CWI Syllabi

CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone + 31 - 20 592 9333
Telefax + 31 - 20 592 4199
WWW page `http://www.cwi.nl/publications_bibl/`

CWI is the nationally funded Dutch institute for research in Mathematics and Computer Science.

Proceedings of the forty-second European Study Group with Industry

Amsterdam, The Netherlands, 18 - 22 February 2002

G.M. Hek (editor)

CWI Syllabus 51

Study groups with industry are meetings where people from industry join forces with mathematicians to jointly tackle industrial problems. The first such meeting was held in the sixties at the University of Oxford. Nowadays study groups with industry are organised in many countries, e.g. Australia, the USA and several European countries.

Since 1998 study groups with industry are also organised in The Netherlands. In 1998 a study group was hosted by the University of Leiden, in 1999 by the University of Eindhoven, in 2000 by Twente University, and in 2002 the study group was hosted by the CWI and the University of Amsterdam. The next study group will take place in Leiden again, in February 2003. For up-to-date information please follow the links `www.wiskgenoot.nl/swi` or `www.stw.nl`. The 42nd European Study Group with Industry was organised by the University of Amsterdam and the CWI, the National Research Institute for Mathematics and Computer Science in the Netherlands. The programme was supported by the section *Industriële en Toegepaste Wiskunde* (ITW) of the *Wiskundig Genootschap*, with financial support from the technology programme *Wiskunde Toegepast* of the technology foundation STW and the EW (Exact Sciences) programme of The Netherlands Organisation for Scientific Research (NWO), and with additional support from the CWI, the University of Amsterdam, and the European Consortium for Mathematics in Industry (ECMI).

Organising committee:
G.M. Hek
M.F.M. Nuyens
M.A. Peletier
R. Planqué
H. van der Ploeg
G. M. Terra

2000 Mathematics Subject Classification:
00B25, 35QXX, 60JXX, 92D40, 92C05, 31AXX, 90B80, 90C10.

ISBN 90 6196 516 0
NUGI-code: 811

pp. v – vi

Voorwoord

De positie van de wiskunde in Nederland heeft in de afgelopen maanden veel aandacht gekregen. Op 18 april 2002 presenteerden de wiskundige onderzoekscholen de nota *Nieuwe dimensies, ruimer bereik*, waarin gepleit wordt voor een versterking van het wiskundig onderzoek in Nederland, ondanks de teruglopende studentenaantallen. Die studentenaantallen vormen een bedreiging voor het voortbestaan van de negen wiskundeopleidingen die ons land rijk is. Dat de kwaliteit van de opleidingen over het algemeen goed is bleek uit het rapport van de onderwijsvisitatiecommissie voor de wiskunde dat op 25 juni 2002 werd aangeboden aan de voorzitter van de VSNU. Op dit moment voeren de universiteiten overleg over het Bachelor Convenant Natuurwetenschappen, waarbij de vraag aan de orde is of zelfstandige bacheloropleidingen in de wiskunde (en de natuurkunde en scheikunde) in stand zouden moeten blijven, of dat zij zouden moeten opgaan in bredere bachelorprogramma's.

Het NWO programma Wiskunde Toegepast heeft tot doel het toegepast wiskundig onderzoek te versterken. Dat gebeurt in de eerste plaats door de financiering van promotieonderzoek en post-doc posities. Maar het programma rekent het ook tot haar taak om activiteiten te ondersteunen die kunnen bijdragen aan een betere beeldvorming van het vakgebied. De Studiegroep Wiskunde met de Industrie is een jaarlijks terugkerend evenement dat een uitstekende bijdrage levert aan die beeldvorming en dan ook van harte wordt ondersteund door Wiskunde Toegepast. Elk jaar zijn het vooral jonge wiskundigen van de verschillende universiteiten, die het initiatief nemen voor de organisatie van de studiegroep. Daarmee laten zij zien dat de wiskunde een springlevend vak is, dat onmiddellijk toepasbaar is op concrete problemen uit de praktijk.

Ook in 2002 was de studiegroep weer een groot succes. U kunt zich hiervan overtuigen door de bijdragen te lezen in deze proceedings. Namens de programmacommissie van Wiskunde Toegepast wil ik mijn dank uitspreken aan de organisatoren en aan iedereen die actief heeft deelgenomen aan de studiegroep. Ik ben ervan overtuigd dat de uitkomsten van deze studiegroep zullen bijdragen aan het beeld van de wiskunde dat wij allen willen uitdragen: niet een stoffig middelbare-schoolvak, maar een uitdagend vakgebied, vooral

ook voor aankomende studenten. Die hebben wij nodig om de wiskundigen op te kunnen leiden waaraan onze samenleving dringend behoefte heeft.

Jos de Smit,
Voorzitter van de Programmacommissie
Wiskunde Toegepast

Contents

Summaries

For the 42nd European Study Group with Industry six problems were selected. The summaries of these problems are presented here.

1. Artis – cooling overheated fish

The Artis Zoo has a problem in its Aquarium and in the adjacent Zoological Museum.

Situation:
Part of the Aquarium is a corridor which contains so called mammoth tanks which measure 5 by 2.5 by 20 meters and are filled with water. Because of the tropical fish inside, the water should have a temperature of 24 degrees Celsius. As there is not much daylight, a dozen lamps have been placed just above the aquarium to make sure the fish inside are visible. However, these big lamps produce a lot of heat.

Problem:
When in summer time the outside temperature reaches 25 degrees, the temperature in the corridor containing the mammoth tanks increases up to 30 degrees. The water itself becomes 27 degrees, which is too much for the fish inside. In the neighbourhood of the lamps, the temperature rises to 40 degrees. Just under the roof sometimes temperatures of 60 degrees have been measured. The museum, adjacent to the aquarium, is suffering from the heat as well: a lot of objects (like stuffed animals) are no longer allowed to be displayed. There are some fans, but these can not do the job, especially not if doors are opened to fight the heat.

Question:
How can we change this situation with a minimum of cost and inconvenience for the visitors, employees and fish?

2. Philips Natlab – compression of audio-signals

Philips Natlab is looking for new ways to compress audio-signals.

Situation:
A new method for the digital representation of high quality audio signals has been introduced as an alternative to the widely used 16-bit recording format

used for CD signals. This new method produces 1-bit samples at a rate that typically is 64 times higher as for CD. For CD the samples are generated at a rate of 44.1 kHz.

The new method results in a raw audio data volume which is 4 times as large as for ordinary CD signals. New storage media provide a huge storage capacity, nevertheless it is beneficial to reduce the required storage capacity. Since the new format is intended for high quality audio signals, popular compression techniques that do change the signals, lossy coding, are unacceptable. This opens up a whole new research area of lossless coding of 1-bit audio signals.

Currently, two main methods have been developed for lossless coding of such 1-bit audio streams. The first, low complexity, scheme uses an adaptive prediction table with run-length residual signal coding. The latter, more elaborate, scheme uses linear prediction with arithmetic coding of the residual signal. In combination with buffering techniques, the methods realise typical average coding gains of 1.3 and 2.1, respectively.

Question:
Can we make compression methods that do better?

To evaluate new proposals, a few short excerpts of 1-bit audio signals will be made available, together with the coding gains achieved with the methods mentioned above.

3. Natuur & Techniek – diffusion of euro coins over Europe

On January 2002 twelve European countries have welcomed the euro as their new coin. The euro coins have a national side, which is different for every country. On top of that three mini states San Marino, the Vatican and Monaco have issued coins with their own image. So there are fifteen different euro coins that can be used in every one of those 15 countries. Therefore, unlike in the past, the coins will not be collected and brought back to their home country. The coins will slowly but surely be spreading over the 15 countries. This is the diffusion of the euro, or euro diffusion.

Because the Belgian and Dutch euro coins form only a fraction of the total number of euro coins, it is to be expected that foreign euros will replace most of the native euros. Interesting questions are: how quickly will the foreign euros take the place of the Dutch euros? How many French coins will we find in one year's time in our wallets? But other questions are also possible!

The diffusion of the euro can be studied in two ways, namely both practically and theoretically. The practical side consists of organizing measurements done by school classes and individuals in the Netherlands and Belgium.

For the theoretical side of the problem the science magazine Natuur & Techniek turned to the Study Group. Unlike the usual way, the problem is not fixed; during the week of the Study Group, the participants will be free to raise interesting questions and hopefully answer them as well. Natuur & Techniek is very interested in the discussion on this problem, and will use the results of the Study Group in an article. A preliminary article already appeared in the January 2002 issue.

The eurodiffusion project is an initiative by the Study Group and the science magazine Natuur & Techniek.

4. Phytocare – better advice to rose cultivators

Phytocare is looking for parameters to grow roses.

Situation:
In Agriculture the key point is optimizing the production at limited costs. For decennia already, one has tried for the best, and currently one cannot think of optimizing the production without a computer anymore. In present-day greenhouses the control of the inner climate is fully automated.

The inner climate could be held constant, but for the optimization of the production it is necessary to adapt the inner climate to the conditions outside the greenhouse. The importance of this can be illustrated by the effect of passing showers: if a rose grower does not anticipate with the, possibly sharp, temperature decrement that is due, this could mean a delay of one week for the production. Hence a swift and adequate reaction is of utmost importance.

Many theoretical models have been made that try to connect the climatic conditions to the resulting production of the weed. Unfortunately, growers do not profit much from the insight obtained by the present models. Most studies aim at one particular type of weed, but the characteristics of various types often differ significantly. Perhaps even more important: the characteristics are not constant throughout the year, whereas the present models account for them with fixed parameters.

Problem:
Phytocare thinks of a new approach. The idea is the following. The climate computer applies specified amounts of moist, light, nutrients, etc. to the plants in a greenhouse. At the same time, the inner climate is measured by the same computer: every 5 minutes the computer provides data on a.o. temperature, humidity and luminiscence in the greenhouse. One thus knows the previous living conditions of each plant with a precision of 5 minutes. One can also measure the production per plant per period from the plants themselves: for tomatoes, for example, the total weight of fruit procuced by a plant within a certain period can be counted. For roses, the weed most

of Phytocare's advises deal with, the production can be measured by the growth of branches per week; indeed, a branch can be harvest as soon as it reaches the required length to be sold. Measuring the production can only be done on a longer timescale. Usually the production is measured every week.

Using the measured climatic conditions in the greenhouse and the production per week, Phytocare would like to find species-specific parameters for the plants, for example by fitting them to the data. As explained before, one of the complications is the fact that the climate measurements are carried out every few minutes, whereas the production can be measured on a weekly basis only. With the parameter values obtained, the approach can be reversed again: are their rules of thumb to be given to the growers, by which their climate control can increase the production at reasonable costs? Different scenarios for advice could be calculated.

Questions:
Can the Study Group make a general model in order to facilitate Phytocare's advice to growers? Can growers be advised how to optimize their production using this model, fitted to the individual grower and his greenhouse by the above described approach? Or will Phytocare at least be able to find out under which conditions the photo-synthetic process of the plants is optimal (optimal production is closely related to optimal photo-synthesis)? Is the model accurate enough in order to calculate whether, for example, certain investments in the greenhouse (to optimize the production) will increase the production sufficiently to justify the investments?

5. Magma Design Automation – component placement on chips

The 'holey cheese' problem

One of the steps in the design process of chips is the positioning of every single component or 'cell' on the chip. The cells are mutually connected by wires. The wiring scheme is given, and in this phase of the design process the positioning of the various cells must be determined. Some (relatively few) cells have a prescribed position.

For the classical positioning problem one considers the chip as a two-dimensional plane; the cells are modelled by rectangles of various sizes. The positioning has to satisfy some conditions:

1 the cells must be placed within a certain rectangle (core area),
2 cells are not allowed to overlap,
3 the total wire length must be minimized.

For this problem many algorithmes are known, each one with its specific pros and cons.

The problem becomes more difficult when large parts of the core area are excluded from positioning, often caused by large, functional components that were placed beforehand (one could think of memory, or components that are designed by other companies). The remaining 'free area' within the core area is usually comparable to a cheese with holes, or 'holey cheese'. Obviously, the cells cannot be placed on the blockages, and this additional requirement makes the positioning problem significantly harder. For a typical 'holey cheese' the free area is strongly disconnected.

