако изразът представлява сума или произведение тази процедура добавя още един израз в края на сбора или произведението;

`procedure Insert(E:TMathExpr);` - ако изразът представлява сума или произведение тази процедура вмъква още един израз в началото на сбора или произведението;

`function Value:TNumValue;` - ако изразът може да бъде пресметнат тази функция връща стойността на израза;

`function HasValue:boolean;` - логическа функция, която връща истина, ако изразът може да бъде пресметнат или неистина – ако не може;

`function Text:string;` - връща стринг, който представя математическия израз във вид на текст;

`function Priority:integer;` - приоритет на израза (0 – ако изразът е константа или променлива, 1 – ако е сума; 2 – ако е произведение; 3 – ако е степен; 4 – ако е друга функция) използва се за определяне реда на изчисленията и за поставяне на скоби при представяне на израза във вид на текст;

`function E(i:integer):TMathExpr;` - когато изразът е съставен от други изрази тази функция връща *i*-ия подизраз.

За да се осигури някакво съгласуване на изразите, използвани за определена цел, (например, всички променливи с едно и също име да се представят от един обект) е съставен друг тип обекти: **TMathModel=class(TObject)**. Предназначението на този тип е да обединява в някакво цяло (математически модел) математически изрази, представяни с обекти от тип TMathExpr.

Може би най-важния метод в този тип е: `procedure AddExpr(const s0:string);`. Тази процедура получава като параметър низ, представящ в текстов вид математически израз. (При представянето на изразите в текстов вид аритметичните действия следва да са означени със знаците + - \* /, повдигането на степен със знак ^, а функциите – с общоприетите им означения.) Преди започване конструирането на обекти от тип TMathExpr входния низ чрез една помощна функция `function StandardStr(const s:string):string;` се конвертира в стандартен вид, при което се премахват интервалите, а всички букви се обръщат в малки. Конструира-

нето на обект, съответстващ на даден низ се осъществява с една рекурсивна вътрешна функция `function MathExpr(const s:string):TMathExpr;`. Функцията `MathExpr` създава променливи посредством друга вътрешна функция `function AddVariable(const s:string):TMathExpr;`, която проверява, дали вече е създадена променлива с дадено име. Ако такава е създадена връща обекта представящ тази променлива, а ако не е създадена я създава и връща новосъздадения обект. Така се осигурява, всички променливи с еднакви имена да се представят с един и същи обект.

Променливите в рамките на един обект от тип `TMathModel` могат се разделят на два вида, които ще наричаме съответно: същински променливи и параметри. Същински са променливите носещи имена от вида 'x1', 'x2', 'x3', ..., т.е започват с буквата 'x' и съдържат цяло число след тази буква. Параметри са всички останали променливи, чиито имена не са съставени по този начин. Същинските променливи се използват за представяне на опитни данни. Полето `Data:TMatrix;` служи за представяне, когато е необходимо, на матрица с опитни данни. Методът `procedure SetVarPointers(i:integer);` насочва вътрешните указатели `PV1` на същинските променливи към елементи от *i*-ия ред на матрицата `Data` (например ако има променлива 'x3', нейния `PV1` указател се насочва към `Data.X[i,3]`). Стойности на параметрите се задават с помощта на метода `procedure SetParams(V:TVector);` и се съхраняват в полето `Param:TVector;`

Идеята е обектите от тип `TMathModel` да се използват за представяне на математически модели, въвеждани от потребителя по време на изпълнение на дадена програма и моделиране с тях на опитни данни. Например, ако са налични данни във вид на матрица, в първия стълб на която има стойности на независима променлива, а във втория стълб съответни стойностите на зависима променлива и трябва да се моделира зависимостта между тях с квадратична функция, потребителят може да въведе формулата: 'a + b \* x1 + c \* x1^2'. Променливата с име 'x1' в тази формула е същинска променлива и нейните стойности ще са числа от първия стълб на матрицата `Data`. Променливите с имена: 'a', 'b' и 'c' са параметри техните, въведени от потребителя начални стойности се задават с метода

SetParams. Методът `function ModelError(c:integer):TNumValue;` в обектите от тип `TMathModel` пресмята сумата от квадратите на разликите между изчислените по зададена формула стойности и стойностите в стълб `c` на матрицата `Data`. Чрез съставяне на една помощна функция и използване на функцията `MaxOfFunction` от модул `OptimUnit.pas` може да се намерят стойности на параметрите, които минимизират грешката на модела.

Така представените идеи са реализирани в програмата `AnalysisTools`.

### **3.6. Модул за кластер анализ (*ClustAnal.pas*)**

В този модул е обединен програмният код, предназначен за осъществяване на кластер анализ. Ядрото на модула е:

`function Clusters(Mt:TMatrix; k:integer):TMatrix;` - една функция, която извършва кластер анализ над данните от матрицата `Mt` и връща като резултат нова матрица. Приема се, че редовете във входната матрица `Mt` са координати на точки в `Mt.n` - мерно пространство, които постъпково се обединяват до образуване на един кластер, съдържащ всички точки. Преди преминаване към постъпково обединяване в кластери входната матрица се подлага на стандартизация чрез трансформация от вида ( 9 ), изпълнена върху всеки стълб на матрицата.

Разстоянието между две точки в `n`-мерно пространство се изчислява по формулата за евклидово разстояние:

$$(10) \quad l = \sqrt{\sum_{i=1}^n (x_i - y_i)^2},$$

където  $x_i$  е  $i$ -тата координата на едната, а  $y_i$  -  $i$ -тата координата на втората точка.

Когато кластерите съдържат повече от една точка, за разстояние между кластери може да се приеме:

1. разстоянието между центровете на тежестта на точките от двата кластера;
2. минималното разстояние между точка от единия и точка от другия кластер

или

3. максималното разстояние между точка от единия и точка от другия кластер. Коя от тези три дефиниции се използва се определя от входния параметър `k:integer` на функцията `Clusters`.

На всяка стъпка се обединяват в един кластер, кластерите между които разстоянието е най-малко.

Матрицата, която връща функцията като резултат от постъпковото обединяване има размерност  $(Mt.m-1 \times 4)$ . Всеки ред на тази матрица представя една стъпка на обединяване. Първият елемент от всеки ред е номера на стъпката, вторият и третият - номерата на кластерите, които се обединяват на тази стъпка, а четвъртият - разстоянието между кластерите, които се обединяват.

Резултатът получен с функцията `Clusters` може да се визуализира и разгледа по различни начини. Може да се види във вида, в който е получен или да се преобразува в текстов или графичен вид.

`function ClustRaport(Mt:TMatrix; k:integer):string;` - представя резултата от кластер анализа в текстов вид. За входен параметър в тази функция трябва да се използва матрица получена с функцията `Clusters`. `k:integer` определя коя от трите дефиниции за разстояние между кластери с повече от една точка е използвана.

`function Dendrogram(Mt:TMatrix; ls,Pos:integer; Lb:TStringList):TDendrogram;` - представя резултата от кластер анализа във вид на дендрограма. Дендрограмата е обект от тип `TDendrogram`, описан пак в този модул. `TDendrogram` е наследник на базовият тип за графични изображения `TGraphObject`, който е обявен в описания по-надолу модул `GraphUnit`. Входните параметри на функцията `Dendrogram` са: `Mt:TMatrix;` - матрица върната от функцията `Clusters`, съдържаща резултата от кластер анализа; `ls:integer;` - определя номера на стъпката на която спира обединяването в кластери; `Pos:integer;` - определя положението на дендрограмата (вертикално при `Pos=0` или хоризонтално при `Pos=1`); `Lb:TStringList` - е списък от описателни наименования, които може да се използват вместо номерата на кластерите. (Ако `Lb=nil` или `Lb.Count <> Mt.m`, се използват номерата.)

### 3.7. Модул за графични изображения (*GraphUnit.pas*)

Това е един помощен модул за представяне на графични изображения: графики на функции, дендрограми и т.н.

В модула се дефинира един абстрактен тип `TGraphObject`, от който следва да произхождат типове наследници за представяне на различни по вид графични изображения. Този тип съдържа един абстрактен метод:

`procedure Plot(C:TCanvas; R:TRect); virtual; abstract;` който трябва да се преопредели в типовете наследници, за да се извърши самото изчертаване на съответното графично изображение на екрана или на принтер.

В същия модул е определен типа `TChart`, за изобразяване графиката на представена таблично, зависимост на две или повече променливи. Всяко изображение, описано с типа `TChart` се състои от една или повече серии точки.

За представяне на една серия точки се използва типа `TGraphSerie`, който обединява два масива `X` и `Y`, състоящи се от координатите на точките.

Най-съществените методи на `TChart` са:

`procedure Add(s:TGraphSeries);` - прибавя серия точки, която да бъде изобразена, заедно с другите серии.

`procedure Plot(C:TCanvas; R:TRect); override;` - построява изображението на екрана или на принтер.

За построяване на променлива от тип `TChart`, която да послужи за изобразяване на зависимости между данни, представени с матрица, е написана функцията:

`function MatrixChart(s:string; Mt:TMatrix):TChart;` Входната променлива `s:string` трябва да съдържа номера на стълба на матрицата `Mt`, който играе роля на независима променлива, последван от номерата на стълбовете, играещи роля на зависими променливи. Символите отделящи споменатите номера могат да бъдат произволни нецифрови символи.





## Глава 4

### Експериментални изследвания - съставени модели

В тази глава се описват конкретни модели, които илюстрират изказаната идея за приложение на моделиране при обработката на данни от селскостопански изследвания и от изследвания на храни. Всички тези модели са намерили вече място в различни публикации, посветени на конкретни изследователски проблеми, в които моделите и моделирането представляват само една част от работата, макар и доста съществена. В изложението на тези публикации много от въпросите свързани с моделирането и най-вече как е осъществено, са разгледани много бегло. В повечето случаи там само се споменават, без да се излагат достатъчно подробно, затова тук се разкрива по-подробно и тази страна от работата, която не е намерила място в други публикации.

#### *4.1. Моделиране толерантността към вируси на сортове и селекционни линии краставици и фасул*

Намирането на подходящ формален начин за сравняване толерантността на сортове и линии растения към вируси има важно значение за работата на селекционерите. Проблем има с това, че заразяването с вируси се проявява в появата на множество различни симптоми. Реакцията на даден сорт към заразяване с вирус се отразява с двойката противоположни понятия: чувствителност и толерантност. Един сорт се счита за чувствителен, когато проявява симптоми, а за толерантен, когато не показва такива. Но това са качествени понятия, за които липсва точно количествено изражение, ето защо беше направен опит за придаване на количествено изражение на тези понятия, съставяне на модел и обработка на опитни данни на базата на такъв модел. Това е пример за задача от вида, описан в раздел “Системен анализ” от глава II.

