CASE STUDIES FROM AN INTEGRATED MATHEMATICS AND SCIENCE COURSE

P.E. McNEIL
Dept. of Mathematics, Norfolk State University
Norfolk, VA 23504

Introduction

In a previous article, the essentials of an experimental, interdisciplinary mathematics and science course was outlined [1]. The course was developed and taught by a team of mathematics and science professors at Norfolk State University (NSU), under the auspices of the Virginia Collaborative for Excellence in the Preparation of Teachers (VCEPT), and funded by the National Science Foundation (NSF). The article gave brief descriptions of the investigatory modules developed for the course and the mathematical modeling approach to solving the problems posed in the modules. In this paper, we go into greater detail on some actual problems posed in the modules and the approaches that students used to investigate them. The problems and activities are grouped under headings called "studies" to indicate that they were compiled by studying and observing students in actual classroom situations. The course makes use of an introductory mathematical modeling text [2], however the majority of the course activities are generated from the written modules. Students are given a lengthy bibliography to assist in their investigations, some of the items from which are included in this paper's references [3-8].

The course attracts approximately twenty students per semester, mostly sophomores and some juniors, who use the course as a lower level math or science elective. The prerequisites for the course are two semesters of mathematics and two semesters of laboratory science. The mix of students in the course includes mathematics and natural science majors, as well as elementary and secondary education majors. This diversity of majors contributes to the interdisciplinary flavor that the course attempts to maintain.

Case Studies

<u>Study 1: Exponential and Logistical Curve Fitting</u> — At the outset of the course, students are invited into the world of mathematical modeling through a number of exercises involving curve fitting and linear regression. In a "greenhouse effect" exercise, students are given the following

The Journal of Mathematics and Science: Collaborative Explorations Volume 5 (2002) 59 - 73

data, which show the rise in the average temperature of the earth each year above that of the base year 1860 over a period of years beginning with the year 1880.

Year	**Average Temperature Rise of the Earth Above the 1860 Figure (°C)**
1880	0.01
1896	0.02
1900	0.03
1910	0.04
1920	0.06
1930	0.08
1940	0.10
1950	0.13
1960	0.18
1970	0.24
1980	0.32

Students are prompted to describe the devastation that would be associated with a rise in average temperature of 7°C above its 1860 value (melting of the polar ice caps, massive flooding, etc.), and then they are challenged to come up with a model of the data which will permit them to predict when the earth's temperature will reach that value should the trend in the chart continue. After some instruction on procedures for finding lines and curves of best fit, some students came up with the model:

$$y = (0.013301236)(1.033447189)^x$$

where x represents the years, zeroed at 1880, and y represents the average temperature above the 1860 value. At this point, students were challenged to solve the equation for x, given that $y = 7$. This presented an excellent opportunity for teaching the value and the utility of logarithms.

In another exercise, students are provided the data in the table below, which gives the distance of each of the planets from the sun and the planet's period of revolution around the sun.

Planet	**Distance, R, From the Sun (millions of km)**	**Period of Revolution, T, Around the Sun (days)**
Mercury	57.9	88
Venus	108.2	225
Earth	149.6	365
Mars	227.9	687
Jupiter	778.3	4,329
Saturn	1427	10,753
Uranus	2870	30,660
Neptune	4497	60,150
Pluto	5907	90,670

Students are then asked to come up with the polynomial equation of the form $y = ax^n$ that "best" models the data. In effect, they are being asked to discover or to verify the well-known Kepler's Third Law which establishes a relationship between a planet's distance from the sun and its period of revolution around the sun. The exercise produces some excitement when students see how closely they can come to deriving Kepler's formula: $T^2 = kR^3$, where T is the planet's period of revolution, R is its distance from the sun, and k is a multiple of the gravitational constant. In one instance, a group of students working cooperatively came up with the model,

$$y = 0.2001x^{1.4997},$$

but were hard pressed to show how closely the latter comes to expressing the law that the square of the period of revolution of a planet around the sun is directly proportional to the cube of its distance from the sun. We took this opportunity to review the basic laws of exponents, and then we had the students square both sides of the model equation to get the result:

$$y^2 = 0.04004x^{2.994}.$$

Now, they were able visually to make the connection with Kepler's formula, and they showed, through elementary error analysis, that their model for the given data varies from the Kepler formulation by less than 5%.