The current algorithmes of Magma DA suffice for more or less convex areas. However, in 'holey cheeses' they often end up in local minima that are far from optimal.

Question:
How can Magma DA find good solutions in case of holey cheeses?

6. NIOZ – reconstruction of sea-surface temperatures using fossil marine plankton

The predictive quality of climate models can be enhanced by incorporating information about temperatures from the past. A number of methods have been developed to determine the ancient temperatures of the upper ocean, and one of these is based on the use of deep sea micro fossils.

For many millions of years a large number of species of the invertebrate group planktic foraminifera have lived in the upper water level of the world's oceans. These organisms produce little shells of calcite (CaCO3) that function as a skeleton. Without changing the isotope ratio of oxygen in the dissolved CO2 in the ambient water, the water temperature has a direct influence on the isotope composition in the calcite shells of the plankton. If one would know the isotope composition of the ocean water, one could hence deduct the ocean water temperature from the isotope composition in the calcite shells. The isotope composition of the ocean water from ancient times is practically unknown, however, and, for theoretical reasons, it's not advisable to try to model it either. One way to get around this problem is to take more than one species of plankton:

The different species of plankton don't prefer the same ecological conditions. Some are adapted to live under colder conditions than others. One may hence in principle infer absolute temperature differences from plankton that has lived during the same time-span in the same water level in the same region: the isotope ratio of the water remains fairly constant during relativey short time-spans, but temperatures differ considerably both regionally and in time. Information could thus be obtained about both average temperatures in a given era, and of the variability of these temperatures. The variability is quite large, and therefore interesting to know. Other methods

to determine ancient sea water temperatures only obtain mean temperaturs, but the NIOZ is interested to obtain variances as well.

Results obtained so far indicate that the isotope composition in the calcite shells are not only determined by the temperature of the water in which the plankton lived, but also by other ecological influences. For instance, food availability also has its influences on the relative abundance of the different species of foraminifera. It is hence difficult to produce direct conclusions from the isotope composition data from the fossil calcite shells.

Questions:
The challenge now is to construct a model that encompasses the ecology of the organisms, that can still be used to infer the temperatures of the past from the isotope composition data found in the deep sea sedimentary record. It will be very easy to make this model extremely complicated, considering the number of side effects involved. How can we reduce this to a reasonable model that still simulates enough of the observed phenomena?

Participants

Andrei Abramian
Technische Universiteit Delft
andrei@dv.twi.tudelft.nl

Alexei Beliaev
Vrije Universiteit Amsterdam
beliaev@cs.vu.nl

Jan Bouwe van den Berg
University of Nottingham
jan.bouwe@nottingham.ac.uk

Franziska Bittner
Universiteit Utrecht
bittner@math.uu.nl

Piet van Blokland
Hogeschool Holland
pjvanblokland@chello.nl

Onno Bokhove
Universiteit Twente
o.bokhove@math.utwente.nl

Lorna Booth
Universiteit Utrecht
booth@math.uu.nl

Rachel Brouwer
CWI
rachel.brouwer@cwi.nl

Thijs Brouwer
Universiteit van Amsterdam
thijs.brouwer@student.uva.nl

Fons Bruekers
Philips Research
fons.bruekers@philips.com

Chris Budd
University of Bath
cjb@maths.bath.ac.uk

Adriaan van der Burgh
Technische Universiteit Delft
a.h.p.vanderburgh@its.tudelft.nl

Edi Cahyono
Universiteit Twente
e.cahyono@math.utwente.nl

Carlota Cuesta
Vrije Universiteit Amsterdam
carlota@cs.vu.nl

Natalia Davydova
Universiteit Utrecht
davydova@math.uu.nl

Dee Denteneer
Philips Research
dee.denteneer@philips.com

Arjen Doelman
Universiteit van Amsterdam
doelman@science.uva.nl

Johan Dubbeldam
Technische Universiteit Eindhoven
j.l.a.dubbeldam@tue.nl

Koen van Eijk
Magma Design Automation
koen@Magma-DA.com

Barbera van de Fliert
bvandefliert@yahoo.com

Philipp Getto
Universiteit Utrecht
getto@math.uu.nl

Dion Gijswijt
Universiteit van Amsterdam
gijswijt@science.uva.nl

Geertje Hek
Universiteit van Amsterdam
ghek@science.uva.nl

Piet Hemker
CWI
pieth@cwi.nl

Kirankumar Hiremath
Universiteit Twente
k.r.hiremath@math.utwente.nl

Michiel Hochstenbach
Universiteit Utrecht
hochsten@math.uu.nl

Bas van 't Hof
VRtech Computing
bas@vortech.nl

Ale Jan Homburg
Universiteit van Amsterdam
alejan@science.uva.nl

Ed Huijbregts
Magma Design Automation
ed@Magma-DA.com

Joost Hulshof
Vrije Universiteit Amsterdam
jhulshof@cs.vu.nl

Cor Hurkens
Technische Universiteit Eindhoven
wscor@win.tue.nl

David Iron
Universiteit van Amsterdam
diron@science.uva.nl

Lute Kamstra
CWI
lute.kamstra@cwi.nl

Yaroslav Kondratyuk
Universiteit Utrecht
kondratyuk@math.uu.nl

Ger Koole
Vrije Universiteit Amsterdam
koole@cs.vu.nl

Simon Kronemeijer
Universiteit van Amsterdam
kronemj@science.uva.nl

Andreas Kyprianou
Universiteit Utrecht
kypriano@math.uu.nl

Andre Leger
University of Bath
mapajpl@bath.ac.uk

Mervyn Lewis
CWI
mervyn.lewis@cwi.nl

Martijn van Manen
Universiteit Utrecht
manen@math.uu.nl

Jaap Molenaar
TUE/UT
j.molenaar1@tue.nl

Carolynne Montijn
CWI
carolynne.montijn@cwi.nl

Daniel Nitzpon
Universiteit van Amsterdam
nitzpon@gmx.net

Misja Nuyens
Universiteit van Amsterdam
mnuyens@science.uva.nl

Simona Orzan
CWI
simona@cwi.nl

Nick Ovenden
Technische Universiteit Eindhoven
n.c.ovenden@tue.nl

Jun Pang
CWI
jun.pang@cwi.nl

Simon van der Pal
Artis
vanderpal@artis.nl

Frank Peeters
NIOZ
peeters@nioz.nl

Mark Peletier
CWI
peletier@cwi.nl

Gemma Piella
CWI
piella@cwi.nl

Derk Pik
Universiteit Leiden
drpik@math.leidenuniv.nl

Bob Planqué
CWI
rplanque@cwi.nl

Harmen van der Ploeg
Universiteit van Amsterdam
hvdploeg@science.uva.nl

Iuliu Sorin Pop
Technische Universiteit Eindhoven
i.pop@tue.nl

Georg Prokert
Technische Universiteit Eindhoven
g.prokert@tue.nl

Marieke Quant
Katholieke Universiteit Brabant
quant@kub.nl

Vivi Rottschäfer
Universiteit Leiden
vivi@math.leidenuniv.nl

Jacques Rougemont
Heriot-Watt University
j.rougemont@ma.hw.ac.uk

Dick van der Sar
Phytocare
dick.vdsar@phytocare.nl

Jan A.M. Schreuder
Universiteit Twente
j.a.m.Schreuder@math.utwente.nl

Piet Sondervan
Artis

Bernadetta Tarigan
CWI
bernadetta.tarigan@cwi.nl

Guido Terra
UvA/NIOZ
gmterra@science.uva.nl

Paul Dario Toasa
University of Kaiserslautern
toasa@mathematik.uni-kl.de

Evgeny Verbitskiy
Eurandom
verbitskiy@eurandom.tue.nl

Erik Vermeulen
Natuur & Techniek
evermeulen@natutech.nl

JF Williams
University of Bath
j.f.williams@maths.bath.ac.uk

Djoko Wirosoetisno
Universiteit Twente
djoko@maths.ed.ac.uk

Dmitri Znamenski
Vrije Universiteit Amsterdam
dznamen@cs.vu.nl

pp. 11 – 20

CHAPTER 1

The Artis Problem

Chris Budd, Mark Peletier, Geertje Hek, David Iron, Andre Leger, Edi Cahyono, Ignacio Guerra, Paul Dario Toasa, JF Williams.

Abstract. The Artis aquarium has had difficulty maintaining a reasonable temperature in the recently install mammoth sea water tanks during the peak of summer. At this time the approximately 400 000 liters of water may be as much as 3 degrees Celsius too hot. This represents a considerable amount of energy to dissipate. Any solution to this problem must take into account the limited budget of the zoo, the heritage status of the building and the health of the fish in the tank. In this report, we analyse the major sources of energy entering and leaving the system. From this analysis, we find that the most effective method of reducing the water temperature is to increase the amount of evaporation from the system.

Keywords: energy balance, water temperature, conductive and radiative energy

1. Introduction

The Artis zoo has a number of aquaria in one heritage building, each with its particular environmental requirements. A recent addition of a large mammoth tropical sea water tank has introduced some problems. This tank is situated in a corridor and measures 5 by 2.5 by 20 meters. The ideal temperature for the tropical fish in the mammoth tank is 24 degrees Celsius. As there is not much daylight, a dozen lamps have been placed just above the aquarium to make sure the fish inside are visible. However, these big lamps produce a lot of heat. When in summer time the outside temperature reaches 25 degrees, the temperature in the corridor containing the mammoth tanks increases up to 30 degrees. The water itself becomes 27 degrees, which is too hot for the fish inside. In the neighbourhood of the lamps, the temperature rises to 40 degrees. Just under the roof sometimes temperatures of 60 degrees have been measured. This information is summarized in figure 2(a).

The problem presented to us is to reduce the temperature of the water in the mammoth tank at a minimal cost with the constraints of minimal modification to the heritage building.

Throughout this report, we speak about observations and assumptions. The observations were done during two visits to the aquaria, where we could,

FIGURE 1. Observation of the aquarium roof.

among other things, inspect the basins in the catacombs and condensation above them, the (lack of) ventilation, and the construction of the whole building including basins, tanks and even the roof. See figure 1. Throughout the report, we also use various material constants. These are all taken from [1].

2. Facts about the Problem

We will restrict our analysis to the largest tank, the mammoth tank. The water in this system, approximately 440 000 l, is divided in two volumes. The first volume is the tank in the public area. The second is the reservoir in the catacombs where the water is oxygenated. The water is pumped through the system with a circulation time of approximately 4 hours. The water of the total system can get up to 3 degrees Celsius too hot. Although the water in the system appears to be well mixed, there is a difference in temperature between the reservoir and the tank of about 0.1 degrees Celsius. Since the tank contains sea water, it is not possible to use a simple heat exchanger to cool the tank as any metals introduced into the water will release ions which is bad for the fish.

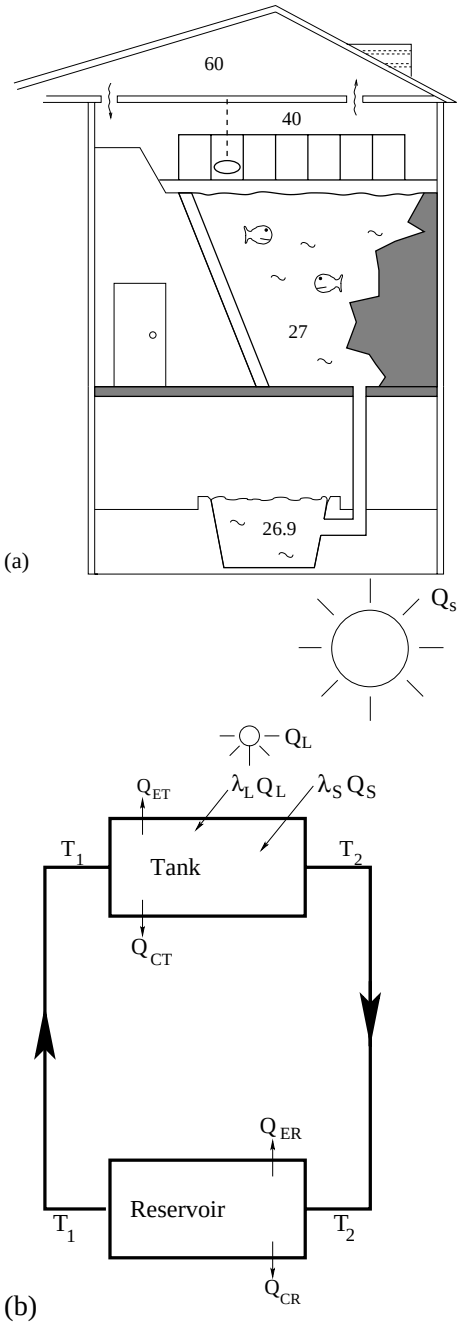


FIGURE 2. (a) Schematic picture of the mammoth tank and reservoir, (b) The energy inputs and outputs of the tank and reservoir.