Съставеният модел беше използван за анализ на данни от изследвания върху краставици по отношение на Краставичен мозаечен вирус (CMV) (Иванова, Георгиев и Костова, 2001) и на фасул по отношение на кукумовирус изолат Ch 39-BG (Ivanova, Kostova, Georgiev, 2002). И при двете изследвания от всеки сорт и линия са отглеждани по две групи от по 10 растения. Едната група е оставяна не

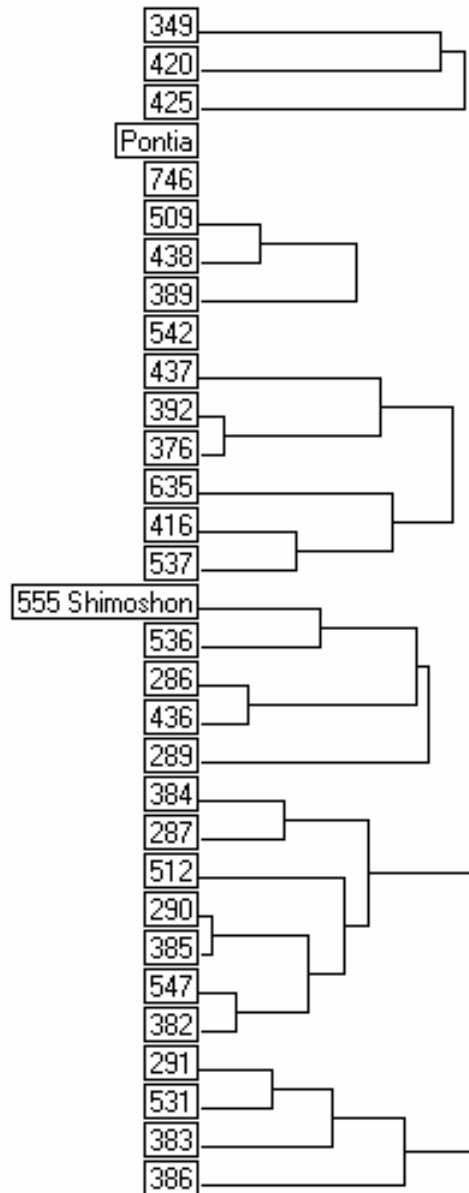
заразена, а втората група е заразена. Един месец след заразяването са извършени отчитания на три параметъра: височина, маса и брой листа на всяко от растенията. Получените данни са осреднени за 10-те незаразени и за 10-те заразени растения от всяка линия.

Прилагайки понятията на системния анализ в това изследване приемаме, че изследваната система се състои от няколко селекционни линии. Всяка от селекционните линии се наблюдава в две състояния: заразено и незаразено. Заразяването с вирус води до изменение на всеки от отчитаните параметрите и задачата за установяване на толерантността към вируса, се свежда до намиране на величина представляваща комплексна оценка на различието между тези две състояния, отразено чрез наличните данни.

При конструиране на такава величина се разсъждаваше така: Най-директно сравнение на две стойности  $x$  и  $y$  се получава чрез пресмятане на разликата между тях  $y-x$ . Но тази разлика има размерност, нейната стойност зависи от мерните единици, използвани за изразяване на  $x$  и  $y$ . Това прави безсмислено съпоставянето на изменението на една величина със изменението на друга величина, когато тези величини се изразяват с различни мерни единици. За да се извърши сравнение е по-подходящо разликата на две величини да се отчита относително чрез частното  $e=(y-x)/x$ . Тази относителна разлика е безразмерна и показва каква част от  $x$  е разликата между  $y$  и  $x$ . Относителната разлика може да се изрази в проценти. Ако се направи сравнение между стойността на всеки от параметрите, при заразена линия със стойността на същия параметър при не заразена, се получава една тримерна (векторна) величина на относителната разлика между заразените и не заразените линии:  $E(e_1, e_2, e_3)$ . Тази векторна величина моделира понятието чувствителност/толерантност на съответната линия към вируса. За сравнение на линиите може да се използва модула на този вектор, а чрез кластер анализ по този вектор, линиите може да се групират в групи (кластери) от линии с близка чувствителност. Бяха извършени и двата вида сравнение.

Пресмятането на относителните разлики на параметрите беше извършено с помощта на електронна таблица на MS Excel, а кластерния анализ с написаната

от автора програма Analysis Tools. На Фигура 2 е показана дендрограмата на групирането на линиите (сортовете) краставици според толерантността им спрямо CMV (Иванова, Георгиев, Костова, 2001). Обединяването в кластери е извършено на 24 стъпки, като на всяка стъпка се обединяват кластерите с най-малко евклидово разстояние между центровете на тежестта им.



Фигура 2. Дендрограма на групирането на линиите (сортовете) краставици според толерантността им спрямо CMV.

#### ***4.2. Групиране на сортове и селекционни линии чрез кластер анализ***

След като беше написана компютърна програма Analyses tools, същата беше използвана за групиране на други обекти: селекционни линии и сортове пипер (Panayotov, Gueorguiev, Ivanova, 1999; Panayotov, Gueorguiev, Ivanova, 2000), сортове картофи (Георгиев и др., 2001), селекционни линии краставици, сортове домати (Ivanova, Krasteva, Georgiev, 2002), диплоидни образци червена детелина (Горанова и др. 2004)

#### ***4.3. Модел на зависимостта концентрация-ефект на инсектициди***

По данни от лабораторни опити за изпитване действието на инсектициди в доклада “Един модел на зависимостта “концентрация-ефект” за някои инсектициди спрямо ларви на колорадски бръмбар” (Василева, Матеева, Ковачева, Георгиев, 1989) са съставени линейни емпирични модели на зависимостта концентрация-ефект. Изчисленията са проведени чрез съставена от автора компютърна програма.

#### ***4.4. Модел на замърсяването на почвите с тежки метали в района на КЦМ***

Във връзка с установяване степента на замърсяване на почвата с тежки метали в района на КЦМ – Пловдив възниква необходимост от картографиране на съдържанието на тези замърсители. (Лечов и Георгиев, 1992) За пресмятане концентрацията на даден замърсител в произволна точка от района беше използван емпиричен модел, съставен по данни от измерване на концентрацията в мрежа от точки. Бяха използвани данни, получени при измерване съдържанието на подвижни форми на оловото в почвени проби (Сенгалевич, 1990). Пробите са взети от радиално разположени в 8 посоки от завода точки, на разстояния: 100, 200, 300, 400, 900 и 1900 м. По предложение на Г. Лечов зависимостта на концентрацията  $C$  на олово в почвата от разстоянието до завода  $R$ , по всеки от лъчите, беше моделирана с два вида функции:

$$C = A \cdot R^{-N} \text{ - степенна и}$$

$$(11) \quad C = B \cdot e^{-k \cdot R} - \text{експоненциална.}$$

За пресмятане параметрите на степенните и експоненциални функции беше използвана, написаната от автора програмата М1 (Глава III). С по-висок коефициент на корелация с данните се оказа експоненциалната зависимост (Таблица 1).

Таблица 1

Изчислени коефициенти за степенен и експоненциален модел на зависимостта на концентрацията на подвижни форми олово в почвата от разстоянието от КЦМ по осемте посоки на вземане на пробите.

I	Степенен модел			Експоненциален модел.		
	A	N	Коеф. на корелация	B	$k \cdot 10^4$	Коеф. на корелация
1.	261,58	-0,897	0,952	442,3	-5,34	0,884
2.	179,55	-0,771	0,984	346,6	-6,41	0,953
3.	556,93	-1,025	0,877	1057,4	-6,12	0,779
4.	1745,05	-0,570	0,775	2671,8	-4,08	0,756
5.	1499,92	-1,276	0,947	4938,6	-10,0	0,815
6.	347,15	-1,598	0,829	1405,1	-12,2	0,926
7.	134,04	-0,753	0,755	274,9	-5,82	0,594
8.	285,7	-0,825	0,876	50,9,5	-5,38	0,809

За пресмятане на концентрацията  $C$  в точка от равнината с полярни координати  $(R, \vartheta)$  се правеше линейна интерполация по азимуталния ъгъл  $\vartheta$ :

$$(12) \quad C(R, \vartheta) = C(R, \vartheta_i) + [C(R, \vartheta_{i+1}) - C(R, \vartheta_i)](\vartheta - \vartheta_i) / (\vartheta_{i+1} - \vartheta_i),$$

където  $\vartheta_i$  е азимуталният ъгъл на  $i$ -ия лъч,  $C(R, \vartheta_i)$  - концентрацията на разстояние  $R$  от завода пресметната по формула ( 11 ), в която коефициентите са взети от Таблица 1. Чрез пресмятане на концентрацията на олово в почвата в мрежа от точки, беше построена карта на замърсяването с олово в района на КЦМ.

Намерените коефициенти бяха използвани от друга, написана от автора специално за целта програма, която пресмяташе концентрацията на подвижните форми на оловото по формула ( 12 ) в правоъгълна мрежа от точки и съпоставя-

ше на пресметнатата стойност подходящ символ, така че след разпечатка да се получи карта на замърсяването - изображение в различни степени на сивото - най-тъмно при максимална концентрация и най-светло при минимална.

#### **4.5. Моделиране динамиката на изменение на витамин С в стафиди**

Беше извършена обработка на данни от изследване за влиянието на съхранението на стафиди от 2 безсеменни сорта грозде и 4 различни технологии за сушене върху съдържанието на витамин С. Съдържанието на витамин С в пробите стафиди е измерено след периоди на съхранение от 1, 3 и 12 месеца. Обработката на данните беше извършена от автора в следната последователност. След представяне на данните в графичен вид с помощта на програмата MS Excel, стана ясно, че зависимостта на концентрацията на витамин С от времето може да се моделира със линейна функция:  $C = A \cdot t + B$ , в която  $t$  е времето на съхранение в месеци. Коефициентите на функцията  $A$  и  $B$ , и коефициента на детерминация  $R^2$  са пресметнати с MS Excel и представени в Таблица 2. За пресмятането са използвани наличните в Excel функции.

Таблица 2

Коефициенти на модела и коефициент на детерминация на зависимостта на съдържанието на витамин С в стафиди от времето на съхранение.

<b>Вариант</b>	<b>A</b>	<b>B</b>	<b>R<sup>2</sup></b>
P1HH	-0,0769	2.2110	0.7873
P1TC	-0.0299	0.3297	1
KCHH	-0.6312	8.4161	0.9795
KCHC	-0.6803	9.3386	0.9620
KCTH	-0.3777	5.0470	0.8977
KCTC	-0.1413	1.7118	0.8444

Означенията на вариантите са следните:

P1HH - Русалка 1, конвекционно сушене с горещ въздух;

P1HC - Русалка 1, слънчево сушене за 3 седмици и следващо конвекционно сушене с горещ въздух;

P1TH - Русалка 1, третиране и конвекционно сушене с горещ въздух;

P1TC - Русалка 1, третиране, слънчево сушене за 3 седмици и следващо конвекционно сушене с горещ въздух

КСНН - Коринт Сидлес, конвекционно сушене с горещ въздух;

КСНС - Коринт Сидлес, слънчево сушене за 3 седмици и следващо конвекционно сушене с горещ въздух;

КСТН - Коринт Сидлес, третиране и конвекционно сушене с горещ въздух;

КСТС - Коринт Сидлес, третиране, слънчево сушене за 3 седмици и следващо конвекционно сушене с горещ въздух;

Резултатите от това изследване са описани в доклад с автори Михайлова, Георгиев, и др., (1996).