In one of the population growth exercises, students are given the following population data for the United States:

Year	**USA Population (millions)**
1850	23.2
1860	31.4
1870	38.6
1880	50.3
1890	62.9
1900	76.0
1910	92.0
1920	105.7
1930	122.7
1940	131.7
1950	150.7
1960	179.3
1970	203.2
1980	226.5
1990	252.0
2000	281.0

Next, they are asked to determine the periods in which the growth was approximately exponential and those in which the growth was approximately logistic. The results are predictably varied amongst the student groups. The most interesting part of this exercise occurs when students are asked to cite historical reasons for the rapid growth or the slowed growth in a given period. The approach in the majority of these exercises is admittedly empirical, but the approach lends itself to nudging students toward the theoretical aspects of modeling.

Study 2: Improving the Quality of Water — This section relates to a module that addresses some causes and remedies of water pollution in the Hampton Roads area. The unit begins with a lecture by a representative of the Chesapeake Bay Foundation who emphasizes the importance of oysters to a clean bay and who describes efforts by the foundation to increase the oyster population in the bay. At the end of the lecture, the representative presents a spreadsheet (see Appendix A, "Oysters in the Chesapeake Bay") that gives information on the oyster density, the fecundity, and the retention and mortality rates of oysters in seven tributaries around the state of Virginia. Given the following definitions and formulas, students are asked to complete the spreadsheet and to make projections concerning the future oyster population in the selected tributaries.

Oyster Density: Number of oysters per square meter of bottom
%Fert: The percentage of eggs fertilized %Fert = 0.49*(Density)^0.72
Fecundity: Number of eggs per female oyster
Area: Area of oyster habitat at the average density given in column 1 of the sheet
Offspring: Number of larvae produced: Female density × fecundity × %Fert × area
Survivors: Number of larvae left after 2 weeks; mortality rate is 99.5% first 2 weeks
%Retention: Percentage of oyster offspring retained for settlement
# of Settlers: Number of offspring that survive to settle in a suitable habitat
Juvenile Mortality: % Mortality of juveniles prior to reaching sexual maturity
New Oysters: Number of oysters reaching sexual maturity after one year

After computing the number and density of the oysters in each of the tributaries at the end of one year, students are now in position to use the new oyster density to project the number of oysters after a second year and for subsequent years. One of the students, a computer science major, became so enthused with this process that he wrote a fairly elegant program in *C++* to project the density of the oysters in each of the tributaries in any given year up to fifty years in the future! The exercise is an excellent one for gaining facility in the use of a spreadsheet. Also, students see the relevance of their projections relative to the future health of the Chesapeake Bay and its environs. A version of the completed chart after one year of propagation is in Appendix B ("Completed Oyster Chart").

<u>Study 3: Tracking the Spread of Diseases</u> — The activities described in this study result from a module entitled, "Epidemics and the Spread of Diseases." In this module, students are provided background and data on some of the dread diseases that plague our society, along with information on some of the epidemic outbreaks of the diseases. Initially, students are presented

with the standard Susceptibles-Infectives-Removals (SIR) model for the spread of diseases which produces a discrete-time collection of difference equations:

$$x_{n+1} - x_n = -kx_n y_n \tag{1}$$

$$y_{n+1} - y_n = kx_n y_n - ry_n \tag{2}$$

$$z_{n+1} - z_n = ry_n \tag{3}$$

where x_n denotes the number of susceptibles (people who are currently uninfected but who may become infected), y_n denotes the number of infectives (people who are infected and capable of spreading the infection), and z_n denotes the accumulated number of removals (persons who died from the disease or who recovered and are now immune) on day n of the epidemic. We make several simplifying assumptions, including: the three groups are mutually exclusive and comprise the entire population being studied; and, there are no births or deaths due to other causes during the epidemic. We further assume that the population size N is fixed throughout the epidemic so that $N = x_n + y_n + z_n$ for all n. The model is particularly useful because if the constants of proportionality, k and r, are known or can be approximated for a population, then values of x_n, y_n, and z_n can be generated on a spreadsheet from initial values. Students illustrate these notions in solving the following problem:

> Suppose that in a population of 1,000 people twenty people are initially immune from a disease and one person is initially infected. Generate values of x_n, y_n, and z_n if it is known that k = .0005 and r = .05. Display your results in a spreadsheet, and generate graphs of each function. Based on your spreadsheet and graphs, approximate answers to the following: How many days will it take for the susceptibles to be eradicated? On what day did the epidemic produce the largest number of infectives? Approximately how long does the epidemic last?

Their solutions typically produce charts and graphs similar to those illustrated in Appendix C ("Introductory Epidemics Problem") and Appendix D ("SIR Epidemic Curves"). In subsequent investigations, students attempt to model specific epidemics by fitting data to these "ideal" SIR curves. Their task, when given data on some epidemic, is to determine values for the parameters k and r in equations (1) – (3) that produce the SIR curves that "best fit" the given data.

Students get the opportunity to apply these notions in an investigation of the Ebola outbreak of 1995 in Zaire, Africa. They are given data in the form of a histogram, downloaded from the Centers of Disease Control website, that contain the number of deaths recorded daily during the 85-day epidemic (see Appendix E). Also, they are instructed that the incubation period for Ebola is between 2 and 21 days, that approximately 77% of infected people usually die from the disease, and that the population size is about N= 900.

Their first task is to generate a table for the removals, z_n. Then, taking the average incubation period to be say nine days, they are led to see that the number of deaths in any nine-day period during the epidemic should be roughly 77% of the number infected on the first day of the period. Using this information, they are able to come up with approximations for k and r for this epidemic. To test the accuracy of their approximations, they generate the theoretical SIR curves for their values of k and r, and then compare them with the given data by way of error analysis. The idea is to juggle the values of r and k until some acceptable level of error is acquired. Using this approach, one student group came up with the values k = .00013 and r = .04542, which yield pretty good results.

Study 4: Studying the Flow of Heat — The activities in this study introduce students to the world of thermodynamics by discussing the problems associated with heat flow in a variety of circumstances. Initially, students are introduced to Newton's Law of Cooling: the change in temperature of a substance is proportional to the difference between the temperature of the substance and room temperature, symbolically:

$$T_{n+1} - T_n = k\left(T_n - R\right). \tag{4}$$

Students then test the law by way of experiments in the physics laboratory involving the cool down rate of a hot liquid. The same experiment is modeled in the classroom using a Texas Instruments Calculator Based Laboratory (CBL) and curve fitting techniques. Their laboratory introduction to Newton's Law is contained in the following instructions relative to a cup of hot liquid:

1. Collect temperature data at two-minute intervals for at least fifty minutes.

2. Plot a graph of the temperature data over time. Describe the curve. Does your curve suggest an asymptote?

3. Test to see if Newton's law holds by graphing (T_{n+1} - T_n) vs (T_n - R). If Newton's Law does hold for your data, what kind of curve should you get? How can you approximate the constant of proportionality, *k*? (Hint: find the slope of a certain regression line.)

At this point, students are asked to verify that the general solution to the difference equation in (4) is:

$$T_n = R + (T_0 - R)(1+k)^n \tag{5}$$

where T_0 represents the original temperature of the liquid. Now, they are led to perform error analysis to determine how closely the data they collected in the cooling experiment approximates that generated by equation (5).

One student, a pre-service teacher, used these notions to develop a lesson plan that he adapted from a scenario he found in *The Mathematical Universe*, by William Dunham [9]. The scenario involves heating a potato ("Mr. Potato Head") in a microwave and then removing the potato prior to students' arrival in a classroom. The removal of the potato simulates the death of Mr. Potato Head and with some clues, students are required to engage in forensic science and estimate the time of death, i.e., the time at which the potato was removed from the microwave.