3. Energy Balance

Here we consider the various energy sources and sinks in the system. To perform the energy balance, we will consider the mammoth system as two separate systems, i.e. the water in the tank and the water stored in the reservoir in the catacombs. In the tank the two main sources of energy are sunlight and the lamps heating the water. This energy is then either dissipated or stored in the water. The dissipation is conductive through the tank glass and walls. The water in the reservoir then dissipates more energy by evaporation and further conduction into the ground. We list each of these terms here.

Sources and Sinks	
Q_S	Solar Energy
Q_L	Energy from Lights
Q_{ET}	Evaporative Energy Loss from Tank
Q_{CT}	Conductive Energy Loss from Tank
Q_{ER}	Evaporative Energy Loss from Reservoir
Q_{CR}	Conductive Energy Loss from Reservoir
Constants and Variables	
λ_S	Percentage of Solar Energy Absorbed by Water
λ_L	Percentage of Light Energy Absorbed by Water
T_1	Temperature of Water in Tank
T_2	Temperature of Water in Reservoir
ΔT	Change in Temperature

Now we balance the energy in the tank and in the reservoir separately. First in the tank,

$$\lambda_S Q_S + \lambda_L Q_L = Q_{ET}(T_2) + Q_{CT}(T_2) + \theta \Delta T\,, \tag{1}$$

where θ is the amount of energy to change the temperature by 1 degree in the 4 hour cycle. The term $\theta\Delta T$ is the energy that is stored in the water as the temperature in the tank increases by ΔT. We calculate θ as

$$\theta = c\,\Phi\rho = \frac{4.18\times 10^3 \cdot 4\times 10^5}{4\cdot 3600} \sim 10^5 \mathrm{WK}^{-1}. \tag{2}$$

Here c is the specific heat of water, Φ is the flow rate, and ρ the density. The energy balance in the reservoir is given by

$$\theta\Delta T = Q_{CR}(T_1) + Q_{ER}(T_1)\,. \tag{3}$$

Note that the decrement in the water temperature in the reservoir is equal to the increment in the tank. This decrement corresponds to the energy loss by conduction and evaporation. The situation is illustrated in figure 2(b). We may make some simplifications from the observations made at the aquarium. The air in the space above the tank appeared to be trapped and at 100 percent humidity, thus the evaporative energy losses in the tank system are

negligible. The temperature change between the two systems, $\Delta T = T_1 - T_2$, is approximately 0.1 degrees Celsius, and thus we may take $T_1 \sim T_2 \sim T$. Under these assumptions, (1) and (3) may be simplified to

$$\lambda_S Q_S + \lambda_L Q_L = Q_{CT}(T) + Q_{CR}(T) + Q_{ER}(T)\,. \tag{4}$$

4. Detailed Analysis

In this section we will examine each term in (4) and find estimates of each one.

4.1. Energy Inputs. As very little convection was observed, we will assume that most of the energy enters the system by radiation. The total energy used by the lamps is approximately 10 kW. Since the lights are incandescent and very inefficient we will assume that all of this energy is transfered to the water. To find the energy added to the water from solar radiation, we use Stefan-Boltzmann law for black body radiation. In bad cases the temperature above the roof can reach 50 degrees Celsius or more and the water may be 25 - 27 degrees Celsius. We are finding the energy flux through a plane and thus divide the total energy by 2 as half will be transmitted and half will be reflected up. Finally, we estimate $\lambda_S \sim 1/2$. Thus, we have the following estimates for the energy inputs:

Lights

$$Q_L \sim 10\text{kW total}\,, \tag{5}$$

$$\lambda_L \sim 1\,, \text{estimate.} \tag{6}$$

Sun

$$Q_S \sim \frac{\sigma}{2}(T_{\text{roof}}^4 - T^4)A \sim 10\text{kW}\,, \tag{7}$$

$$\lambda_S \sim 0.5\,,\ \text{estimate.} \tag{8}$$

Here $\sigma = 5.6 \times 10^{-8}\ \text{Wm}^{-2}\text{K}^{-4}$ is the Stefan-Bolztmann constant, T_{roof} is the temperature just under the roof, for which we took $T = 323$ K in this estimate, and $A = 5 \times 20\ \text{m}^2$ is the area of the top of the tank. Furthermore, we took $T = 298$ K. For really hot days, when $T_{\text{roof}} = 333$ K (60 degrees C) and $T = 300$ K (27 degrees C), Q_S reaches about 20 kW.

Note here, that the order of magnitude of energy input by the lights and by the sun on warm days is the same. Even if we would have $\lambda_S \sim 1$, this is still the case.

4.2. Conductive Energy Losses. We will now estimate the energy loss in the system due to conduction. In the tank, we will assume that all of the energy is lost through the glass wall of the tank, as the back of the tank is made of thick rock and is a much better insulator. We will also

assume that the energy is convected away from the glass ideally. Under these assumptions, we get

$$Q_{CT} = \mu_T (T - T_{\text{ambient}}) , \tag{9}$$

where μ_T is the total thermal conductivity of the glass wall in the tank and is given by

$$\mu_T = k \frac{A}{d} . \tag{10}$$

Here, A is the area of the glass, d is the thickness of the glass and k is the thermal conductivity of glass. When we substitute in these constants we find that $Q_{CT} = 0.93 \times 50/0.1 \sim 0.5$ kW.

The estimation of the energy loss in the reservoir is more complicated, as the heat will flow into the ground and we may not assume that the energy is being removed ideally at the outer surface of the reservoir. To simplify the calculations, we will assume that the reservoir is hemispherical in shape. Although this will overestimate the heat loss to the ground, the result will be of the same order of magnitude as the actual losses. To find the thermal conductivity of such a reservoir, we assume that the temperature profile in the soil outside of the reservoir is radially symmetric and given by

$$T_S(r) = \frac{T - T_{\text{soil}}}{r} R + T_{\text{soil}} \quad \text{for } r > R. \tag{11}$$

Here, T is the temperature of the water, T_{soil} is the temperature of the soil far from the reservoir, r is the distance from the centre of the reservoir and R is the radius of the reservoir. Since the approximate dimension of the reservoir is 15 meters by 35 meters, we approximate R by 10 meters. Moreover, we approximate T_{soil} by 16 degrees Celsius. The conductive loss of the reservoir into the ground is then given by

$$\begin{aligned} Q_{CR} &= -k(2\pi R^2) \left. \frac{\partial T_S}{\partial r} \right|_{r=R} , \\ &= 2\pi R k (T - T_{\text{soil}}) . \end{aligned} \tag{12}$$

In other words, the thermal conductivity μ_R of the reservoir is $2\pi Rk$. Substituting all values into (12), we find that on a hot summer day, so for $T = 27$ degrees Celsius, the conductive loss in the reservoir is approximately 0.5 kW.

4.3. Evaporative Energy Losses. We now calculate the amount of energy removed from the system due to evaporation. We first make a few assumptions based on observations of the aquarium. We will assume that all of the evaporation occurs in the reservoir, since we observed much condensation near the air outlet in the catacombs, and no condensation above the tank. Moreover, the air just above the tank has the same temperature as the water inside the tank, and the area of the tank is much smaller than the area of the reservoir. We also assume that air enters the reservoir at 80%

saturation (typical conditions for Amsterdam in mid-summer) and then is fully saturated and leaves, via the vent, at 100% saturation. Under these assumptions, the evaporative loss is given by

$$Q_{ER} = \frac{\rho A V L}{p_{\text{atmos}}}(pp_{\text{inside}} - pp_{\text{outside}})\,, \tag{13}$$

where pp are the partial pressures of water vapour inside the catacombes and outside the building, ρ is the density of moist air (~ 1.3 kg/m^3), p_{atmos} is the atmospheric pressure (~ 100 kPa), L is the latent heat of water ($\sim 22.6\times 10^5$ J/kg), A is the area of the vent (~ 0.04m^2) and V is the velocity of air through the vent (and above the water). Since we have 100% saturation inside, we have $pp_{\text{inside}} = sp(T)$. Here sp is the saturation pressure at temperature T, and the air temperature is approximated by the water temperature. Outside, saturation is again expected to be 80% on average on a hot summer day, so $pp_{\text{outside}} = 0.8 \times sp(T_{\text{ambient}})$. In the summer, approximately 2 cubic meters of distilled water per week must be added to the system to maintain the volume. This is equivalent to approximately 6 kW of evaporation. Plugging this into (13), with $T = 27$ and an estimate of $T_{\text{amb}} = 20$ for the outside temperature averaged over a full hot day in summer, results in an air velocity of 2.4 m/s.

4.4. Summary of Energy Balance. Before proceeding to the full energy balance, we may compare one of our calculations with observed results. We will calculate the change of temperature in the reservoir and compare this to the observed change which is approximately 0.1 degrees Celsius. Recall, the temperature change in the reservoir is given by

$$\theta \Delta T = Q_{CR} + Q_{ER}\,, \tag{14}$$

where $\theta \sim 10^5 \text{WK}^{-1}$ is the amount of energy to change the temperature by 1 degree in the 4 hour cycle, Q_{ER} is the evaporative loss which is observed to be about 6 kW, and Q_{CR} is the convective loss which was calculated to be about 0.5 kW. Substituting this results in a value of $\Delta T = 0.065$ degrees Celsius, which is the right order of magnitude. We may now approximate $T_1 \sim T_2 \sim T$ indeed with some confidence. Substituting (13), (5), (7), (9) and (12) into (4) results in the following equation relating the temperature of the water with the ambient temperature, the heating effect of the sun (via the roof's temperature) and the velocity of air through the vent in the reservoir,

$$\begin{aligned}\tfrac{\sigma}{4}(T_{\text{roof}}^4 - T^4)A + 10 = \mu_T(T - T_{\text{amb}}) + \tfrac{\rho A V L}{p_{\text{atmos}}}(sp(T) - 0.8sp(T_{\text{amb}})) \\ +\mu_R(T - T_{\text{soil}})\,.\end{aligned} \tag{15}$$

We now approximate $sp(T)$ following estimates in [2], fitted to the values $T = 293$ and $T = 303$. This yields

$$sp(T) = 0.16 \times 10^1 2 \cdot \exp(-5.3 \times 10^3/T).$$

Roughly, this gives pressure values of 3 - 4 kPa in the range of interest. Using this estimate we provide graphs of relation (15) in figures 3, 4, 5 and 6.

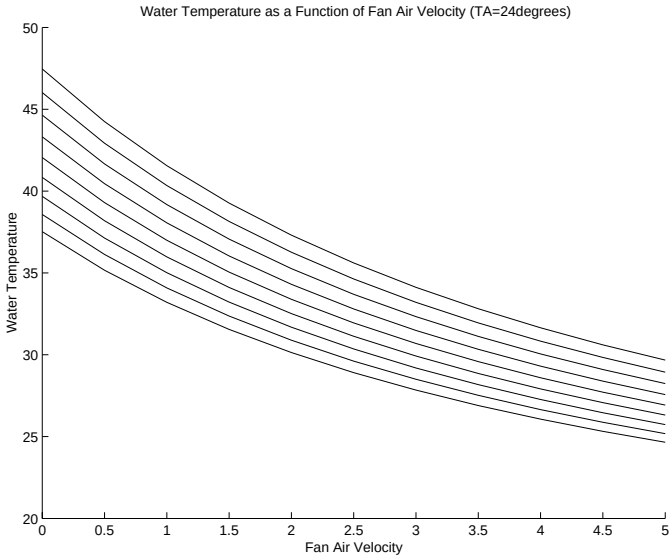


FIGURE 3. Water temperature versus fan air velocity. Higher lines correspond to higher roof temperature by the sun.