#### ***4.6. Моделиране динамиката на цветни характеристики на сушени гъби***

Цветни характеристики се наричат координатите  $L^*$ ,  $a^*$  и  $b^*$  на цвета на изследвания образец, в цветното пространство. Стойностите им се измерват със спектрофотометър (в случая е използван: PU 8800, Pye Unicam Ltd., England). Обработени са данни от измервания, извършени след време на съхранение на пробите сушени гъби от 1, 2 и 12 месеца.

От физични съображения приемаме, че изменението на състоянието на сушените гъби представлява процес на релаксация, при който всички техни параметри се изменят с времето експоненциално с различни времена на релаксация, стремяйки се да достигнат до постоянни равновесни стойности. Ето защо моделираме зависимостта на всяка от цветните характеристики  $L^*$ ,  $a^*$  и  $b^*$  от времето  $t$  с функция от вида:  $F(t) = A + Be^{\frac{-t}{c}}$ . Определянето на коефициентите  $A$ ,  $B$  и  $c$ , които минимизират сумата от квадратите (1) е извършено по метода на случайните направления за намиране екстремум на функция на много променливи (Стоянов, 1990). Този алгоритъм беше осъществен чрез написана от автора програма, а резултатите представени в доклада с автори Михайлова, Георгиев и др. (1996).

#### ***4.7. Моделиране многогодишната тенденция в изменението на добива от пшеница и царевица, при отглеждането им в двуполка и прилагане на постоянен режим на торене***

Беше съставен модел с цел обработка на данни от стационарен полски опит провеждан от 1968 до 2000 г. В опита се редуват две култури пшеница и царевица, и се изпитват различни варианти на торене. Опитът е бил заложен по метода решетков квадрат в четири повторения с големина на опитната парцела  $63 \text{ m}^2$ , а на реколтната -  $20 \text{ m}^2$ . Обработени са данните за добива от пшеница и царевица при варианти на торене:  $N_0P_0K_0$ ;  $N_{12}P_0K_0$ ;  $N_0P_{12}K_0$ ;  $N_0P_0K_{12}$ ;  $N_{12}P_{12}K_0$ ;  $N_0P_{12}K_{12}$ ;  $N_{12}P_{12}K_{12}$ .

При моделиране динамиката на изменение на добива в различните варианти на опита, изходихме от съображението, че когато в продължение на много години се прилага една постоянна земеделска практика (в случая вариант на торене), системата на производство постепенно се доближава до едно състояние на динамично равновесие, след установяване, на което добивът остава постоянен. Както е известно процесът на установяване на равновесие в дадена система се описва с експоненциални функции. Ето защо избраният модел за описание на изменението на добива във времето е от вида:

$$y = y_e + A \cdot \exp[-a \cdot (t - t_0)],$$

в което  $y$  е средният добив в  $\text{kg/dca}$ ,  $t$  е годината на отчитане,  $t_0$  - първата година на опита,  $y_e$ ,  $A$  и  $a$  са подлежащи на определяне параметри. Параметрите в уравнението имат ясна практическа интерпретация:

$y_e$  представлява средният добив в системата на производство, който ще се установи след време, когато системата достигне състоянието на равновесие. Да го наречем равновесен добив.

$A$  представлява отклонението от равновесния добив в началото, при започване на прилагането на съответния вариант на торене.

$a$  е параметър, който определя бързината за достигане до състоянието на равновесие. Неговата реципрочна стойност  $\tau = 1/a$  може да се интерпретира като време на релаксация на системата.



За определяне на параметрите на модела  $y_e$ ,  $A$  и  $a$  може да се използват възможностите, които предлага съставената от автора програма AnalysisTools, или след съставяне на подходяща таблица може да се определят и чрез процедурата tools/solver с програмата MS Excel. Получените резултати и с двете програми съвпадат. Резултатите от моделирането са представени в две отделни публикации: Кирчев, Хр., В. Георгиев, М. Василева, 2001 и Кирчев, Хр., В. Георгиев, М. Василева, 2002.

#### ***4.8. Изследване влиянието на някои абиотични фактори върху динамиката на диапаузиране при гъсениците на ябълковия плод червей***

Извършен е корелационен анализ върху данни за минималната температура на въздуха, максималната температура на въздуха, средноденонощната температура, относителната влажност, валежите и процента диапаузирали гъсеници на ябълковия плод червей. Въз основа на този анализ са направени изводи:

1. Върху процента на диапаузирали гъсеници най-силно влияе температурата, като тенденцията е този процент да се увеличава с понижаване на температурите.

2. За установяване на регресионни уравнения, приложими за прогнозиране на диапаузирането са необходими повече данни.

Приносът на автора в посочената работа се състоеше в написване на програмата за извършване на изчисленията. Това беше програма на Basic, която поради технически причини е изгубена, но сега същите изчисления могат да се направят и с написаните по-късно програми M1 или Analyses tools.

#### ***4.9. Изследване на ентомофауната по люцерната***

##### ***4.9.1 Изследване на гостоприемниковата способност на люцерната***

За проверка на статистическата хипотеза за принадлежност на плътностите на насекомни популации в два сорта люцерна към една и съща генерална съвкупност, беше приложен  $U$  критерият на Уилкоксон. Основанието за приложение на този критерий беше, че той е непараметричен и не зависи от вида на разпределението на величините. Плътността на много насекомни популации е устано-

вено, че има поасоново разпределение, затова в случая би било некоректно да се прилагат критерии за проверка на статистически хипотези, отнасящи се за нормално разпределение като критерият на Стюдънт или критерият на Фишер. За изчисляване на прилаганата при U критерия, статистиката - броя инверсии между стойности от двете извадки, от автора беше написана подходяща функция на Turbo Pascal. Процедура за пресмятане на инверсиите сега има в модул StatUnit.pas и в програмата AnalysisTools.

#### **4.9.2 Изследване на корелацията между полезни и вредни видове насекоми по люцерната**

При това изследване (Babricova, Georgiev и др. 1994) са приложени стандартни статистически тестове (t – критерий на Стюдънт и интервална оценка за коефициента на корелация). Пресмятанията са извършени със съставени от автора програми.

#### ***4.10. Имитационен модел на популация на люцернов листояд***

Този модел беше построен с цел да послужи при провеждане на борба с люцерновия листояд - един важен насекомен неприятел по люцерната в България. Той представлява, може би най-пълна илюстрация на изложените в теоретичната част идеи - следване на посочените етапи за построяване на модела и получаване на информация за изследваната система.

##### **4.10.1 Системен анализ**

Изследваната система, на която се строи имитационен модел в случая е популация на люцернов листояд в люцернов посев. Популацията се състои от индивиди, намиращи се в различни стадии на развитие: яйца, ларви и възрастни. Броят индивиди населяващи единица площ (в случая  $1 \text{ m}^2$ ) в растителната защита се нарича плътност на популацията. Плътността на популацията се дефинира отделно за всеки стадий, т.е. различаваме: плътност на яйцата  $E$ , плътност на ларвите  $L$  и плътност на бръмбарите (възрастните) на люцерновия листояд  $W$ . Времето отчитаме в дни, като номерираме дните с индекси  $i$ , започвайки от 1 март ( $i = 0$ ) - датата след която започва появата на неприятеля напролет. Експерименталните данни, с които разполагахме бяха данни за плътността на яйцата, ларви-

те, и възрастните, отчитани ежедневно, върху 10 квадратни опитни площадки с лице  $1 \text{ m}^2$ . Ето защо, следвайки изтъкнатото в теоретичната глава изискване за максимална простота на модела, трябва да се построи модел, който описва динамиката само на тези наблюдавани в опита величини.

Знанията за биологичните особености на люцерновия листояд показват, че развитието на популацията протича в следната последователност. Люцерновият листояд зимува в почвата в стадий на възрастно насекомо - бръмбари, заровени в почвата. Напролет със затопляне на времето бръмбарите започват да излизат на повърхността и да се хранят. Женските бръмбари през това време яйцеснасят. От снесените яйца се излюпват ларви, които след какавидиране в почвата и превръщане във възрастни формират следващото поколение. Част от бръмбарите от това поколение излизат и се появяват на повърхността, друга остава в почвата. С приближаване на есента всички бръмбари се заравят в почвата, изпадат в диапауза и така преживяват до следващата пролет. Именно тези процеси намериха отражение в съставеният модел.

#### 4.10.2 Съставяне на модела

Излизането на презимувалите бръмбари от почвата моделирахме със функцията от вида:

$$P_i = P_0 \cdot \exp\left[-\frac{(D-i)^2}{S_w}\right], \quad i \in [0, 1, 2, \dots]$$

където  $P_i$  е броят бръмбари на квадратен метър, излезли от почвата на  $i$ -ия ден,  $D$  - номера на деня на най-масово излизане,  $P_0$  и  $S_w$  - параметри, подлежащи на определяне.

Коефициентите на преживяемост, показващи каква част от началния брой индивиди остават живи след  $k$  дни, за ларвите  $A_{wk}$  и за възрастните  $A_{Lk}$  приехме, че намаляват линейно с времето по формулите:

$$A_{wk} = \begin{cases} 1 - \frac{k}{2 \cdot T_4}, & \text{при } k < T_4 \\ 0, & \text{при } k > T_4 \end{cases} \quad \text{и} \quad A_{Lk} = \begin{cases} 1 - \frac{k}{2 \cdot T_2}, & \text{при } k < T_2 \\ 0, & \text{при } k > T_2 \end{cases}.$$

$T_4$  е максималната продължителност на живот на възрастните,  $T_2$  - максималната продължителност на живот на ларвите.