Study 5: Explorations in Human Genetics — In this final study, the activities are designed to introduce students to the elementary principles of genetics and the application of these principles to the study of certain genetic diseases and other genetically based phenomena. The genetics principles are introduced through individual examples using intuitive probability notions. In this introduction, students become familiar with basic genetic terms and concepts: gene, allele, genotype, phenotype, dominant/recessive traits, etc., and they get some experience in producing and interpreting Punnett squares. This approach extends in a natural way to discussions of population genetics and illustrations of the Hardy-Weinberg Principle. Students examine case studies involving genetic diseases, such as cystic fibrosis, Huntington's disease, Tay-Sach's disease, albinism, and sickle-cell anemia. In a typical example, students compute the probability that two parents who are carriers of the sickle cell anemia trait will produce a child who has the disease or who carries the trait. In another example, students consider a population in which one child in 400 has the sickle cell disease and are asked to find the genotypic frequencies of the normal and abnormal alleles relative to sickle cell anemia in this population. After computing these frequencies for the first filial generation of the population, they are prepared to discuss the question of whether the Hardy-Weinberg Principle appears to hold for the population.

Another activity involves students in an experiment comparing blood type frequencies on Norfolk State's campus with established blood type proportions in the United States and elsewhere. First, they conduct a survey to determine the blood group type of each member of the class. They list the number in each blood group and the blood group frequencies as percentages of the total class size. Next, using the table below, which shows the result of blood type testing among certain populations, they calculate the expected numbers in the blood groups based on the proportions listed in the table. Finally, they use the chi-square test with three degrees of freedom to determine if there is a significant difference between the expected and the observed numbers in the blood groups. Students are encouraged to increase the size of their sample to get better results.

Proportions of ABO Phenotypes in Selected Populations

Population	**A**	**B**	**AB**	**O**
Chinese	0.27	0.23	0.06	0.44
French	0.45	0.09	0.04	0.42
Indians	0.25	0.38	0.07	0.30
Nigerians	0.21	0.23	0.04	0.52
US Blacks	0.27	0.21	0.04	0.48
US Whites	0.41	0.10	0.07	0.52

As a part of their study of genetic diseases, students obtain first-hand information about the research efforts at the Eastern Virginia Medical School, Norfolk, VA, to isolate a diabetes gene and, after viewing a Public Broadcast System video, they discuss the ethical implications of the discovery of a breast cancer gene. Their final challenge is the gathering of information on the Internet about the Human Genome Project. A good deal of excitement resulted from the public announcement, while we were engaged in these studies, that scientists had completed the mapping of the entire human genetic system.

Evaluation and Conclusion

The evaluation of the course is based on three items: a) student participation in group activity; b) student performance on examinations covering basic mathematics and science skills; and, c) student evaluation questionnaires. Student evaluation surveys are conducted after the completion of each module, and one final questionnaire is administered at the end of the course. Similar final questionnaires are administered in three traditional sophomore/junior level courses in mathematics, biology, and chemistry, each taught by a member of our teaching team.

We find that students, after an initial period of adjustment, adapt very well to working in cooperative groups. In order to discourage students from depending on one or two persons in the group to do all of the work, we require a "division of labor" statement to be included with each of their submissions. Students soon realize that each of them must contribute his/her expertise in order for his/her group to successfully complete a module.

Relative to item b), we identify items on the final examination which appeared to deal with basic mathematics and science concepts and skills. We do the same in the three regular courses. We found that performance on these items in the interdisciplinary course is comparable to that in the three regular courses. Our conclusion is that the innovative, experimental elements in the course are not detrimental to the basic skill building that should accompany a mathematics or science course.

As a result of our evaluation surveys, we find that students in the interdisciplinary course display great enthusiasm for the course topics and methodology. They appreciate the relevance of our modules to real-world problems and issues, and they tend to be surprised and amazed at the interconnectedness of the disciplines. Moreover, we discern from the surveys a definite increase in students' perceived confidence in their ability to do mathematics and science and we observe, especially in the cooperative group activity, an increase in their readiness to tackle challenging problems.