5. Conclusions

The above figures may be used as an indication of the relative influences on the water temperature of fan air velocity, ambient temperature and 'solar temperature', i.e. temperature just under the roof caused by the sun. Figures 4 and 5 show the dependence of the water temperature on the ambient and solar temperatures for fixed fan air velocities. The lines in these figures have slopes between 0.1 and 0.4, so the dependence is not very strong. In figure 3 however, the lines have slope ~ -2 in the regime around the current fan air velocity. For fixed solar and ambient temperature, the water temperature decreases by about 3 degrees if the fan air velocity is doubled from 2.4 to 4.8 m/s.

Therefore we conclude from the figures that the largest gains may be obtained by increasing the flow of air through the reservoir. This is relatively inexpensive and should not interfere with the appearance of the building. However, it must be noted that in the construction of relation (15), it was

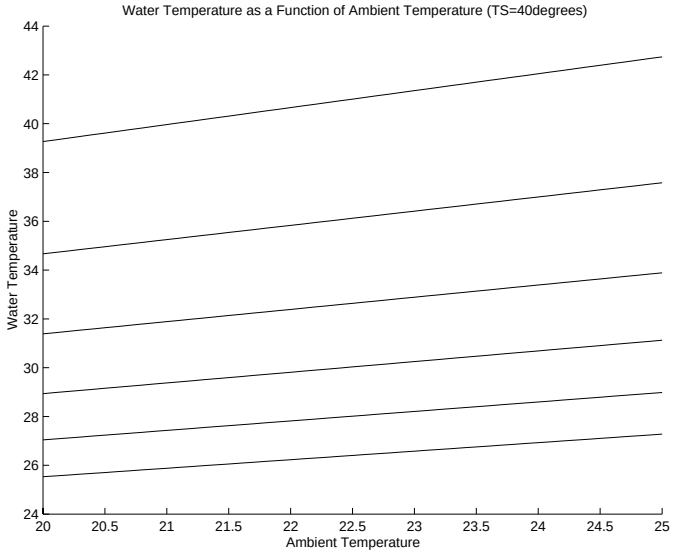


FIGURE 4. Water temperature versus ambient temperature. Lower lines correspond to increasing fan air velocity.

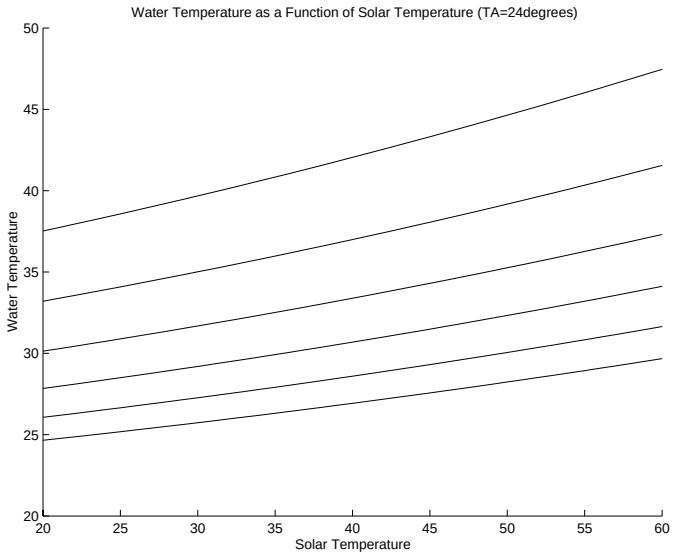


FIGURE 5. Water temperature versus solar temperature. Lower lines correspond to increasing fan air velocity.

assumed that an increase in the fan velocity will be proportional to the increase in water vapour leaving the system. This will not be true unless some care is taken. The area around the fan is very leaky and unless this

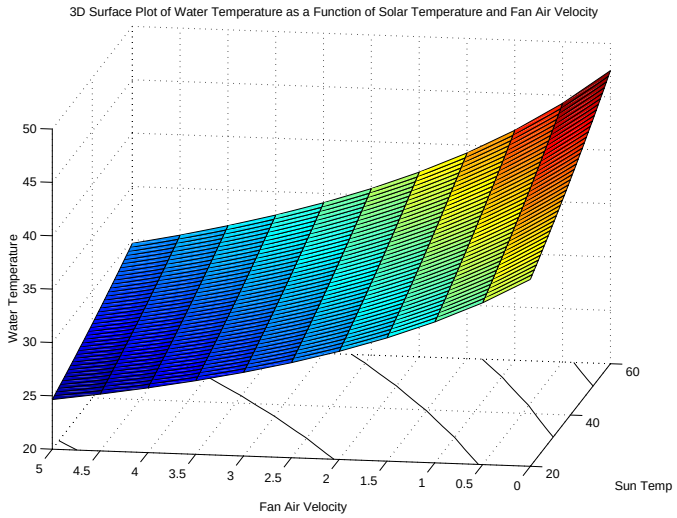


FIGURE 6. Overall picture.

is addressed, increasing the fan velocity will only draw outside air from the source near the fan. This will do nothing to increase evaporation. We also note that even though figure 5 suggest that reducing the amount of solar radiation will have a minimal effect on the temperature of the water, a reduction in the influx of solar energy could be achieved quite cheaply by reflective blinds or the growth of ivy and should thus also be considered.

Bibliography

[1] *Binas*, Wolters Noordhoff, Groningen, 1986.

[2] Education and outreach provided by the Planetary Atmospheres Node, *Encyclopedia of Standard Planetary Information, Formulas, and Constants*, `http://atmos.nmsu.edu/education_and_outreach/encyclopedia/sat_vapor_pressure.htm`

pp. 21 – 39

CHAPTER 2

On Lossless Compression of 1-bit Audio Signals

Franziska Bittner, Dee Denteneer, Simon Kronemeijer, Misja Nuyens, Simona Orzan, Jacques Rougemont, Evgeny Verbitskiy, Dmitri Znamenski.

Abstract. In this paper we consider the problem of lossless compression of 1-bit audio signals. We study the properties of the existing solution proposed in [**5, 6**]. We also discuss possible improvements. Other methods have been cosidered, and the results are reported.

Keywords:Audio compression, linear prediction, Markov predictors, boosting, machine learning.

1. Introduction

Lossless compression of audio signals is an active area of research, which already has a wide range of practical applications such as Compact Disks (CD's) and Digital Versatile Disks (DVD's). The problem posed to the Study Group by the Philips Research Laboratories comprised two different goals. Firstly, Philips is interested in the highest possible compression ratio which can be achieved without, more or less, any restriction to the complexity of algorithms, computing power required, etc. Thus the first part of the problem has a predominantly theoretical flavor. It should indicate the limits of compression techniques for audio data in the form of binary sequences. The second part of the problem is more practical. It concerns the evaluation of the current compression technique described in [**5, 6**]. An interesting question here is whether the proposed algorithm is optimal in some sense, and whether any improvements are possible.

There is one important practical aspect to keep in mind. If a new superior compression algorithm is proposed within the class of models, currently accepted and implemented as hardware (*encoders*/*decoders*), and one would like to implement this algorithm, then only a relatively small number of *encoders* should be upgraded. The 'old' *decoders* (such as CD players) would still be capable of reproducing the audio signal.

Philips supplied us with 4 generic audio sequences and with the corresponding compression ratios achieved by their coding techniques. We will refer to these samples as samples A, B, C and D. The first sample A is considered to be 'easy', the last sequence D is considered to be 'difficult'. The

remaining sequences B and C are of average complexity. The efficiency of the compression algorithm is measured by its *gain* R,

$$R = \frac{\text{length of the original sequence in bits}}{\text{length of the compressed sequence in bits}}.$$

For reference, the Philips algorithm achieves compression gains of around 3 for sample A, 2.4–2.6 for samples B and C respectively, and 2.2 for sample D.

This paper is organised in the following way. In section 2, we describe a general approach to data compression based on statistical modelling. We also describe in more detail the algorithm proposed by Philips, and formulate the criterion for the optimal predictor within the class of linear predictors. In section 3 we discuss the efficiency of the current linear predictor. We discuss ways of improving Markov predictors in section 4. In section 5, we propose ways of improving the current linear predictor. Especially schemes based on weighted least squares seem to be very promising.

2. Statistical modelling and prediction

A large number of compression methods is known and used in practice. These include among others well-known algorithms such as Lempel–Ziv, Huffman, or arithmetic coding. We will refer to this type of compression techniques as entropy coding. Given a binary sequence of length N, an efficient entropy coder will produce an output of approximately Nh bits, where h is the *entropy* given by

$$h = -p \log_2 p - (1-p) \log_2 (1-p),$$

where p is the probability of observing a 0 in the source sequence. From this we can see that binary sequences in which one symbol occurs more often than the other – 0 say, and hence $p \gg 1-p$ – will be compressed efficiently.

It is known, however, that applying entropy coding to audio signals is not very efficient due to the presence of long-time correlations in the signal. These should be exploited in order to gain efficiency. Therefore, a preprocessing step is required, which eliminates the statistical dependencies, and leads to an almost uncorrelated source with one dominating symbol. Then standard entropy coding techniques can successfully be applied.

Suppose $\{x_i\}$, $i = 1, \ldots, N$, is a binary sequence we want to compress. A typical approach to the development of such a preprocessing stage would be the design of a scheme that tries to predict the next symbol of the sequence based on k previous symbols

$$\hat{x}_n = F(x_{n-1}, \ldots, x_{n-k}). \tag{1}$$

Next we define a new sequence $\{y_i\}$ as follows. If the prediction is successful, i.e. $x_n = \hat{x}_n$, then we let $y_n = 0$, otherwise $y_n = 1$. It is also clear that

given the parameters of the predictor F, the first k bits $(x_1, \dots, x_k)$, and the values $\{y_i\}$, $i = k+1, \dots, N$, we can reconstruct the original sequence, $\{x_i\}$, $i = k+1, \dots, N$. In practice, we would like to have N much larger than k.

If we would be able to design a predictor F which makes only very few mistakes, then 0 will be a dominating symbol in the sequence $\{y_i\}$, and we should expect a good compression gain for this sequence by entropy coding. At the same time, the decoder must be supplied with the parameters of the predictor F. We will refer to the storage space in bits, required for these parameters, as the *overhead*. Clearly, the overhead should not be too large. The gain of such a coding scheme is

$$R = \frac{\text{length of the original sequence}}{\text{overhead} + k + (N-k)h_y},$$

where h_y is the entropy of the sequence $\{y_i\}$. Hence, in this setup, the problem of efficient audio compression is reduced to the design of a prediction scheme with a minimal value of overhead $+ Nh_y$. Here we used our assumption that k is much smaller than N.

When designing the prediction scheme for a sequence $\{x_i\}$, $i = 1, \dots, N$, we are allowed to use the whole sequence. For example, even an uncompressed piece of music of just a few minutes long takes up several GigaBits of storage space. Hence, N can be quite large. On the other hand, we would like to be able to start decoding (playing music) from a more or less arbitrary position. So in practice predictors are designed for much shorter sequences (frames) of length $N \approx 40000$. In this way possible problems arising from the non-stationarity of the audio signal are avoided; we can expect a single predictor to perform reasonably well on the whole frame.

2.1. Markov predictors. Let $k \in \mathbb{N}$ be the length of words on which the prediction will be based. The state space of our Markov chain is the set $\mathcal{S}^{(k)} = \{-1, 1\}^k \approx \{1, \dots, 2^k\}$ of words of length k on a two-symbol alphabet[1].

The Markov chain will make transitions between a word $\pi = x_1 \cdots x_k$ and the words $\pi^{\pm} = x_2 \cdots x_k k_{k+1}$ for $x_{k+1} = \pm 1$. Figure 1 shows the graph of the Markov chain for $k = 3$.

Each edge of the graph will be assigned the empirical transition probability measured from the data set. Namely for each word $\pi = x_1 \cdots x_k$, we compute

$$\begin{aligned} U^{(k)}(\pi) &= \text{number of times the word } x_1 \cdots x_k, +1 \text{ is found in the data ,} \\ D^{(k)}(\pi) &= \text{number of times the word } x_1 \cdots x_k, -1 \text{ is found in the data .} \end{aligned}$$

[1]In this paper we are dealing with binary data. Everywhere below we will be using equivalent representations of the data in alphabets {0,1} and {-1,1} freely, without announcing it explicitly. In every case, however, it will be clear which representation we are using.