Броят на снасяните от един женски бръмбар яйца на  $k$ -ия ден от излизането му от почвата приехме, че се изразява с формулата:

$$B_k = B_0 \cdot \exp\left[-\frac{(K-k)^2}{S_E}\right].$$

Частта от новоимагиниралите бръмбари излизаци на повърхността приехме, че намалява експоненциално по формулата:

$$A_{vi} = A \cdot \exp[-B \cdot (i - T_3)].$$

Плътноста на индивидите от трите стадия се изчисляваше по формулите:

$$E_i = \sum_{k=0}^{i-1} B_k \cdot Q \cdot A_{wk} \cdot P_{i-k}$$

$$L_i = \begin{cases} 0, & \text{когато } i < T_1 \\ \sum_{k=0}^M A_{Lk} \cdot A_i \cdot E_{i-T_1-k}, & \text{когато } i > T_1 \end{cases}$$

$$M = \max\{T_2, i - T_1 - 1\}$$

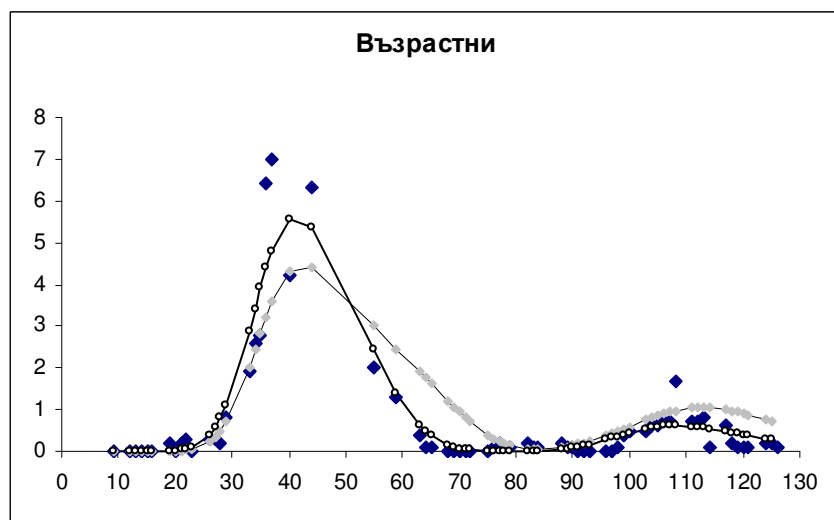
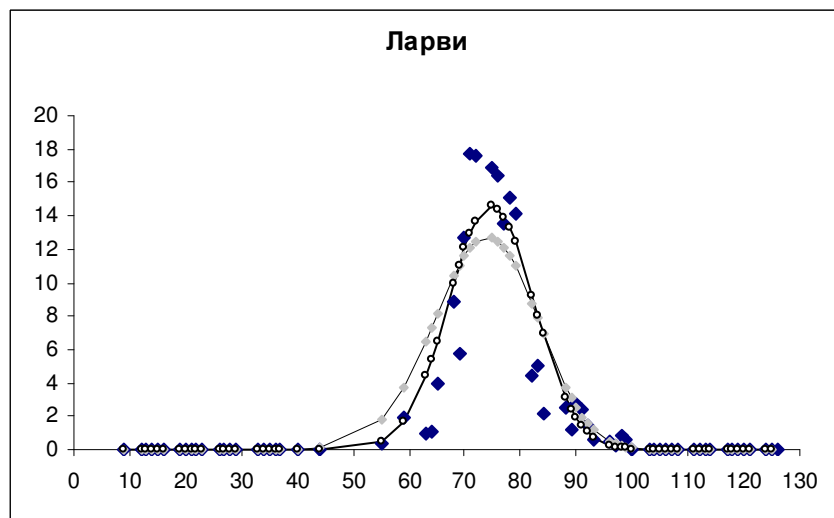
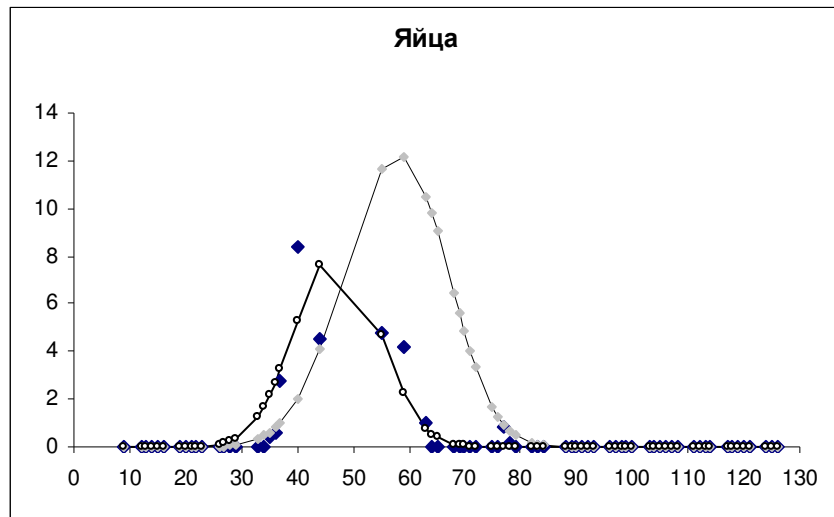
$$W_i = \sum_{k=0}^{i-1} A_{wk} \cdot P_{i-k} + A_{vi} \cdot \sum_{k=0}^{i-T_3-1} E_{i-T_3-k}.$$

#### 4.10.3 Параметризиране на модела

В така съставеният модел, присъстват редица подлежащи на определяне параметри, които не са известни предварително. Параметризирането беше извършено по метода на случайните направления за намиране екстремум на функция на много променливи (Стоянов, 1990), приложен за намиране на стойности на неизвестните параметри, които минимизират сумата от квадрати на разликите между изчислените с модела и наблюдаваните в опита стойности.

#### 4.10.4 Валидиране

След параметризиране, моделът възпроизвежда с добро приближение наблюдаваните опитни данни, което се вижда от приложените графики (Фигура 3). Беше направен опит да се прогнозира развитието на популацията на люцернов листояд с помощта на този модел. За целта моделът беше параметризиран по данните от началните дни от наблюдението и после с него беше определена динамиката на популацията през следващите дни. На Фигура 3 е показана и кривата на прогнозираната по този начин динамика.



Фигура 3. Резултати от имитационно моделиране на популация на люцернов листояд. ◆ - експериментални точки. —○— - динамика, получена с оптимизиран по данните модел. —◆— - прогнозирана с модела динамика.

#### ***4.11. Математическо моделиране на взаимоотношенията между колорадски бръмбар и седемточкова калинка***

Този модел беше описан в доклад пред Втората национална конференция по ентомология (Георгиев и Матеева, 1997). Моделът е построен с цел интерпретиране на данните от полски опит, в който се наблюдава потискащо действие върху популацията на колорадския бръмбар в посев картофи, при наличие на седемточкови калинки в същия посев. Кратко описание на опита:

През 1986 г. в посев от 5 дка картофи сорт Ерлинг се маркираха 2 полета по 50 м<sup>2</sup> в краищата на посева, където плътността на колорадския бръмбар беше най-висока. Двете полета се оградиха с дълбоки бразди и се окантиха с преградни стени от тензух на височина 80 см. Отстраниха се всички плевелни растения за да се неутрализира действието на страничните фактори и се създадат еднакви условия в тях.

Бяха използвани калинки, събрани от житни и люцернови посеви нападнати от листни въшки. Преди да се внесат калинките, беше отчетена плътността на стадията на колорадския бръмбар в двете полета. В едното поле бяха внесени калинки в съотношение 1:10 яйца и 1:2 ларви на колорадски бръмбар.

В продължение на 22 дни се определяше плътността на яйца, ларви и възрастни на колорадския бръмбар в двете полета и на калинките в моделното поле.

Получените стойности са представени в Таблица 3.

За съставяне на модел бяха издигнати следните хипотези: През  $n$ -ия ден увеличението на плътността на ларвите  $L_n$  се дължи единствено на излюпването на яйцата с плътност  $E_n$ . Намаляването на  $L_n$  в контролния вариант се дължи единствено на постоянна естествена смъртност, а в моделния и на постоянната атака на хищните калинки имащи плътност  $H_n$ . Увеличението на плътността на възрастните  $V_n$ , се дължи единствено на постоянно превръщане на част от ларвите във възрастни, а намалението - на постоянната естествена смъртност и на постоянна атака на хищниците.

Таблица 3

Данни за динамиката на плътността на колорадски бръмбар и седемточкова калинка.

Дата	Колорадски бръмбар			Седемточкова калинка			Колорадски бръмбар – контрола		
	Яйца	Ларви	Бръмбари	Яйца	Ларви	Бръмбари	Яйца	Ларви	Бръмбари
29.06	10	255	42	0	0	137	10	242	44
30.06	9	261	39	0	0	136	10	241	43
01.07	5	230	40	0	0	138	11	242	43
02.07	5	211	38	0	0	132	10	244	40
03.07	4	209	33	0	0	131	9	245	37
04.07	0	187	32	0	0	109	7	247	32
05.07	0	161	33	0	0	100	0	247	42
06.07	0	123	17	0	0	91	0	249	41
07.07	0	85	18	0	0	91	0	207	25
08.07	0	72	17	0	0	95	0	201	29
10.07	0	71	26	0	0	76	0	201	37
15.07	0	46	28	0	3	53	0	170	36
20.07	0	32	31	0	5	34	0	114	17

Въведени са следните означения:

$\alpha_L$  - коефициент на излюпване на яйцата, който показва каква част от наличните яйца се излюпват през съответния ден,

$\beta_L$  - коефициент на естествена смъртност на ларвите - показва каква част от наличните живи ларви умират през съответния ден поради фактори, различни от наличието на калинки,

$\gamma_L$  - "консумативна способност" на калинките спрямо ларвите на колорадския бръмбар - представлява средният за ден брой ларви на колорадския бръмбар, унищожавани от една калинка.

Коефициентите  $\alpha_L$  и  $\beta_L$  може да се изразят в проценти.

Сега плътността на ларвите на следващия  $n+1$ -и ден се изчислява по формулата:

$$(13) \quad L_{n+1} = L_n + \alpha_L \cdot E_n - \beta_L \cdot L_n - \gamma_L \cdot H_n.$$

Въведени са и коефициенти  $\alpha_v$ ,  $\beta_v$  и  $\gamma_v$  с аналогичен смисъл но с други стойности за възрастните на колорадския бръмбар и плътността на възрастните през  $n+1$ -я ден  $V_{n+1}$  се пресмята по формулата:

$$(14) \quad V_{n+1} = V_n + \alpha_v \cdot L_n - \beta_v \cdot V_n - \gamma_v \cdot H_n.$$

При моделиране на контролната популация, полагаме  $\gamma_L = \gamma_v = 0$ .

При известни стойности на коефициентите  $\alpha_L$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\alpha_v$ ,  $\beta_v$  и  $\gamma_v$  и на плътностите на ларвите и възрастните в деня преди първото наблюдение  $L_0$  и  $V_0$ , формули (13) и (14) позволяват да се изчисли плътността на ларвите и възрастните през всички следващи дни.

В началото формулите (13) и (14) бяха заложи в електронна таблица на компютърната програма "EXCEL 5.0". Изчислените по формулите стойности графично се сравняваха с наблюдаваните в експеримента стойности от Таблица 3. Задавах се различни стойности на параметрите  $L_0$ ,  $V_0$ ,  $\alpha_L$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\alpha_v$ ,  $\beta_v$  и  $\gamma_v$  докато се получи сравнително добро приближение. Намерените стойности бяха използвани като начални в една процедура, съставена на езика за програмиране Turbo Pascal 5.0, която чрез намиране на нови по-подходящи стойности на параметрите минимизира сумата от квадрати на разликите между изчислените от компютъра и наблюдаваните в опита стойности. Споменатата процедура използва метода за намиране на екстремум на функция наричан "метод на деформируемият симплекс на Нелдер и Мид", така както е описан от Стоянов (1990).

Намерените оптимални стойности са дадени в Таблица 4, а на Фигура 4 и Фигура 5 е представена динамиката на наблюдаваната (означена с точки) плътност на ларвите и на възрастните на колорадския бръмбар в двата варианта на опита, както и изчислената с описания модел (представена с линии).