The success that we have enjoyed in the activities outlined in these studies suggest to us that students with minimal preparation in mathematics and science can become involved in significant mathematics/science concepts. We believe that the secret to success in a course such as ours is first, to involve students in interesting and relevant problems and second, to begin at a concrete level at which they can easily function, gradually leading them to the conclusions and abstractions that we wish them to comprehend. It is gratifying to us that the course has become a popular one for students preparing to teach. Our hope is that when these students become masters in their own classrooms they will employ these methods with their own students. Gratifying also is the evidence, provided in this course, that both liberal arts majors and education majors can coexist and be mutually beneficial in the same classroom setting.

The success of our efforts has led to the approval of a follow-up course using the same format. We look forward to beginning this new endeavor during the Spring 2003 semester. ■

References

[1] P.E. McNeil, "A New Interdisciplinary Mathematics and Science Course," *The Journal of Mathematics and Science: Collaborative Explorations*, **2**(2) (1999) 27-32.

[2] J. Berry and K. Houston, *Mathematical Modeling*, Edward Arnold, London, 1995.

[3] T. Audesirk and G. Audesirk, *Biology: Life on Earth*, Prentice Hall, Upper Saddle River, NJ, 1996.

[4] N.G. Cooper (ed.), *The Human Genome Project: Deciphering the Blueprint of Humanity*, University Science Books, Mill Valley, CA, 1994.

[5] S.P. Gordon, et al., *Functioning in the Real World: A Precalculus Experience*, Addison-Wesley, Reading, MA, 1997.

[6] I.D. Huntley and J.G. James (eds.), *Mathematical Modeling: A Source Book of Case Studies*, Oxford University Press, New York, 1992.

[7] J. Trefil and R. Hazen, *The Sciences: An Integrated Approach*, John Wiley and Sons, Inc., New York, 1995.

[8] M.A. Tribe, I. Tallan, and M.R. Eraut, *Basic Biology Course: Book 12 Case Studies in Genetics, Unit 5 Aspects of Heredity*, Cambridge University Press, New York, 1978.

[9] W. Dunham, *The Mathematical Universe: An Alphabetical Journey through Great Proofs, Problems and Personalities*, John Wiley and Sons, Inc., New York, 1994.

APPENDIX A – Oysters in the Chesapeake Bay										
Area	**Oyster Density (Oyst per m^2**	**% Fert**	**Fecundity Eggs/Fem**	**Area m^2**	**Offspring**	**Survivors**	**% Retention**	**# Settlers**	**%Juv. Mort.**	**New Oysters**
James River	30.9	5.8	10^6	12147	1.1x10^10	5.4x10^7	90	4.9x10^7	99	5.4x10^7
Rappahannock River	1.7		10^6	12147			50		99	
Piankatank River	0.7		10^6	24294			90		99	
Great Wicomico River	36.6		10^6	44539			90			
Coan River	6.3		10^6	12147			75		99	
Yeocomico River	0.9		10^6	40490			75		99	
Tangier Sound	0.4		10^6	336067			25		99	
	Density of New Oysters (number of new/area)		**% Mortality of Adult Oysters (Fishing + Natural)**		**Adult Density Ater 1 year**		**Final Density (Adult + New)**		**Total Oysters After 1 Year**	
James River	40.3		65		10.8		51.1		620712	
Rappahannock River			50							
Piankatank River			40							
Great Wicomico River			50							
Coan River			50							
Yeocomico River			20							
Tangier Sound			5							