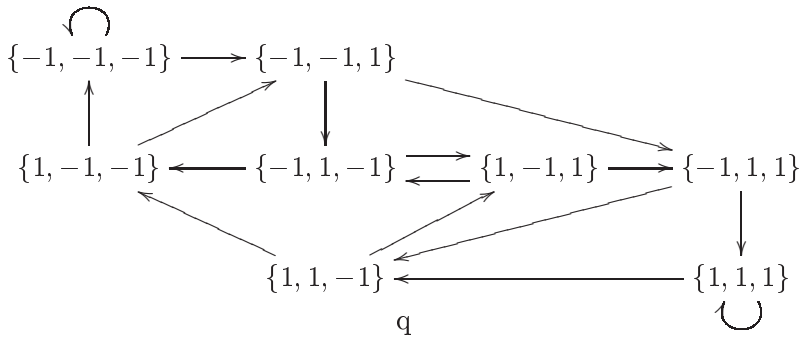


FIGURE 1. Graph of the Markov chain on $\mathcal{S}^{(3)}$

The empirical transition probabilities are given by

$$\begin{aligned} \mathbf{P}^{(k)}\big(\pi \to \pi^+\big) &= \frac{U^{(k)}(\pi)}{U^{(k)}(\pi) + D^{(k)}(\pi)} \,, \\ \mathbf{P}^{(k)}\big(\pi \to \pi^-\big) &= 1 - \mathbf{P}^{(k)}\big(\pi \to \pi^+\big) \,. \end{aligned} \tag{2}$$

These probabilities are not needed in practice, we only provide them for a complete definition of the Markov chain.

The Markov predictor of order k, $\widehat{x}_n = p^{(k)}(\pi)$ on the word $\pi = x_{n-k} \cdots x_{n-1}$ is defined by

$$p^{(k)}(\pi) = \begin{cases} +1 \,, & \text{if } U^{(k)}(\pi) \geq D^{(k)}(\pi) \,, \\ -1 \,, & \text{if } U^{(k)}(\pi) < D^{(k)}(\pi) \,. \end{cases} \tag{3}$$

Hence, in this way we predict the most probable symbol to follow π.

The overhead of the Markov predictor is 2^k bits. Hence, practical application of Markov predictors is feasible only for relatively small k's.

2.2. Linear predictors. Linear prediction is amongst the most successful tools in signal processing. In data compression, schemes based on linear prediction show good compression gains for various data types, e.g. speech. Linear predictors, in general, are trying to predict the next observation by a linear combination of several previous values:

$$\hat{x}_n = \beta_0 + \beta_1 x_{n-1} + \beta_2 x_{n-2} + \ldots + \beta_k x_{n-k},$$

Since we are dealing with binary data, $x_n \in \{-1, 1\}$, a natural way to incorporate this into our predictor is to consider the following predictors

$$\hat{x}_n = \operatorname{sign}\big(\beta_0 + \beta_1 x_{n-1} + \beta_2 x_{n-2} + \ldots + \beta_k x_{n-k}\big),$$

where $\operatorname{sign}(x) = 1$ for $x \geq 0$, and -1 otherwise.

The design of an optimal linear predictor hence consists of the selection of a vector of parameters $\beta = (\beta_0, \ldots, \beta_k) \in \mathbb{R}^{k+1}$ in a such way that the number of instances where $x_n \neq \hat{x}_n$ is minimal. The overhead of a linear

predictor is $(k+1)M$, where M is the number of bits used to represent a real number. Hence, the overhead is growing only linearly with k. This is a great advantage of linear predictors over Markov predictors, and allows them to go to much higher values of k. It is also clear that for the same order k, Markov predictors are at least as good as linear predictors.

2.2.1. *Philips predictor.* In [**5, 6**], a linear predictor of the following form has been used

$$\hat{x}_n = \operatorname{sign}\bigl(\beta_1 x_{n-1} + \beta_2 x_{n-2} + \ldots + \beta_k x_{n-k}\bigr), \tag{4}$$

where the coefficients $\beta_1, \ldots, \beta_k$ have been selected to minimize the following expression

$$\sum_{n=k+1}^{N} \bigl|x_n - (\beta_1 x_{n-1} + \beta_2 x_{n-2} + \ldots + \beta_k x_{n-k})\bigr|^2 \to \min. \tag{5}$$

The problem (5) is a standard least squares problem, which admits an efficient practical solution. This simple approach produces a predictor of a remarkable quality. An optimal value of the order k for such a linear predictor has also been investigated in [**5, 6**]. In fact, optimal k may vary and depends on a particular frame, but typically $k = 128$ gives good results. Increasing k does lead to a better predictor, but the corresponding growth of the overhead is not compensated by the gain in the quality of the prediction.

2.2.2. *Optimal linear predictor.* In fact, solving problem (5) gives only an approximation of the *optimal* linear predictor. An optimal predictor minimizes the number of errors, i.e. instances when $x_n \neq \hat{x}_n$. We can reformulate the criterion for the optimal predictor (4) in the following way: coefficients $(\beta_1, \ldots, \beta_n)$ of the optimal linear predictor are such that in the system of inequalities

$$\begin{array}{cc} \beta_1 x_{k+1} x_k + \beta_2 x_{k+1} x_{k-1} + \ \ldots \ + \beta_k x_{k+1} x_1 & > 0 \\ \vdots & \vdots \\ \beta_1 x_N x_{N-1} + \beta_2 x_N x_{N-2} \ + \ \ldots \ + \beta_k x_N x_{N-k} & > 0 \end{array}$$

the number of valid inequalities is *maximal*.

Let us denote by A the matrix with elements $x_n x_{n-j}$, $n = k+1, \ldots, N$, $j = 1, \ldots, k$. Hence the problem of finding the optimal linear predictor is equivalent to finding a vector β such that a vector $A\beta$ has a maximal number of positive coordinates.

In general, for a given matrix A it is quite easy to check whether there exists a vector β such that *all* the coordinates of $A\beta$ are positive. If such a vector exists, then we say that A defines a *feasible* system of inequalities. This is a best possible case, because the corresponding linear predictor will not make a single error. To check whether A is indeed feasible, we can use

standard methods of Linear Programming, which give an answer in polynomial time. However, one should not expect that the matrix A obtained from the data will lead to a feasible system of inequalities. The problem of finding the optimal linear predictor then is equivalent to the problem of finding the maximal feasible subsystem, i.e. a matrix A', made out of rows of A, which has maximal dimension and still gives a feasible system of inequalities. This problem is known to be NP hard, but also there is no polynomial approximation, see [**1, 2**]. However, there are several heuristic methods, which are known to perform relatively well. Unfortunately, we were not able to pursue this idea further. In the next section however, we will discuss briefly how close the linear predictor, given by (5), is to the optimal linear predictor.

3. Comparing predictors

In the previous section we have seen that the Philips predictor is, in principle, different from the optimal linear predictor. Nevertheless, the Philips predictor performs remarkably well: on average, 1 error for every 10 predictions made. This suggests that maybe the Philips predictor is not that far away from the optimal one. However, as was mentioned above, design of an optimal predictor is a known NP hard problem. On the other hand, we can try to compare Philips predictor with a Markov predictor.

For every pattern, the linear predictor makes at least as many mistakes as the Markov predictor. If it makes more, these extra errors are a measure for the quality of our predictor. We can calculate these numbers for our data. We should not make k too large; then almost every pattern would only occur once at most, and the Markov predictor would make no errors. This is not what we want; for each pattern, we should have a reasonable number of occurrences, or none.

It is useful to develop some terminology for these errors: we say a k-bit linear predictor makes a 'Type I' (or unavoidable) error if it makes an error, but agrees with the Markov prediction. In that case, the predictor can not be improved by repairing this error. For example, if our data contains a pattern of k bits, which appears twice, once followed by 0 and 1 at the second instance, then the Markov (and hence, linear) predictor will make an error. If the predictor makes an error, while the Markov predictor is correct, we call this a 'Type II' (or avoidable) error. This error may be due to the linearity of the predictor, or it may be that we have not found the best linear predictor. We will not try to distinguish between these cases. If we find a type II error, we can, in principle, improve our predictor. However, this could mean we have to leave the class of linear models, which might not be a desirable solution from a practical point of view.

Predictor length	No. of errors	Unavoidable (type 1)	Avoidable (type 2)
7	6301	6301	0
30	5660	3070	2590
90	4328	0	4328
128	3764	0	3764

FIGURE 2. The Philips predictor applied to the first frame of sample A (frame size is 37632 bits)

Table 2 suggests that the Philips predictor is indeed optimal for small k. As expected, for large k the Markov predictor will not make any mistake. This, however, has no practical value. On the other hand, comparison between linear and Markov predictors is not really fair. It would be more interesting to define Type I and Type II errors for the class of linear predictors. For example, if our data contains $1, \ldots, 1$ and $-1, \ldots, -1$ (both k times), followed by 1, then any linear predictor is bound to make a mistake. The precise identification of avoidable/unavoidable mistakes seems to be out of reach at the moment.

4. Improving the Markov predictor

We have seen earlier that Markov predictors of high order have superior quality. At the same time, huge overhead (2^k bits) makes them unpractical. In this section we consider two attempts to produce predictors of comparable quality, but with smaller overheads.

4.1. Adaptive Markov Predictor. For each word π, the number of correct predictions will be $\max\big(U^{(k)}(\pi), D^{(k)}(\pi)\big)$, see (3). Consequently the total number of incorrect predictions is

$$\sum_{\pi\in\mathcal{S}^{(k)}} e(\pi) \;=\; \sum_{\pi\in\mathcal{S}^{(k)}} \min\big(U^{(k)}(\pi), D^{(k)}(\pi)\big)\ .$$

As shown in Figures 3 and 4, most of those errors are typically due to just a few words. An adaptive Markov scheme can be designed in the following way:

: 1) Construct the Markov predictor of order k (see (3)).

: 2) Sort the 2^k elements of $\mathcal{S}^{(k)}$ by the number of errors they account for, that is number the words $\pi(1), \pi(2), \ldots, \pi(2^k)$ in such a way that $e(\pi(1)) \geq e(\pi(2)) \geq \cdots \geq e(\pi(2^k))$.

: 3) Choose $k' > k$, $I \in \{1, \ldots, 2^k\}$ and compute the Markov predictor $p^{(k')}(\rho)$ for each word $\rho \in \mathcal{S}^{(k')}$ of the form $\rho = \pi_1\pi_2$, where π_2 is one of $\pi(1), \ldots, \pi(I)$.

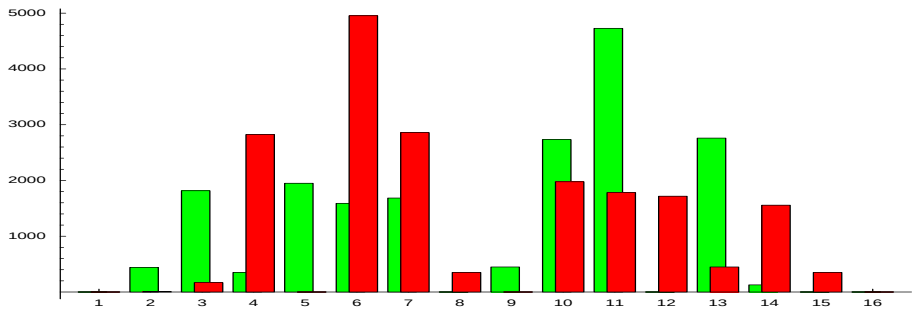


FIGURE 3. Word counts $U^{(4)}$ and $D^{(4)}$ computed for sample A

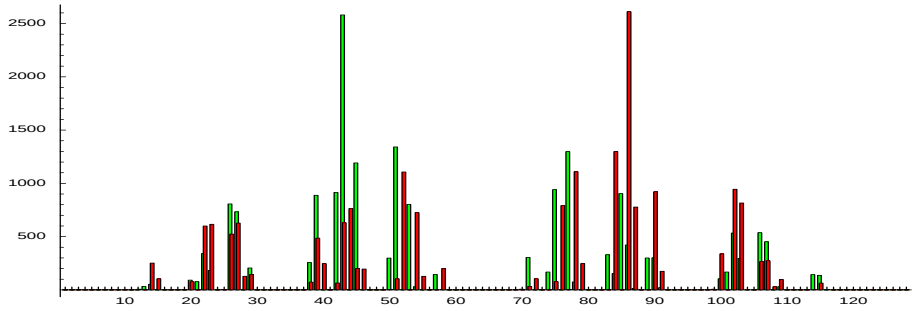


FIGURE 4. Word counts $U^{(7)}$ and $D^{(7)}$ computed for sample A

: 4) Define a predictor $\widehat{x}_n = p_{\mathrm{ad}}^{(k,k')}(\pi_1\pi_2)$, where $\pi_1 = x_{n-k'}\cdots x_{n-k-1}$ and $\pi_2 = x_{n-k}\cdots x_{n-1}$, as follows:

$$p_{\mathrm{ad}}^{(k,k')}(\pi_1\pi_2) = \begin{cases} p^{(k)}(\pi_2)\ , & \text{if } \pi_2 \in \{\pi(I+1),\ldots,\pi(2^k)\}\ , \\ p^{(k')}(\pi_1\pi_2)\ , & \text{if } \pi_2 \in \{\pi(1),\ldots,\pi(I)\}\ . \end{cases} \tag{6}$$

Equation (6) defines our adaptive predictor. The overhead for this is of the order of $2^k + 2^{k'-k}I$ bits.