Вижда се, че получената чрез модела крива на динамиката на плътността на ларвите следва добре наблюдаваните в опита стойности, както при контролната така и при популацията, подложена на въздействие с калинки. Коефициентът на корелация между получените от модела и наблюдаваните в опита стойности е 0,972 за популацията, подложена на калинки и 0,969 - за контролната популация.

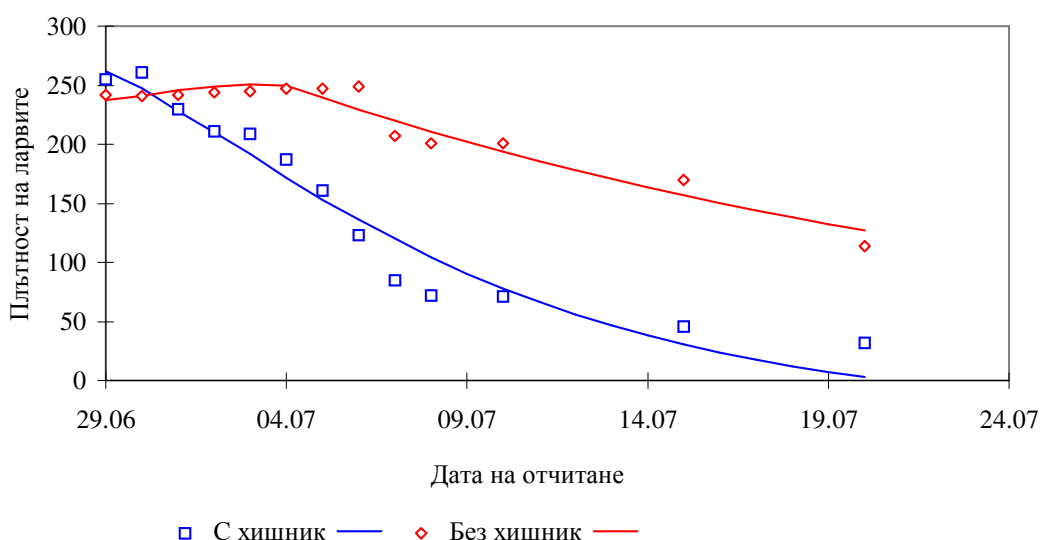


Доброто съвпадение говори, че получените стойности на първите три параметъра, дадени в Таблица 4 са точни.

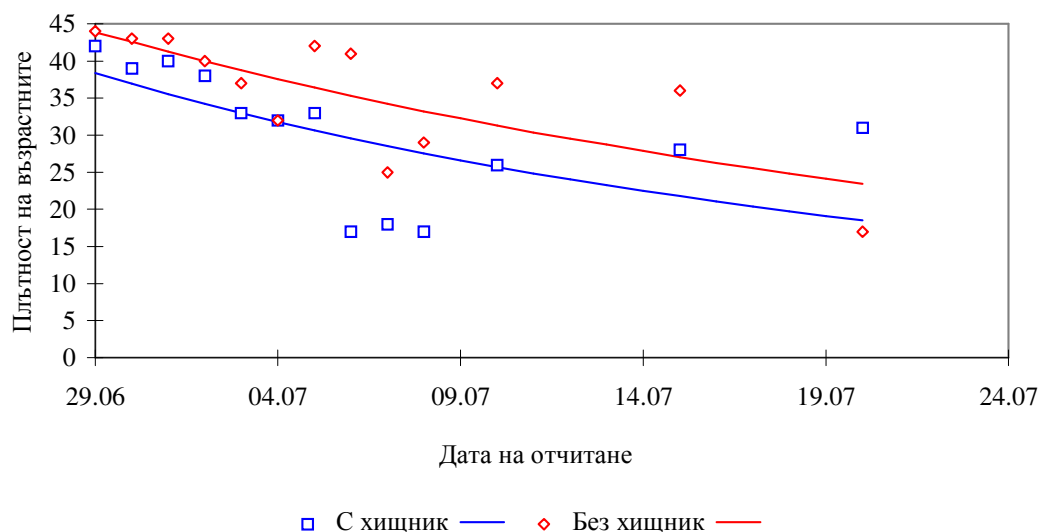
Таблица 4

Оптимизирани стойности на коефициентите на модела на взаимодействие между колорадски бръмбари и седемточкови калинки.

Коефициент на излюпване на яйцата (брой ларви от купчинка на ден)	1.34
Коефициент на естествена смъртност на ларвите (% ларви на ден)	4.14%
Консумативна способност на калинките към ларвите на колорадския бръмбар (ларви на хищник за ден)	0.114
Коефициент на превръщане на ларвите във възрастни (% ларви на ден)	-0.04%
Коефициент на естествена смъртност на възрастните (% индивиди на ден)	2.80%
Консумативна способност на калинките към възрастните на колорадския бръмбар (ларви на хищник за ден)	0.0017



Фигура 4. Динамика на плътността на ларвите на колорадски бръмбар в двете наблюдавани популации.



Фигура 5. Динамика на плътността на възрастните на колорадски бръмбар в двете наблюдавани популации.

Значително несъответствие между наблюдаваната и моделираната динамика на плътността има при възрастните на колорадски бръмбари (Таблица 4). Коефициентите на корелация са малки - 0,739 - при контролата и 0,565 при популацията, взаимодействаща с калинки. Невъзможността оптимизиращата процедура да намери стойности на параметрите, даващи по-малка грешка на модела показва, че самият модел е качествено неприложим и направените при неговото извеждане хипотези следва да се отхвърлят.

Не може да не се отбележи, че за коефициента на превръщане на ларвите във възрастни (Таблица 4) е получена отрицателна стойност, което е абсурдно. Но в случая отрицателната стойност означава, че нарастването на плътността на възрастните превъзхожда намалението в плътността на ларвите, следователно появата на нови възрастни се дължи не само на превръщането на ларви във възрастни, а вероятно и на миграция.

До друг интересен извод стигаме като сравним консумативната способност на калинките по отношение на ларвите и на възрастните на колорадския бръмбар. По отношение на ларвите тя е значително по-висока 0.114 срещу само 0.0017 по отношение на възрастните, което е около 65 пъти повече. Следовател-

но, чрез калинките може да се въздейства върху ларвеният стадий, но почти не се въздейства върху възрастния стадий на колорадския бръмбар.

#### ***4.12. Структурна схема за модел на растежа и развитието на посеви едногодишен лук***

Работата по съставяне на имитационен модел на посев лук беше започната, но не е завършена. Тук ще бъде изложена първоначалната идея за този модел и описание на осъществената част от тази идея. Ще посочим още, че към момента на пристъпване към тази работа ни беше известен един единствен модел на луков посев, описан от De Visser, (1994). Този модел обаче изисква данни за сумарната слънчева радиация на мястото на посева, с каквито ние не разполагаме, затова се пристъпи към съставяне на собствен модел, който да използва като вход данните от наблюденията, които ние провеждахме.

Производството на лук за глави чрез директна сеитба е перспективно и добре застъпено в други страни производствено направление. За популяризиране и по-широко внедряване на това производство в България е необходимо да се разработи надеждна технология, съобразена с местните условия. Два съществени елемента на такава технология са прилагането на набор от подходящи сортове и уточняване на най-подходящите срокове за засяване на всеки сорт при различни климатични условия. Във тази връзка се изисква извършване на обемиста и продължителна експериментална работа по изпитване на нови за страната сортове при различни срокове на засяване. Обемът и сроковете за изпълнение на тази работа биха могли да се намалят чрез приложение на компютърно моделиране.

Целта ни е съставяне на имитационен модел на посев лук, чрез който да се подпомогне изследването на поведението на новите сортове при нашите условия и съставянето на оптимална технология. Съставянето на такъв модел изисква изграждане на система от методологични средства които биха послужили за база при построяване на модели и на други селскостопански култури, както и на други обекти представляващи интерес в селскостопанската научноизследователска работа.

При имитационното моделиране на селскостопанските посеви прилагаме идеята за обектно ориентирано програмиране. Един селскостопански посев разглеждаме като комбинация от сорт, агротехника и местни условия. Всеки от тези три елемента разглеждаме и моделираме като отделен обект.

Луковият посев разглеждаме като съчетание (комбинация) от три независими обекта, подлежащи на самостоятелно моделиране: 1) място на посева; 2) сорт; и 3) агротехнически детайли. Схематично това може да се представи така:



Фигура 6 Структура на модел на посев едногодишен лук.

Всеки от тези обекти беше реализиран програмно на Turbo Pascal for Windows чрез отделен обектен тип, съответно: TSite, TCultivar и TCrop, за всеки от които е съставен отделен Turbo Pascal модул (unit), съответно: Site.pas, Cultivar.pas и Crop.pas (Исходният код на тези модули са представени на компакт диска приложен към настоящата работа – директория \source\Onion\). Полетата на обектите представляват параметрите, които ни интересуват, а методите им отразяват количествените връзки, между стойностите и моделират съответните процеси.

#### **4.12.1 Моделиране на мястото на посева**

Мястото на посева се моделира чрез обектния тип TSite, описан в модул Site.pas.

По-важните полета на type TSite са:

AirTF,SunRF:PChar; - имена на файлове, съдържащи метеорологични данни за съответното място на посева.

Latitude, Longitude:real; - съответно географска ширина и дължина на мястото.

По-важните методи са:

function TimeOf(T:real; W:Word):real; - Функция за изчисляване на времето на най-близкия до момент T изгрев или залез. Вторият параметър W, приема стойности \_SunRise или \_SunSet. Изчисленията се извършват по алгоритъма описан в раздел 4.13.

#### **4.12.2 Моделиране на сорта**

Модул Cultivar.pas. Модулът не е завършен. Следва да се установят кои са сортово специфичните параметри на моделираните посеви и възможно ли е да се определят по данни на семеразпространителните компании или чрез по-икономични и кратковременни опити.

#### **4.12.3 Моделиране на посева**

Посевът се моделира чрез обектен тип TCrop, модул Crop.pas.