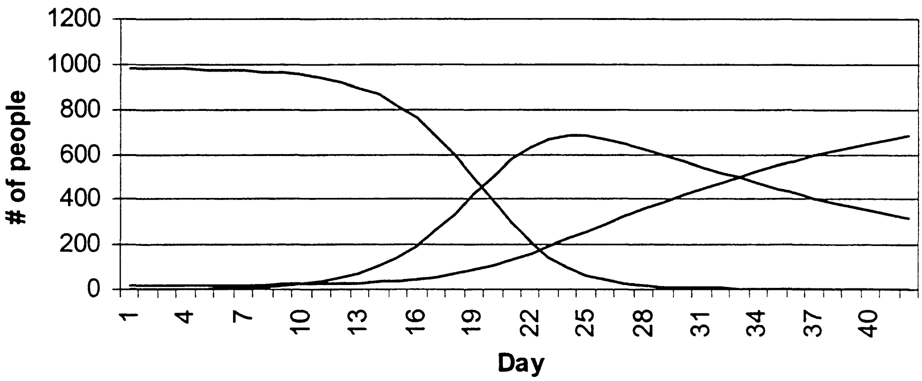
APPENDIX B - Completed Oyster Chart

Region	Oyster Density (Oyst per m^2	% Fert	Fecundity Eggs/Fem	Area m^2	Offspring	Survivors	% Retention	# Settlers	%Juv. Mort.	New Oysters
James River	30.9	5.8	1.0E+06	12147	1.1E+10	5.4E+07	90	4.9E+07	99	4.9E+05
Rappahannock River	1.7	0.7	1.0E+06	12147	7.4E+07	3.7E+05	50	1.9E+05	99	1.9E+03
Piankatank River	0.7	0.4	1.0E+06	24294	3.2E+07	1.6E+05	90	1.5E+05	99	1.5E+03
Great Wicomico River	36.6	6.5	1.0E+06	44539	5.3E+10	2.7E+08	90	2.4E+08	99	2.4E+06
Coan River	6.3	1.8	1.0E+06	12147	7.1E+08	3.5E+06	75	2.6E+06	99	2.6E+04
Yeocomico River	0.9	0.5	1.0E+06	40490	8.3E+07	4.1E+05	75	3.1E+05	99	3.1E+03
Tangier Sound	0.4	0.3	1.0E+06	336067	1.7E+08	8.5E+05	25	2.1E+05	99	2.1E+03
	Number of Original Oysters (Density H Area)		**% Mortality of Original Oysters (Fishing + Natural)**		**Remaining Original Oysters**		**Total Oysters After 1 Year**		**Oyster Density After 1 Year**	
James River	3.8E+05		65		1.3E+05		6.2E+05		51.1	
Rappahannock River	2.1E+04		50		1.0E+04		1.2E+04		1.0	
Piankatank River	1.7E+04		40		1.0E+04		1.2E+04		0.5	
Great Wicomico River	1.6E+06		50		8.2E+05		3.2E+06		72.2	
Coan River	7.7E+04		50		3.8E+04		6.5E+04		5.3	
Yeocomico River	3.6E+04		20		2.9E+04		3.2E+04		0.8	
Tangier Sound	1.3E+05		5		1.3E+05		1.3E+05		0.4	

APPENDIX C - Intro. Epidemics Problem

N = 1000	$x_0 = 979$	$y_0 = 1$	$z_0 = 20$
DAY	x_n(SUS)	y_n(INF)	z_n(REM)
0	979	1	20
1	979	1	20
2	978	2	20
3	977	3	20
4	975	4	20
5	973	6	21
6	970	9	21
7	966	13	21
8	960	18	22
9	951	26	23
10	939	37	24
11	921	53	26
12	897	74	29
13	864	104	32
14	819	144	38
15	760	195	45
16	686	260	55
17	597	336	68
18	497	419	84
19	392	502	105
20	294	576	130
21	209	632	159
22	143	666	191
23	96	680	224
24	63	679	258
25	42	666	292
26	28	647	325
27	19	624	358
28	13	598	389
29	9	572	419
30	6	546	447
31	5	521	475
32	3	496	501
33	3	472	526
34	2	449	549
35	2	427	572
36	1	406	593
37	1	386	613
38	1	367	632
39	1	349	651
40	1	331	668
41	0	315	685

APPENDIX D
SIR Epidemic Curves



Appendix E
Daily Ebola Deaths in the Zaire Epidemic of 1995