Such an adaptive predictor has been constructed based on the first 37632 bits of sample A. Firstly, a Markov predictor of order $k = 4$ is constructed. The values of $U^{(4)}$ and $D^{(4)}$ for each word (numbered $1,\ldots,16$) are displayed in Figure 3.

We see that $I = 4$ words (numbers $6, 7, 10$ and 11) account for 86% of the errors (7034 out of 8137). A Markov predictor of order $k' = 8$ on those words can correct 1911 of those errors (a 23% reduction). Hence the adaptive Markov predictor $p_{\mathrm{ad}}^{(4,8)}$ makes 6226 mistakes. The overhead would be around 100 bits.

Starting from a predictor of order $k = 7$ (see Figure 4), and taking $I = 10$ (accounting for $4699/6217 \approx 76\%$ of all errors), a predictor of order

$k' = 14$ makes 3924 errors on those words (an overall reduction of 12%). The adaptive Markov predictor $p_{\text{ad}}^{(7,14)}$ thus makes 5442 errors, with an overhead of about 1400 bits.

4.2. Compressed Markov predictor. The main weakness of Markov predictors is the large size of the overhead. It might be feasible to try to 'compress' this overhead, by representing the Markov predictor in an equivalent, but more economic form. Our idea is to start with a boolean function which describes the Markov predictor and then to minimize its size using Quine's algorithm ([**8**]) for minimization of boolean functions.

Consider the Markov predictor given by (3). Let $\Pi_1^{(k)}$ be the set of words (patterns) of size k, such that Markov predictor predicts 1 on those patterns:

$$\Pi_1^{(k)} = \{\pi = (\pi_1, \dots, \pi_k) \in \{0,1\}^k : p^{(k)}(\pi) = 1\}.$$

Then the following boolean function gives an equivalent representation of our Markov predictor

$$f(x_1, \dots, x_k) = \bigvee_{\pi \in \Pi_1} \bigwedge_{1 \le i \le k} (x_i = \pi_i).$$

Now we can try to minimize the size of f using the Quine minimization algorithm, which minimizes the number of boolean operators in a logical expression.

EXAMPLE 4.1. Suppose the source sequence is 011001 and $k = 2$. The Markov predictor is given by

$$f(x_0, x_1) = ((x_0 = 0) \wedge (x_1 = 1)) \vee ((x_0 = 0) \wedge (x_1 = 0))$$

This Markov predictor does not make any mistakes for our data. After minimization, we conclude that $f(x_0, x_1) = (x_0 = 0)$.

Some tests performed on the real date are summarized in the following table. Frame size was set to 37632.

Order k	No. errors (all are type 1)	Overhead (bits occupied by f)
2	8038	2
7	7917	98
10	7464	1160

Unfortunately, this method does not seem to lead to any substantial improvement of compression ratios.

5. Improving the linear predictor

A coding approach based upon the linear predictor was shown to be very successful, as compared to approaches based on other prediction schemes

such as the Markov predictor. Therefore, it makes sense to take the Philips-approach based on the linear predictor as a starting point and to try variations on this scheme. This will be the subject of this section. We consider the following variations. In subsection 5.2, we try linear predictors with coefficients chosen from a finite set, either $\{0,1\}$ or $\{-1,0,1\}$. In subsection 5.4, we change the optimization criterion to a weighted optimization criterion. In subsection 5.5, we investigate the effect of changing the prediction order k of the linear predictor. Finally, in subsection 5.6, we consider the effect of a lagged estimation scheme in which $\hat{\beta}$ is estimated from the previous bit string rather than from the current bit string. Firstly, let us discuss briefly how the binary data is obtained from an analog signal. An understanding of this transformation might lead to improvements of prediction schemes, see subsection 5.3 below.

5.1. Sigma-Delta Modulation. An audio signal – a relatively smooth function of time – is digitized using a *sigma-delta modulator* (SDM) (see Figures 5 and 6). So the additional information, which, in principle, can help improving the quality of the prediction is the following:

(1) The original signal $X(t)$ is an audio signal, and is a smooth function of time.
(2) The SDM is, theoretically, a deterministic function of $X(t)$, and it produces the same binary sequences as an output, given the same audio signal as an input. But in practice, some noise is always present in the working circuit of an SDM. This noise can occasionally invert an output bit of the SDM.

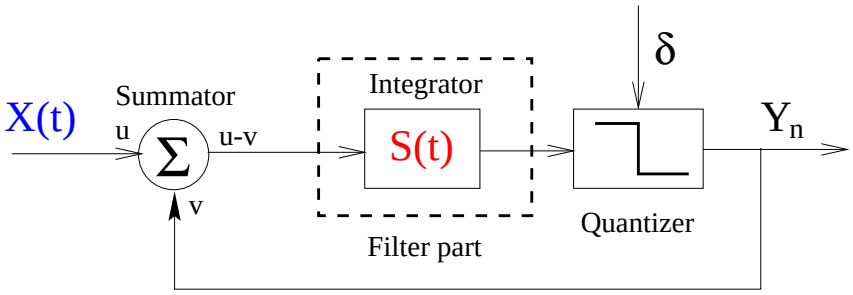


FIGURE 5. A primitive 1-st order SDM with one integrator. The input $S(t)$ to the quantizer is the integral of the difference between the original signal $X(t)$ and Y_n, the output of the quantizer.

It follows from (1), that if some frequency f is essential in the spectrum of $X(t)$, then $X(t)$ and $X(t+1/f)$ are highly correlated. Since $1/f$ contains $1/(\delta f)$ quanta of time, there are dependencies in the ouput binary stream (x_n) at distance $k = 1/(\delta f)$. For example, the frequency $f = 1000$ Hz

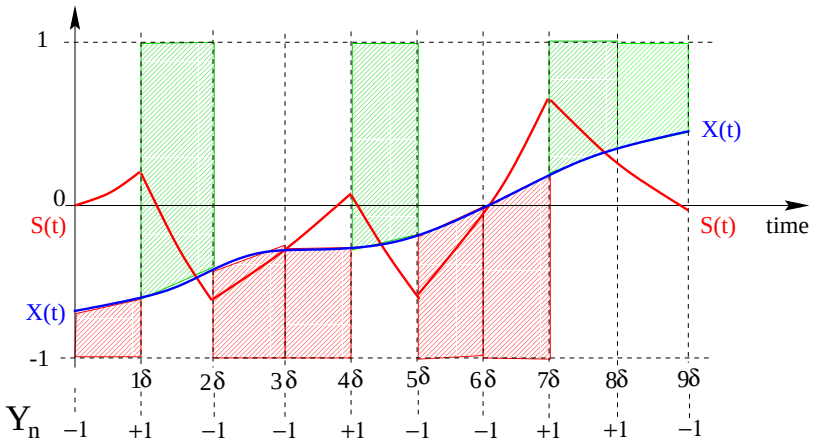


FIGURE 6. A primitive 1-st order SDM. The integrated difference $S(t)$ as a function of the original signal $X(t)$.

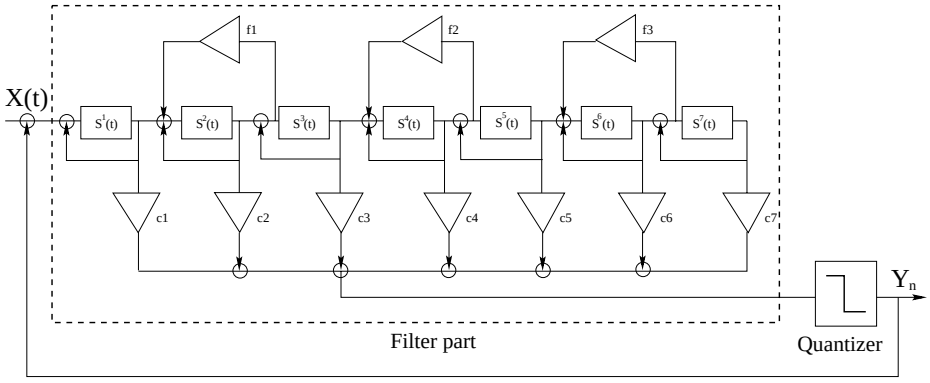


FIGURE 7. A practical implementation of a 7-th order SDM in [**7**]

corresponds to dependencies at distance $k = 44.1 * 64 = 2822.4$ quanta of time, if one quantum of time $\delta = (44100 * 64)^{-1}$ seconds. This suggests that a predictor of length 2000–3000 should be used. In practice however, predictors of much smaller order k are used. Again, long predictors are not efficient due to a large overhead. It is nevertheless possible to use longer predictors, provided we use coefficients that can be stored using only a few bits: for example, binary ($\beta_i \in \{-1, 1\}$) or ternary ($\beta_i \in \{-1, 0, 1\}$).

5.2. Linear predictors of low pression. We have seen that finding the optimal linear predictor is equivalent to finding a vector $\beta = (\beta_1, \ldots, \beta_k) \in \mathbb{R}^k$ such that the scalar product $\beta \cdot Y_n$ is strictly bigger than 0 as often as possible, where $Y_n = (x_n x_{n-1}, \ldots, x_n x_{n-k}) \in \{-1, 1\}^k$. In other words, we

need to find a hyperplane containing 0 in $\mathbb{R}^k$ such that as many Y_n's as possible lie on the same side.

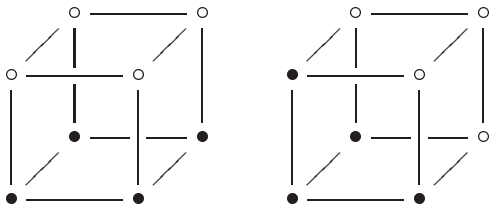
This problem can theoretically be solved. Moreover, the fact that all the Y_n lie on the cube $\{-1,1\}^k$ should help. It looks as if the only β's one needs to try are the ones which have entries in $\{-1,0,1\}$, where one should exclude the β's such that the number of nonzero entries is even, to avoid that $\beta \cdot Y_n = 0$ for some Y_n.

A 'linear' predictor for us is a special function

$$\{-1,1\}^k \longrightarrow \{-1,1\}$$

which takes the value 1 on one side of a hyperplane and the value -1 on the other. In particular it maps half of the elements to 1 and half of them to -1.

For $k = 3$, up to symmetry we get the following possibilities:

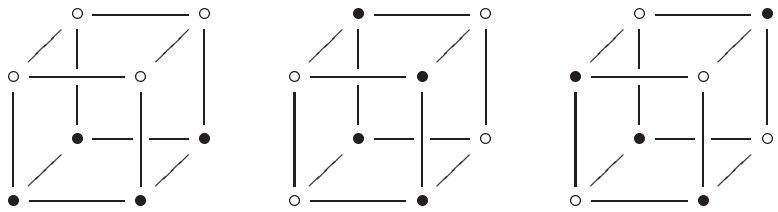


Here we denote points which are mapped to 1 by $\circ$ and points which are mapped to -1 by $\bullet$. There are 6 'linear' predictors of the first type (as the cube has 6 faces) and 8 'linear' predictors of the second type (corresponding to the 8 vertices of the cube).