По важните полета на type TCrop са:

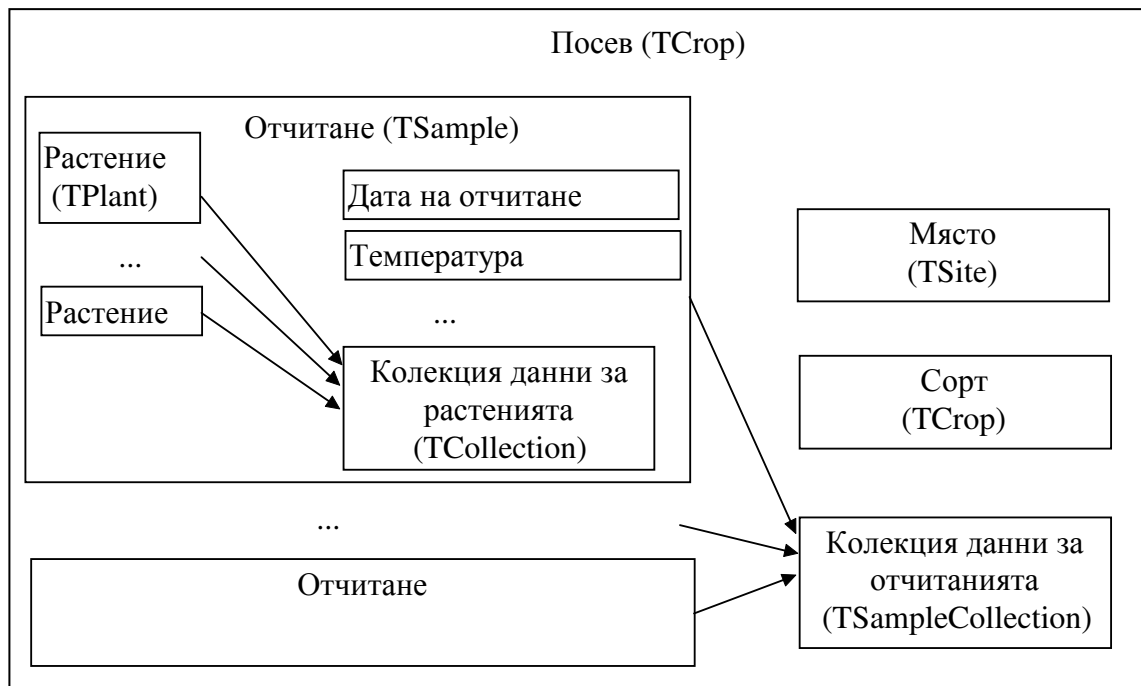
Site:PSite; - указател към обект, представящ мястото на посева.

Cultivar:PCultivar; - указател към обект, представящ сорта.

SowingTime,Emerg:real; - време на момента на сеитбата и на поникването.

Bulbs,Plants:word; - брой на растенията и на луковиците на 1 m от реда в лехата.

Sample:PSampleCollection; - указател към обект, съдържащ данните от експерименталните наблюдения върху посева. Данните от наблюденията се организират в структура, чиито йерархичен строеж и положение спрямо другите данни за посева схематично може да представим така:



Фигура 7. Йерархична структура, представяща данните от наблюденията.

Типът от който започва изграждането на тази структура е TPlant. Това е тип на обект, предназначен да съхранява данните за едно отделно растение. Полетата му съдържат стойностите получени при измерване на растенията от взетите проби. Този тип и произтичащите от него други типове са описани програмно в модул Plant.pas. Данните за няколко растения отчетени в един и същи ден се обединяват в обект от тип TCollection. Указателят към него заедно с други полета, съдържащи данни за наблюдението (дата на наблюдение, температура) образуват нова структура от тип TSample. Обектите от тип TSample, съдържащи данните за всички наблюдения върху даден посев се обединяват в структура от тип TSampleCollection, наследник на тип TSortedCollection.

#### ***4.13. Моделиране положението на Слънцето спрямо посевите***

За моделиране реакцията на посевите спрямо слънчевата радиация и влиянието ѝ върху растежа и развитието на селскостопанските култури се налага изчисляване на положението на Слънцето спрямо посева както и моментите на изгрев и залез и продължителността на деня.

Във всички известни ни модели на селскостопански култури (например: Spitters at al., 1989; Weir at al., 1984) пресмятането на посочените стойности се

извършва по приблизителни формули. Резултатите, които се получават по този начин, се различават съществено от действителните стойности на пресмятаните величини, давани в астрономическите календари и справочници. Не е правен анализ до колко това влияе върху точността на моделите като цяло, но е безспорно, че използването на по-точни алгоритми само може да подобри нещата, още повече, че възможностите на съвременните персонални компютри позволяват такова усложняване без недопустимо увеличаване на времето за обработка. Поради тези съображения е съставен алгоритъм за пресмятане, който се основава на точно (в смисъл на машинна точност - Форсайт, 1986) решаване на уравненията.

#### 4.13.1 Изчисляване положението на Слънцето

Положението на Слънцето спрямо посева в даден момент се описва чрез хоризонталните координати на Слънцето: азимут  $A$  и височина  $h$ . Азимутът  $A$  е ъгъл в хоризонталната равнина между посоката юг и посоката към вертикалната проекцията на центъра на Слънцето върху хоризонта. Отчита се в посока от юг към запад през север и изток и има стойности в интервала  $[0^0, 360^0)$ . Височината  $h$  е ъгъл във вертикална равнина между хоризонта и посоката към центъра на Слънцето. Тя е положителна, когато Слънцето е над хоризонта и отрицателна, когато е под него. Стойностите ѝ са в интервала  $[-90^0, +90^0]$ .

Теоретичната информация, на която се основават изчисленията е изложена достатъчно пълно в учебника по астрономия на Н. Бонев, 1964. Тук ще представим изводите необходими за изясняване на реализираните алгоритми.

Законът за движението на една планета се получава при решаването на механичната задача за две тела (теорията е изложена във всеки учебник по теоретична механика, например: И. И. Ольховский, 1987) и може да се представи в параметричен вид със системата уравнения:

$$(15) \quad r = a.(1 - e.\cos E)$$

$$(16) \quad M = E - e.\sin E$$

$$(17) \quad M = \frac{2\pi}{T}(t_G - \tau)$$

където:

$r$  е големината на радиус-вектора на планетата, спрямо хелиоцентричната отправна система,

$a$  - дължината на голямата полуос на орбитата,

$e$  - ексцентрицитетът на орбитата,

$E$  - параметър, известен под наименованието ексцентрична аномалия,

$M$  - така наречената средна аномалия на планетата,

$t_G$  - универсалното време (времето по Гринуич),

$\tau$  - един от моментите на преминаване на планетата през перихелия (най-близкото положение до Слънцето) по универсално време,

$T$  - интервалът време между две последователни преминавания на планетата през перихелий, нарича се аномалистична година.

Положението на планетата спрямо Слънцето в даден момент се отчита чрез истинската ѝ аномалия  $v$  (ъгълът между радиус-вектора ѝ в този момент и радиус-вектора на перихелия). Връзката между радиус-вектора  $r$  и истинската аномалия  $v$  представлява уравнение на елипса, известно от аналитичната геометрия:

$$(18) \quad r = \frac{A \cdot (1 - e^2)}{1 + e \cdot \cos v}$$

Ако се приравнят (15) и (18) и се изрази:

$$\cos v = \frac{\cos E - e}{1 - e \cdot \cos E},$$

може да се напише:

$$1 - \cos v = 2 \cdot \sin^2 \frac{v}{2} = \frac{(1 + e)(1 - \cos E)}{1 - e \cdot \cos E} = \frac{2 \cdot (1 + e) \cdot \sin^2 \frac{E}{2}}{1 - e \cdot \cos E}$$

$$\text{и} \quad 1 + \cos v = 2 \cdot \cos^2 \frac{v}{2} = \frac{(1 - e)(1 + \cos E)}{1 - e \cdot \cos E} = \frac{2 \cdot (1 - e) \cdot \cos^2 \frac{E}{2}}{1 - e \cdot \cos E}.$$

След деление и коренуване на последните две равенства се получава връзката между истинската аномалия  $v$  и ексцентричната аномалия  $E$  във вида:



$$(19) \quad \operatorname{tg} \frac{v}{2} = \sqrt{\frac{1+e}{1-e}} \cdot \operatorname{tg} \frac{E}{2}.$$

Тази връзка се представя във вид на развитие в степенен ред.

Нека е зададена неявно функция  $y(x)$  с равенството:

$$(20) \quad \operatorname{tgy} = \frac{a \cdot \sin x}{1 - a \cdot \cos x}, \quad |a| < 1.$$

$y(x)$  може да се развие в ред по степените на  $a$ . Наистина след диференциране на (20) по  $a$  се получава:

$$\frac{dy}{da} = \frac{\sin x}{1 - 2 \cdot a \cdot \cos x + a^2}.$$

По метода на неопределените коефициенти за последния израз се намира разложението:

$$\frac{dy}{da} = \sin x + a \cdot \sin 2x + a^2 \cdot \sin 3x + \dots,$$

а след интегриране по  $a$  се получава:

$$(21) \quad y = S(a, x) \equiv a \cdot \sin x + \frac{1}{2} \cdot a^2 \cdot \sin 2x + \frac{1}{3} \cdot a^3 \cdot \sin 3x \dots$$

Последното развитие ще се използва неколкократно, поради което ще го означаваме с  $S(a, x)$ .

Нека е зададена неявно функция  $y(x)$  с равенството:

$$(22) \quad \operatorname{tgy} = n \cdot \operatorname{tg} x$$

След преобразувания се получава:

$$\begin{aligned} \operatorname{tg}(y-x) &= \frac{\operatorname{tgy} - \operatorname{tg} x}{1 + \operatorname{tgy} \cdot \operatorname{tg} x} = \frac{(n-1) \cdot \operatorname{tg} x}{1 + n \cdot \operatorname{tg}^2 x} = \frac{(n-1) \cdot \sin x \cdot \cos x}{\cos^2 x + n \cdot \sin^2 x} = \\ &= \frac{2(n-1) \cdot \sin x \cdot \cos x}{1 + \cos 2x + n - n \cos 2x} = \frac{(n-1) \cdot \sin 2x}{(n+1) - (n-1) \cos 2x} = \left( \frac{n-1}{n+1} \sin 2x \right) / \left( 1 - \frac{n-1}{n+1} \cos 2x \right). \end{aligned}$$

Последният израз е от вида (20), и за  $y-x$  може да се приложи разложението (21), така за  $y(x)$ , зададена неявно с (22) се получава следното разложение:

$$(23) \quad y = x + S\left(\frac{n-1}{n+1}, 2x\right)$$

Сега да сравним (19) със (22) и да приложим (21). Получаваме:

$$(24) \quad v = E + 2.S(a, E) \quad \text{при} \quad a = \left( \sqrt{\frac{1+e}{1-e}} - 1 \right) / \left( \sqrt{\frac{1+e}{1-e}} + 1 \right),$$

което представя в явен вид връзката (19) между истинската аномалия  $v$  и ексцентричната аномалия  $E$  и позволява извършване на изчисленията.

Изчисляването на хоризонталните координати на Слънцето може да се извърши в следната последователност:

**А.** От официалното поясно време на мястото на посева  $t$  се преминава към универсално време  $t_G$

$$(25) \quad t_G = t + 3600.H,$$

където  $H$  е номерът на часовия пояс на мястото на посева.

**Б.** Пресмята се средната аномалия на Земята по формулата (17).

**В.** Числено се решава уравнение (16) относно ексцентричната аномалия  $E$ .

**Г.** От (24) се намира истинската аномалия  $v$ .

**Д.** Изчислява се координата “дължина” на Слънцето  $\lambda$  по формулата:

$$\lambda = v + \omega,$$

където  $\omega$  е дължината на пролетната равноденствена точка.

**Е.** По еклиптичните координати на Слънцето (дължина  $\lambda$ , и ширина  $0$ ) се изчисляват неговите екваториални координати (ректасцензия  $\alpha$ , часов ъгъл  $\theta$  и деклинация  $\delta$ ).

Ректасцензията се намира от формулата:

$$\operatorname{tg} \alpha = \operatorname{cose} . \operatorname{tg} \lambda,$$

в която  $\epsilon$  е ъгълът между равнината на екватора и равнината на еклиптиката. Тази формула е от вида (22) и за намиране на  $\alpha$  можем да приложим (23):

$$\alpha = \lambda + S\left(\frac{\operatorname{cose} - 1}{\operatorname{cose} + 1}, 2\lambda\right).$$

Часовият ъгъл на Слънцето  $\theta$  се намира по формулата:

$$(26) \quad \theta = \frac{2\pi(t_G - \tau_0)}{D_s} - \alpha,$$

в която:

$t_G$  е универсалното време,

$\tau_0$  -универсалното време на едно от преминаванията на Слънцето през меридиана на мястото на посева,

$D_s$  -продължителност на звездното денонощие в секунди.

Стойностите на  $\sin\delta$  и  $\cos\delta$  се намират по формулите:

$$\sin\delta = \sin\lambda \cdot \sin\varepsilon, \quad \text{и} \quad \cos\delta = \sqrt{1 - \sin^2\delta}.$$

**Ж.** От екваториалните координати на Слънцето се намират хоризонталните му координати - височина  $h$  и азимут  $A$  като се използват известните от астрономията формули:

$$(27) \quad \sinh = \sin\varphi \cdot \sin\delta + \cos\varphi \cdot \cos\delta \cdot \cos\theta$$

$$(28) \quad \cosh \cdot \sin A = \cos\delta \cdot \sin\theta$$

$$(29) \quad \cosh \cdot \cos A = -\sin\delta \cdot \sin\varphi + \cos\delta \cdot \sin\varphi \cdot \cos\theta,$$

където  $\varphi$  е географската ширина на мястото на посева.

В представените до тук формули участвуват редица астрономически константи: ексцентрицитетът на земната орбита  $e$ , дължината на пролетната равноденствена точка  $\omega$ , ъгълът между равнината на екватора и равнината на еклиптиката  $\varepsilon$ , които в действителност макар и бавно се изменят с времето. Това изменение е отчетено чрез емпирични формули:

$$e = 0.01675104 - 0.00004180 \cdot t - 0.000000126 \cdot t^2$$

$$\omega = 281^{\circ}13'15.17'' + 6189.03'' \cdot t + 1.63'' \cdot t^2 + 0.012'' \cdot t^3$$

$$\varepsilon = 23^{\circ}27'8.26'' - 0.46844'' \cdot t - 0.0000060 \cdot t^2$$

в които времето  $t$  е в юлиански векове. Тези зависимости са взети от Н. Бонев, 1964, който цитира справочника Newcomb.

#### 4.13.2 Изчисляване времето на изгрева и залеза и продължителността на деня

За момент на изгрева се приема, моментът в който слънчевият диск се допре отдолу до въображаемия хоризонт, а за момент на залеза - моментът на пълно скриване на слънчевият диск под този хоризонт. Очевидно това са моментите, в които Слънцето има височина  $h_0$ , равна на отрицателния ъгъл под който се вижда слънчевият радиус  $R$  от Земята. Прилагайки (15) получаваме:

$$h_0 = -\frac{R}{r} = -\frac{R}{A \cdot (1 - e \cdot \cos E)} = -\frac{2h_1 h_2}{h_1 + h_2 + (h_1 - h_2) \cdot \cos E}$$

В последната формула с  $h_1$  и  $h_2$  са означени минималната и максималната стойност на ъгъла под който се вижда слънчевия диск, когато се намира близо до хоризонта. Върху тези стойности оказва влияние и рефракцията в атмосферата. Поради това, че в прегледаната литература няма информация за стойностите на  $h_1$  и  $h_2$  те са подбрани, така че изчислените стойности на времето на изгрев и залез да съвпада с посоченото в астрономическите календари.

Часовият ъгъл на Слънцето  $\theta_0$ , отговарящ на височина  $h_0$  се изразява от (27):

$$(30) \quad \theta_0 = \arccos\left(\frac{\sin h_0 - \sin \varphi \cdot \sin \delta}{\cos \varphi \cdot \cos \delta}\right)$$

След като в (26) се замени универсалното време  $t_G$  от (25) с официалното поясено време на мястото на посева  $t$  и се реши полученото уравнение спрямо  $t$  се получава едно първо приближение за времето на изгрев и залез:

$$(31) \quad t_{1,2} = D_s \cdot \left[ n + (\alpha \pm \theta_0) / (2\pi) \right] + \tau_0 - 3600 \cdot H$$

С индекс 1 и знак “-” е означено времето на изгрева, а с индекс 2 и знак “+” - времето на залеза.  $n$  е поредният номер на звездното денонощие от приетото начало за отчитане на времето.

При условие, че стойностите на ректасцензията и деклинацията в (30) и (31) се отнасят точно за моментите време  $t_1$  и  $t_2$ , изчислените от (31) стойности, биха

били точни, но понеже  $t_1$  и  $t_2$  не са известни предварително се прилагат последователни итерации. Отначало се пресмятат  $\alpha$  и  $\delta$  за 12 часа по пладне на съответната дата и от (31) се получава първото приближение за времето на изгрев или залез, отново се пресмятат  $\alpha$  и  $\delta$  вече за тези моменти, намират се нови още по-точни стойности и т. н. до постигане на нужната точност.

Така съставените алгоритми са осъществени в програмния модул `Astro.pas` и `Site.pas` (директория `/source/Onion/` на компакт-диска приложение).

#### ***4.14. Изводи:***

1. Има тясна връзка между обработката на опитни данни и съставянето на модели на изследваните системи. Дори и традиционните статистически методи за обработка на данни могат да се разглеждат, като частен случай на обработка на данни чрез използване на модели. Във всеки конкретен случай е необходимо да се състави най-подходящ за него модел - емпиричен, статистически, динамичен, имитационен. Избора на модел се прави при системния анализ на изучавания обект.

2. За съставяне на моделите и извършване на изчисленията в повечето случаи трудно се намира подходящ готов софтуер и това налага да се състави собствена библиотека от софтуерни средства, която непрекъснато да се допълва с нови елементи в зависимост от възникващите нови задачи.

3. Във всеки от случаите се вижда, че след съставяне на подходящ модел се правят научно обосновани изводи, които водят до нови знания за изследвания обект - знания, които трудно биха били достигнати чрез методи, които не включват моделиране.

## Заклучение

Практиката при провеждане на експериментални изследвания в областта на селското стопанство показва, че съществува необходимост от използване и на други методи за анализиране на експерименталните данни, допълващи обичайно прилаганите статистически методи. До тази необходимост се стига, например, при анализиране на стари данни с цел достигане до нови пропуснати по някакви причини преди изводи, или при анализиране на данни, събирани в допълнение към обичайните схеми за залагане и провеждане на полски или лабораторни опити.

Като беше взет предвид напредъкът постигнат в момента в областта на компютърното моделиране на изследваните обекти в редица научни области – селско стопанство, биология, техника, физика и т.н. и след анализ на технологията за създаване на модели, авторът на настоящият труд достигна до решението да съчетае обработката на експериментални данни с компютърно моделиране и да приложи тази идея във всички случаи за обработка на опитни данни, с които се сблъска.

За съставяне на подходящ за всеки от случаите модел беше необходим набор от софтуерни средства, представляващи отворена система, която да може непрекъснато да се подобрява и допълва. Изграждането на такава система, състояща се от няколко програмни модули, написани с Delphi и Turbo pascal е започнато и беше описано в изложението на този труд.

Въпреки, че съществуват и готови програми, с които може да се направят описаните в изложението модели и обработки на данни, съставянето на собствена система от софтуерни модули има редица предимства:

- системата се изгражда по-избрана от нас, съобразена с конкретната работа схема;
- пълното познаване на тази схема, позволява при необходимост бързо и оперативно внасяне на изменения в програмния код;
- разполагането с програмния код позволява многообразно да се съчетават части от този код за съставяне на нови програми и решаване на конкретни зада-

чи; съставените по този начин програми включват само необходимият минимум елементи, икономични са, не изискват големи машинни ресурси;

- отвореността на системата дава възможност за непрекъснато подобряване и допълване с нови елементи, позволяващи използване на системата и при решаване на непредвидени в началото задачи.

Макар и голяма част от извършената работа да се състоеше в рутинно програмиране на познати алгоритми може да се приеме, че направеното е ценно защото винаги е било продиктувано от практическата необходимост от обработка на конкретен набор от данни. Всеки път обработката на данните беше извършвана в необходимите срокове и водеше до извличане на нови ценни изводи за експериментално изследваните обекти. Ето защо може да се посочат следните:

#### **Научни и научно-приложни приноси:**

1. Предложена е една нетрадиционна гледна точка към обработката на данни от опити, провеждани във връзка с изследвания от областта на селското стопанство - обработката на данни е свързана с моделиране на изследваните обекти.

2. Съставени са модели на обекти от селскостопански изследвания (популация люцернов лисояд, взаимоотношение между колорадски бръмбар и седемточкова калинка, посев лук и др.) и са използвани за обработката на опитните данни, получени при тези изследвания.

3. Съставени са програмни модули, написани на Turbo Pascal и Delphi, които са използвани при създаване на моделите и обработката на данните, и може да се използват и в бъдеще за съставяне на други модели и програми.

4. Чрез съставените модели и софтуер са направени изводи за изследваните обекти, до които трудно би могло да се стигне само с традиционни методи за обработка на данни.

5. С написаният софтуер, се поставя начало за бъдещо по-голямо разширяване на работата по обработка на данни от селскостопански експерименти чрез прилагане на метода на моделиране. Дава се и възможност за създаване на програмен продукт, който би могъл да бъде предложен на пазара.

## Литература

1. Айвазян, С. А., З. И. Бежаева, О. В. Староверов, 1974, Классификация многомерных наблюдений, Москва, "Статистика".
2. Бихеле, З. Н., Х. А. Молдау и Ю. К. Росс, 1980, Математическое моделирование транспирации и фотосинтеза растений при недостатке почвенной влаги, Гидрометеиздат, Ленинград.
3. Божанов, Е. И. Вучков, 1979, Статистически решения в производството и научните изследвания, Държавно издателство "Техника", София.
4. Бонев, Н., 1964, Астрономия, "Наука и изкуство", София
5. Вольвач, В. В., 1978, Моделирование влияния агрометеорологических условий на развитие колорадского жука, Гидрометеиздат, Ленинград.
6. Воробьев, Е. М., 1998, Введение в систему "Математика", Финансы и статистика, Москва.
7. Под редакцията на Гвишиани, Д. М., И. Б. Новика, С. А. Пегова, 1986, Природа моделей и модели природы, "Мысль", Москва.
8. Георгиев, В. Г., А. Матеева, 1997, Математическо моделиране на взаимоотношенията между *Leptinotarsa decemlineata* Say и *Coccinella septempunctata*, IV Национална конференция по ентомология, София, 2-6 юни, 1997
9. Горанова, Г., В. Георгиев, Цв. Миховски, 2004, Селекционна оценка на диплоидни образци червена детелина (*Tr. pratense* L.), 50 години Институт по фуражните култури – Плевен, Юбилейна научна конференция "Устойчиво фуражно производство за ефективно животновъдство", 8-10 юни 2004 г., Плевен.
10. Димова, Д, Е. Маринков, 1999, Опитно дело с биометрия, Академично издателство на ВСИ, Пловдив.
11. Запрянов, З., Е. Маринов, 1978, Опитно дело с биометрия, Пловдив, Издателство "Хр. Г. Данов".