Functions $\{-1,1\}^k \longrightarrow \{-1,1\}$ are the same as functions $\{0,1\}^k \longrightarrow \{0,1\}$. Regarding $\{0,1\}$ as the field $\mathbb{F}$ with two elements we can consider special functions of the form

$$\mathbb{F}^k \longrightarrow \mathbb{F}, \qquad (y_1, \ldots, y_k) \mapsto \sum \beta_i y_i + c,$$

where $\beta = (\beta_1, \ldots, \beta_k) \in \mathbb{F}^k$ and $c \in \mathbb{F}$. We assume that β is not the zero vector, such that half of the elements is mapped to 0 and half of the elements is mapped to 1. As multiplication and addition in $\mathbb{F}$ are just AND and XOR (where 1 corresponds to TRUE and 0 corresponds to FALSE) these functions are easy to handle on a computer. The class of functions one gets is different from the class of 'linear' predictors. Again we draw the case of $k = 3$:



Now we denote points which are mapped to 0 by ∘ and points which are mapped to 1 by •. We get 6 functions of the first and second kind and 2 of the third kind. Here the first kind corresponds to a β containing two zero entries, the second kind to a β containing one zero entry, and the third one to $b = (1, 1, 1)$.

The predictors on $\mathbb{F} = \mathbb{Z}_2$. Let for the moment $x_n \in \{0, 1\}$ and consider predictors

$$\hat{x}_n = \beta_0 + \sum_{i=1}^{k} \beta_i x_{n-i} \ (\mod 2), \quad \beta_i \in \mathbb{F}.$$

We compared, for every $k \leq 12$, all 2^{k+1} predictors and selected the optimal one. The minimal percentage of errors in the frames varies from 9% to 22%, and is equal to $\sim 17.5\%$ on average, independent of the sample.

The ternary predictors. Here we return to $x_n \in \{-1, 1\}$ and to predictors of the form

$$\hat{x}_n = \operatorname{sign}\Big(\beta_0 + \sum_{i=1}^{k} \beta_i x_{n-i}\Big),$$

where $\beta_i \in \{-1, 0, 1\}$. We have used the fast gradient method to minimize the number of errors. The percentage of errors p equals, on average, 18.5% independent of the sample. The optimal k was of order 10.
We remark here, that it is strange that ternary predictors give worse results than binary. Probably it is due to a poor optimisation that we do get worse results here.

As a conclusion we would suppose that the idea to use the low resolution predictors is not promising and could not provide the desirable $p \sim 5\%$.

5.3. SDM predictors. An SDM is an almost deterministic mapping of a signal $X(t)$ into the binary sequence $\{x_i\}$. This means that, given a signal $X(t)$, one could write an adjustable model of an SDM to produce the output sequence $\{x_i\}$ almost without errors. We could write it in a form

$$\hat{x}_n = M\Big(X(t), t \leq \delta n; x_i, i < n\Big), \tag{7}$$

where M is some fixed map, depending on the realization of the SDM, and $\hat{x}_n$ predicts x_n almost without errors.

In our setup, the signal $X(t), t \leq \delta n$, is unknown. However, we can reconstruct and extrapolate it from $x_j, j < n$ applying the low pass filter with some coefficients. One can obtain those coefficients by analyzing the DAC of the decoder. Therefore, we have

$$\hat{X}(t) = f(t;\ x_i, i < n), \qquad t \leq \delta n.$$

Substituting the above estimate for $X(t)$ in (7), we should obtain an accurate estimate for x_n.

We have not completed this approach because it demands the exact knowledge of the SDM diagram and DAC filter characteristics. Nevertheless, to see how promising this approach can be, we have tried a simplified version of the algorithm. The following predictor

$$\begin{aligned} \hat{X}(t) &= (x_{n-1} + \cdots + x_{n-300})/300, \\ S(t) &= S(t-\delta) + x_{n-1} * \text{const}, \\ \hat{x}_n &= \text{sign}\Big(\hat{X}(t) - S(t)\Big). \end{aligned}$$

gives about $\sim 18\%$ of errors, which is quite promising.

5.4. Weighting. A promising approach to improve predictors is 'boosting', see e.g. [**9**] or [**4**]. In boosting, a number of predictors is constructed that are combined to yield the final predictor. The individual predictors that make up the final one are constructed by training a given base predictor from the same training data, using different weights. The weights are such that cases that were predicted wrong frequently with the predictors already constructed, are given more weight during the construction of the next one.

We have not fully explored the possibilities inherent in boosting. However, we did an experiment with the reweighting of badly predicted cases that is at the heart of boosting. Thus, we arrived at the following scheme to estimate the coefficient vector of the linear predictor.

$$\begin{aligned} \hat{\beta}_0 &:= \text{argmin}_\beta \sum_{n=k+1}^{N} \Big(x_n - \sum_{j=1}^{k} \beta_j x_{n-j}\Big)^2 \\ w_n &:= \begin{cases} 1 & \text{if } \Big|\sum_{j=1}^{k} \beta_j x_{n-j}\Big| \leq 1/2 \\ 0 & \text{otherwise, for } n = k+1, \ldots, N \end{cases} \\ \hat{\beta}_1 &:= \text{argmin}_\beta \sum_{n=k+1}^{N} w_n \Big(x_n - \sum_{j=1}^{k} \beta_j x_{n-j}\Big)^2. \end{aligned}$$

Thus the algorithm starts by constructing the basic linear predictor using ordinary least squares. It then throws away all 'sure' predictions by giving weight 0 to all cases for which the absolute predicted value exceeds 1/2. It

then constructs a coefficient vector from the remaining cases in the sample. The coefficient vector $\hat{\beta}_1$ thus found will be used for the linear prediction.

This simple scheme performs remarkably well; some numerical results are displayed in Figure 8. It can be observed that weighting improves the basic scheme by something between 0.5% and 2%. It is remarkable that the improvement was uniform: on all frames tried in all sequences did weighting improve.

Of course, many variants of this basic scheme are possible. Obvious possibilities are

- Changing the threshold from 1/2 to other values
- Iterating the scheme

However, the basic scheme given above proved to be very difficult to improve upon, and it will be used in the remainder.

5.5. Changing the prediction order. In this subsection, we investigate the effect of the prediction order on the prediction accuracy. Now, increasing the prediction order will 'obviously' improve the prediction accuracy. However, this will not necessarily lead to a more efficient coding scheme: the predictor must also be transmitted and longer predictors require more bits.

We have not investigated the efficient coding of the linear predictor separately. However, to compensate for this effect, we have simultaneously changed the length of the bit sequence on which the linear predictor is based. We expect that, roughly, the coding of a linear predictor is proportional to the length of the predictor. Hence it will be as efficient to use a linear predictor of length k for a bit sequence of length N as it is to use a linear predictor of length ck on a bit sequence of length cN for real valued c.

Qualitatively, it was found that decreasing the order of the predictor did always deteriorate the prediction performance. Increasing the order of the predictor to 256 did yield improvement; but there was no further gain in increasing the order to 512. The gain was only marginal with the basic linear predictor, but more substantial with the linear predictor obtained by reweighting as described in the previous section. It was found that the increased complexity of the predictor could be compensated for by increasing the frame size.

Figure 8 gives qualitative results and displays the prediction error of the linear predictor with prediction order $k = 256$ estimated from frames of size $N = 100000$. These can be compared with the errors with the other approaches described above. It is easily observed that the proposal from this section outperfroms the other approaches. Again, the improvement was uniform in that the current approach was better on all frames tried in all sequences.

5.6. Reducing the overhead. A final possibility to improve on the basic linear predictor scheme is to reduce the overhead by computing the predictor from the bits that have already been transmitted, at the decoder, rather than transmitting the predictor from encoder to decoder. Of course, the former scheme has several disadvantages as compared to the latter:

- It requires computational power at the decoder that is not needed with the current set-up.
- It makes the music less accessible: we cannot decode a given bit sequence without decoding its full past.

There are two basic strategies to implement such a scheme: adapting a current solution or recomputing a solution. In the former strategy, the coefficient vector of the linear predictor is updated continuously, i.e. after every new bit. The second possibility is to recompute the coefficient vector from a previous block of bits: the coefficient vector to predict the bits in the current frame is computed from the bits in the previous frame.

We have experimented only with the latter scheme, which we will call the lagged scheme. We give qualitative results only. Generally, the predictions from the lagged scheme are as good as the predictions from the original scheme. This holds true both for the original unweighted optimization and the proposed weighted optimization. This is the case for sequences A, B, and C. For sequence D the results are mixed. For the major part of the sequence the results are comparable. However, in the middle of sequence D the lagged predictor works very badly.

For this reason, the lagged predictor cannot be used without precautions. However, some approximations to the lagged scheme can be practical. We suggest the following options:

- The keep bit, which informs the coder to keep the coefficient vector from the previous frame. As the lagged predictor is almost as good as the predictor based on the current frame, this amounts to a saving of about 50% in the transmission of the coefficient vector.
- the link bit, which tells the decoder where it can find the required coefficient vector
- the recompute info, e.g. the lag, which tells the decoder which parameters to utilize in the recomputation

However, the savings that can be expected from this lagged scheme are limited as it can save at most the transmission of the coefficient vector. Because of these limited savings and the inherent problems, we have not pursued this lagged scheme much further.

Another possibility of saving on the transmission (storage) of the coefficients, is to use some methods of the theory of machine learning. For example, one of its oldest algorithms, the so-called *Perceptron Algorithm*. Let $\beta = (\beta_1, \dots, \beta_k)$ be a real vector, and $X_{n-k,n-1} = (x_{n-k}, \dots, x_{n-1})$ are

the last k observed bits. We predict $\hat{x}_n = 1$ if $(\beta, X_{n-k,n-1})$ is positive, and -1, otherwise. However, we are going to update β after each mistake:
a) if we predict $\hat{x}_n = -1$, while $x_n = 1$, let $\beta' = \beta + X_{n-k,n-1}$;
b) if we predict $\hat{x}_n = 1$, while $x_n = -1$, let $\beta' = \beta - X_{n-k,n-1}$.
To start the algorithm we may choose $\beta = (1, 1, \dots, 1)$. Motivation for the updating rules as above is the following:

$$(\beta', X_{n-k,n-1}) = (\beta, X_{n-k,n-1}) \pm (X_{n-k,n-1}, X_{n-k,n-1}),$$

so the value of $(\beta', X_{n-k,n-1})$ is closer to x_n, then $(\beta, X_{n-k,n-1})$.

We have applied the Perceptron Algorithm to our data. The quality of the prediction is substantially worse. On average, the perceptron algorithm makes twice the number of errors of the least squares predictor. This poor quality of the predictor is not compensated by the space we saved by not transmitting the coefficients. Therefore, the overall compression of the scheme based on Perceptron algorithm is lower. It is interesting to mention that after training the perceptron algorithm on a long sequence, the corresponding vector of coefficients β is very close to the one obtained from the Philips prediction scheme (5).

5.7. Conclusion. The linear prediction scheme suggested in [**5, 6**] leads to an efficient compression algorithm. It seems that in many cases the Philips linear predictor is quite close to the optimal linear predictor. Nevertheless, further improvements are still possible. Let us summarize our results on possible ways of improving the quality of linear predictors.

Switching to the longer predictors of low precision would probably not improve the overall performance of a linear predictor.

Weighting improves. Note that this improvement can be achieved with the current decoders as it requires change only for the encoder.

Some further benefits can be obtained by fine-tuning the frame size and the prediction order: it is particularly advised to increase the prediction order to 256. This can be achieved without loss in efficiency if the frame size is increased from 40000 to, say, 100000. Further optimizations are possible here.

Note that the improvements for frame A are then substantial: they amount to a reduction of approximately 75% of the storage of the bit string; the reduction obtained with the original approach is roughly 66%. The improvements on the other frames are less impressive. Whether these changes are worth the trouble will obviously depend on which of the frames is more typical.

Moreover, the cost of transmission of the coefficient vector can be reduced by at least 50% if a keep bit is defined.

Adaptive schemes, like the perceptron algorithm, seem to be unfeasible. In our opinion, it simply takes a very long time (and hence, a lot of mistakes

will be made) for such a scheme to achieve a quality comparable to that of the least squares predictor.