12. Кабишева, Н. В., 1971, Косвенные расчеты климатических характеристик, Гидрометеорологическое издательство, Ленинград.
13. Клевцов, А. В., 2002, Математически модели и анализ на данни при почвен агротехнологичен трансфер, Национален център за аграрни науки, Институт по почвознание "Н. Пушкиров", София.
14. Колективен труд, 1983, Контроль и прогноз – основа целенаправленной защиты растений. Принципы, методы и технология выявления и прогноз распространения вредных организмов в странах СЭВ, Коллективный труд ученых стран, участвующих в разработке темы “Контроль и прогноз” в рамках СЭВ, Академия сельскохозяйственных наук германской демократической республики.
15. Лакин, Г. Ф., 1973, Биометрия, “Высшая школа”, Москва.
16. Мандель, И. Д., 1988, Кластерный анализ, Москва, "Финансы и статистика".
17. Митков, А., Д. Минков, 1989, Статистически методи за изследване и оптимизиране на селскостопанска техника, I част, “Земиздат”, София
18. Митков, А., С. Кардашевски, 1977, Статистически методи в селскостопанската техника, “Земиздат”, София
19. Ольховский, И. И., 1987, Курс теоретической механики для физиков, Издательство московского университета, Москва.
20. Перегудов, Ф. И., Ф. П. Тарасенко, 1989, Введение в системный анализ, "Высшая школа", Москва.
21. Петросян, Л. А., В. В. Захаров, 1986, Введение в математическую экологию. Издательство Ленинградского университета, Ленинград.
22. Потемкин, В. Г., 1999, Система инженерных и научных расчетов Matlab 5.x, Диалог – МИФИ, Москва.

23. Садовски, А. Н. 1986, Статистическо моделиране на зависимостта “почва-добив” – теоретични и експериментални основи. Дисертация за получаване на научната степен “Доктор на селскостопанските науки”, София.
24. Сенгалевич, Г., 1990, Отчет за изследването на екологическата обстановка в района на КЦМ, Пловдив.
25. Сендов, В., В. Попов, 1976, Числени методи (в два тома: първа и втора част), Издателство “Наука и изкуство”, София
26. Сиротенко, О. Д., 1981, Математическое моделирование теплового режима и продуктивности агроэкосистем, Гидрометеиздат, Ленинград.
27. Стойчев, Любомир Щ., Анатолий С. Антонов, Иван А. Филипов, 1989, Програмни езици по изкуствен интелект, Държавно издателство "Техника", София
28. Стоянов, Стоян, 1990, Методи и алгоритми за оптимизация, Държавно издателство “Техника”, София.
29. Тихонов, А. Н., М. В. Уфимцев, 1988, Статистическая обработка результатов экспериментов, Издательство Московского Университета, Москва.
30. Тончева, Р. Т., 2001, Математически модели на продуктивността на царевица в зависимост от гъстотата на посева, дисертация за присъждане на образователната и научна степен "доктор", Институт по почвознание "Н. Пушкиров", София.
31. Фишер, Р. А. 1958, Статистические методы для исследователей, Москва, "Госстатиздат".
32. Форсайт, Дж., 1986, Компютърни методи за математически пресмятания.
33. Франс, Дж., Х. М. Торнли, 1987, Математические модели в сельском хозяйстве, Агропромиздат, Москва.
34. Хог, Робърт, Алън Крейг, 1982, Увод в математическата статистика, Издателство "Техника", София.

35. Шанин, Йордан Иванов, 1977, Методика на полския опит, Издателство на БАН, София.
36. Benz, J., R. Hoch and T. Legovi, 2001, ECOBAS — modelling and documentation, *Ecological Modelling*, Volume 138, Issues 1-3 , 15 March 2001, Pages 3-15
37. Francis, B. 1994, *Statistical Software Reviews*, *Applied Statistics*, 43, No 4, pp. 679-682
38. Francis, B. 1995, *Statistical Software Reviews*, *Applied Statistics*, 44, No 1, pp. 139-141
39. De Visser, C. L. M., 1994, ALCEPAS, an onion growth model based on SUCROS87. I. Development of the model, *Journal of Horticultural Science*, (1994) 69 (3) 501-518.
40. Dumouchel, William; Lane, Thomas. 1995, *Journal of Computational & Graphical Statistics*, Dec95, Vol. 4 Issue 4, p251
41. Hoch, R., T. Gabele and J. Benz, 1998, Towards a standard for documentation of mathematical models in ecology, *Ecological Modelling* Volume 113, Issues 1-3 , 2 November 1998, Pages 3-12
42. Ivanova, I., D. Kostova, V. Georgiev, 2002, Grouping of *Phaseolus vulgaris* cultivars and breeding lines according to their susceptibility to Cucumovirus CH 39-BG on the basis of cluster analysis, First symposium on horticulture, 16-20 October, 2002, Ohrid, Republic of Macedonia.
43. Knorrenschild, M., R. Lenz, E. Forster and C. Herderich, 1996, UFIS: a database of ecological models, *Ecological Modelling*, Volume 86, Issues 2-3 , May 1996, Pages 141-144
44. Kornbrot, Diana 2002, Statistical software review, *British Journal of Mathematical and Statistical Psychology*, 55, 391–392
45. Linfield, Kevin 2002, Open Window to Unix, *Computing Canada*, August 23.

46. Pregibon, Daryl. 1995, *Journal of Computational & Graphical Statistics*, Dec95, Vol. 4 Issue 4, p257.
47. Review of the SYSTAT v.10.2 software package, 2003, *British Journal of Mathematical and Statistical Psychology*, 56, 189–195
48. Spitters, C. J. T., H. van Keulen and D. W. G. van Kraalingen, 1989, A simple and universal crop growth simulator: SUCROS87, in *Simulation and systems management in crop protection*, R. Rabinge, S. A. Ward and H. H. van Laar (Editors), Pudoc, Wageningen.
49. Thornley, John N. M. & Ian R. Johnson, 1990, *Plant and Crop Modelling. A Mathematical Approach to Plant and Crop Physiology*, Clarendon Press, Oxford.
50. Weir, A. H., P. L. Bragg, J. R. Porter and J. H. Rayner, 1984, A winter wheat crop simulation model without water or nutrient limitations, *Journal of Agricultural Science, Cambridge*, (1984), 102, 371-382

## Приложения

Приложен е компакт диск съдържащ описаните в работата софтуерни средства.

### Публикации на автора по темата на дисертацията

1. Ангелова, Р., Г. Желепов, В. Георгиев, 1988, Проучване влиянието на някои абиотични фактори върху динамиката на диапаузиране при гъсениците на ябълковия плодов червей, *Laspeyresia pomonella* L. (Tortricidea, Lep.), с помощта на ЕИТ”. Научни трудове на Висшия селскостопански институт - Пловдив, т. XXXIII, кн. 3, 1988 г.
2. Василева, М., А. Матеева, Д. Ковачева, В. Георгиев, 1989 г. Един модел на зависимостта “концентрация – ефект” за някои инсектициди спрямо ларви на колорадски бръмбар, Висш селскостопански институт – Пловдив, Научни трудове, т. XXXIV, кн. 2, 1989, Трета научна конференция с международно участие “Интензификация и екологизация на селскостопанското производство”.
3. Георгиев В., М. Василева, А. Матеева, 1991, Върху гостоприемниковата способност на люцерната, Първа национална конференция по ентомология, 28 - 30 октомври, 1991 г., София.
4. Георгиев В., М. Василева, А. Матеева, 2001, Характеристика и групиране на сортове картофи на основата на кластер анализ чрез специфични показатели. Юбилейна научна сесия: 50 години Добруджански земеделски институт - Генерал Тошево, 1 юни 2001 г., гр. Добрич
5. Георгиев В., Р. Ангелова, 1993, Имитационен модел на популация на люцернов листояд *Phytodecta fornicata* Brüg. (Coleoptera chrisomelidae), сп. "Растениевъдни науки", 1993, год XXX, N. 9-10, стр. 77-80.
6. Иванова, И., В. Георгиев, Д. Костова, 2001, Кластер-анализ на селекционни линии краставици в зависимост от степента на устойчивост спрямо краставично-мозаечния вирус (CMV), Юбилейна научна сесия “80 год. висше агро-

номическо образование в България”, Аграрен университет – Пловдив, Научни трудове, том XLVI, кн. 3.

7. Кирчев, Хр., В. Георгиев, М. Василева, 2001, Многогодишни тенденции в изменение на добива в двуполката пшеница-царевица в следствие на продължително минерално торене. I. Пшеница. Юбилейна научна сесия: 50 години Добруджански земеделски институт - Генерал Тошево, 1 юни 2001 г., гр. Добрич.
8. Кирчев, Хр., В. Георгиев, М. Василева, 2002, Многогодишни тенденции в изменението на добива в двуполката пшеница-царевица в следствие на продължително минерално торене. II. Царевица., Аграрен университет – Пловдив, 100 години от рождението на акад. П. Попов, Юбилейна научна конференция, Научни трудове, том. XLVII, кн. 1, 2002.
9. Лечов, Г., В. Георгиев, 1992, Модел на замърсяването на почвите с тежки метали в района на КЦМ - Пловдив, Как да решим екопроблемите на България - сборник доклади, Висша книжна школа "СИРИУС 4", В. Търново.
10. Михайлова, Е., В. Георгиев, В. Ройчев, И. Панчев, З. Велчев, 1996, Динамика на изменението на витамин С в стафида, Висш селскостопански институт - Пловдив, Научни трудове, т. XLI, 1996 г., стр. 27 - 34
11. Михайлова, Е., В. Георгиев, И. Панчев и З. Велчев, 1996, Приложение на цветните характеристики за прогнозиране на срока за съхранение на сушени гъби, Висш селскостопански институт - Пловдив, Научни трудове, т. XLII, 1996 г.
12. Babrikova, T, V. Georgiev, M. Vassileva, A. Mateeva, 1994, Correlation between the harmful and predatoty entomofauna in alfalfa, Зборник на трудови од XVIII, советовање за заштитата на рестенијата, Охрид, 9-10 XI 1993, година V, том V.
13. Ivanova, I., D. Kostova, V. Georgiev, 2002, Grouping of *Phaseolus vulgaris* cultivars and breeding lines according to their succceptibility to Cucumovirus CH 39-

BG on the basis of cluster analysis, First symposium on horticulture, 16-20 October, 2002, Ohrid, Republic of Macedonia.

14. Panayotov N., V. Gueorguiev, I. Ivanova, 2000, Characteristics and grouping of F<sub>1</sub> pepper (*Capsicum annuum L.*) hybrids on the basis of cluster analysis by morphological characteristics of fruits. Capsicum & eggplant newsletter. № 19, 2000, University of Turin, Italy, p 62-65.