Finally, the optimal linear predictors cannot be computed (or estimated) efficiently at the present time.

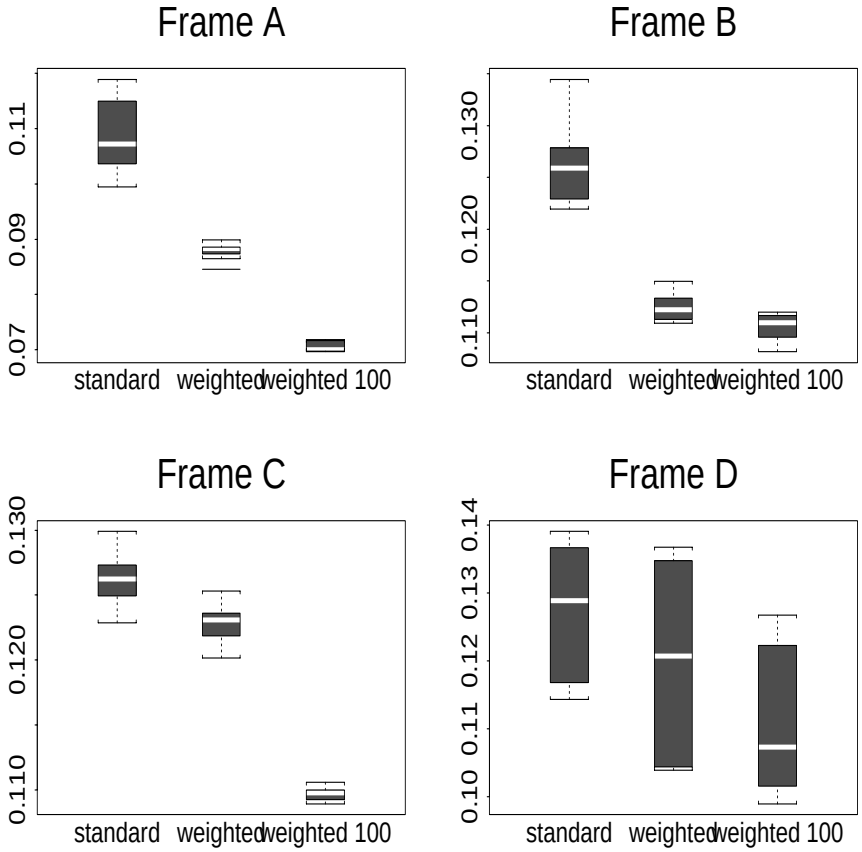


FIGURE 8. Prediction errors for several frames in given sequences for several approaches based on linear prediction. Standard: coefficient vector of length 128 estimated from sample of size 40000 using unweighted least squares. Weighted: coefficient vector of length 128 estimated from sample of size 40000 using weighted least squares. Weighted 100: coefficient vector of length 256 estimated from sample of size 100000 using weighted least squares.

Bibliography

[1] E. Amaldi, M.E. Pfetsch, L.E. Trotter, Some structural and algorithmic properties of the maximum feasible subsystem problem. *Integer programming and combinatorial optimization (Graz, 1999)*, 45–59, Lecture Notes in Comput. Sci., 1610, Springer, Berlin, 1999.

[2] E. Amaldi, V. Kann, On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems. *Theoret. Comput. Sci.* 209 (1998), no. 1-2, 237–260.

[3] H. Arora, A. McLean, Stability Analysis of 1st and 2nd Order Sigma Delta Analog to Digital Converter, *preprint* `www.duke.edu/~ha/HimanshuArthur.pdf`

[4] L. Breiman, Arcing classifiers. The Annals of Statistics 26, 3, pp. 801-849, 1998.

[5] F. Bruekers, W. Oomen, R. van der Vleuten, and L. van de Kerkhof, Lossless coding of 1-bit audio signals. *AES 8th Regional Convention, Tokyo, Japan,* 1997.

[6] F. Bruekers, W. Oomen , R. van der Vleuten, and L. van de Kerkhof, Improved lossless coding of 1-bit audio signals. *AES 103rd Convention, New York*, 1997.

[7] D. Reefman, P. Nuijten, Why Direct Stream Digital is the best choice as a digital audio format, *Audio ES, Convention Paper, preprint*, 2001.

[8] W. Quine, The problem of simplifying truth functions, American Mathematical Monthly, 1952, 59, pp.521-531.

[9] R. Schapire, Y. Freund, P. Bartlett, and W. Lee, Boosting the margin: a new explanation for the effectiveness of voting. The Annals of Statistics 26, 5, pp. 1651-1686, 1998.

pp. 41 – 57

CHAPTER 3

The Euro Diffusion Project

Piet van Blokland, Lorna Booth, Kirankumar Hiremath, , Michiel Hochstenbach, Ger Koole, Sorin Pop, Marieke Quant, Djoko Wirosoetisno.

Abstract. From 1st January 2002 we have the unique possibility to follow the spread of national euro coins over the different European countries. We model and analyse this movement and estimate the time it will take before on average half the coins in our wallet will be foreign.

Keywords: Markov chains, diffusion

1. Introduction

On January 1, 2002 a total of 12 European countries replaced their national currencies with the euro. These coins are not identical, one side of each of the 8 denominations of coins differs from country to country, and as people travel these coins mix. The national banks have decided that no redistribution will take place, and therefore it is expected that in the long run a close to perfect mixture of coins will take place. In this paper we analyse this problem and come up with a mathematical model for the movement of coins that enables us to estimate the speed at which this process will occur. This speed strongly depends on the value of certain parameters that are hard to estimate and for this reason we have to be careful with conclusions. However we estimate that it will take around 12 months before roughly half of all coins in Dutch wallets will be foreign.

Our other main conclusions are:

- Markov chains are the obvious models to model the movement of euro coins,
- Continuous models are the natural approximation of these,
- The data on the Euro Diffusion web site are unreliable and must be used for parameter estimations only with care.

In this paper each section or subsection is marked with a number of stars to indicate approximate level of difficulty:

* means that the section can be understood without effort,

** means that the section contains mathematics which is explained very gently,

*** means that the section would be accessible to anyone who has done a little university level mathematics, or who is prepared to work a bit more at understanding.

2. Basic facts*

The euro was introduced in 12 European countries at the beginning of 2002. A total of 64.9 gigacoins (1 GC = 1 000 000 000 coins) were made, of which 3.3 GC are Dutch, but not all these coins were brought into circulation right away: in Holland 1.6 GC were put into the market on 1st January 2002. The number of coins made by each country is not in proportion with the population: e.g., France made around 190 coins per person, Germany 280, and Holland 200. The reason for these differences is unknown and is part of the policy of each national bank.

New coins are still being brought into circulation after the introduction of the euro, mainly to compensate for savings and wastage. At the end of 2001 all money-boxes were emptied of their national currencies, and now they are slowly being filled again with euros. The Dutch national bank (DNB) expects that this will require about 100 MC per month in the Netherlands. The amount of coins saved is eventually expected to equal the amount in active circulation; in the guilder age it was estimated that 1.5 GC out of a total of 3.0 GC were saved. Wastage is considerably less: in the guilder age it was about 50 to 100 MC per year, largely due to loss (people dropping coins, etc.) and coins going abroad forever. We think that wastage will be a little less as tourists from other countries can also use the Dutch euro coins in other European countries, and therefore 60 MC, 5 MC per month, seems to be a reasonable estimate of this type of loss. The fact that there are new coins brought into circulation will mean that the mix will never be perfect, but as the waste is tiny compared to the total number of coins in circulation, this effect will be small.

A new effect, which is hard to estimate, is the collection of euro coins. Collection used to be an effect of negligible size; but nowadays many people seem to collect foreign euro coins. The size of this effect and when these coins will be brought into active circulation again are difficult to estimate.

When counting the number of coins in our wallets, we usually find 10 to 20 coins, and this is confirmed by a study from DNB which shows that we carry on average 15 coins. However, per person 100 coins were brought into circulation. The remaining 85% of all coins can be found at check-outs in shops and at banks. This may imply that movement of the euro coins will be relatively slow: only 15% of all coins can be taken abroad at any one time! For example, suppose that 2/3 of the Dutch population go abroad during the summer holidays, and assume that after their foreign visits their coins are representative of the coin mixture in the country they visited (which will

certainly contain a (small) percentage Dutch coins!). This replaces only 10% of all Dutch coins with (mainly) foreign coins.

People travelling and taking their coins abroad is not the only possible reason for the spread of euro coins. Another is the possibility that Dutch banks will buy euro coins from somewhere other than DNB, which might be cheaper, for example, if a depot of the Belgian or German national bank is closer to the particular Dutch bank. The effects of this and how national banks will react to it are difficult to predict; nobody knows if it will occur and to which extent. This effect might even create an imbalance between the quantities of coins in a country: there might be a net flow of coins in or out a country.

3. Data

3.1. Observations*. To fit the parameters in our model, we use the data gathered in the web-site

`http://www.wiskgenoot.nl/eurodiffusie/`

which is built up by voluntary observations, mainly from different parts of the Netherlands and Flemish Belgium, and from a few points elsewhere. Each data point consists of the date and locale, and the number of coins the correspondent has, along with the breakdown according to denominations and countries of origin. Various summaries are available on the web-site.

We summarise the observations for the percentage of 1-euro coins in the Netherlands in Table 1.

t	nl	be	de	fi	fr	gr	ie	it	lu	os	pt	es	#no.
1	90.9	1.1	4.2	0.0	1.1	0.0	0.0	1.1	0.3	0.0	0.3	0.8	353
2	84.7	3.6	4.4	0.0	2.9	1.7	0.4	0.1	0.3	1.0	0.0	0.8	724
3	82.7	5.2	5.1	0.2	3.6	0.0	0.1	1.0	0.7	0.6	0.0	0.8	1186
4	85.2	2.8	4.7	0.2	2.3	0.2	0.2	1.1	0.1	2.2	0.2	0.6	3524
5	80.6	5.5	5.7	0.2	4.3	0.0	0.0	1.0	0.1	1.3	0.2	0.9	976
6	75.3	5.8	8.9	0.1	3.3	0.2	0.5	1.6	0.0	2.5	0.2	1.7	1260
7	79.4	5.4	6.9	0.2	2.5	0.1	0.4	1.2	0.4	2.2	0.2	1.2	3977
8	74.7	10.3	5.4	0.1	3.7	0.3	0.1	1.1	0.5	1.9	0.6	1.3	874
9	80.0	5.6	7.1	0.4	1.9	0.6	0.4	1.2	0.0	1.3	0.4	1.2	520

TABLE 1. Percentage of 1-euro coins of various types in the Netherlands

Here line `1` corresponds to an average over the first ten days in January and so on; the countries are labelled by their internet abbreviations; the last column is the total number of coins counted, to provide some indications of the error. Similarly we see the number of 1-euro coins for Belgium in Table 2.

In Figure 1, we plot the percentages of local coins in the Netherlands and Belgium, with the horizontal axis being the time in days.

t	nl	be	de	fi	fr	gr	ie	it	lu	os	pt	es	#no.
1	0.0	94.6	2.7	0.0	2.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	37
2	5.3	89.8	1.9	0.0	1.9	0.0	0.0	0.3	0.6	0.0	0.0	0.3	323
3	5.1	87.6	1.8	0.1	3.3	0.1	0.0	0.6	0.3	0.5	0.0	0.4	930
4	5.4	85.7	2.5	0.1	3.4	0.0	0.0	0.6	0.6	0.8	0.1	0.6	4045
5	5.1	74.2	4.5	0.3	10.3	0.1	0.3	1.2	1.5	1.4	0.0	1.3	1027
6	6.7	74.4	4.9	0.1	7.3	0.0	0.2	1.5	1.0	2.5	0.2	1.2	2275
7	6.8	75.9	4.4	0.1	6.3	0.2	0.2	1.7	1.8	1.3	0.1	1.2	4263
8	5.7	82.6	3.0	0.4	3.1	0.1	0.4	1.2	0.6	0.4	0.1	2.3	1376
9	5.2	77.2	3.2	0.0	5.5	0.0	0.8	1.5	1.2	0.2	0.0	5.0	400

TABLE 2. Percentage of 1-euro coins of various types in Belgium

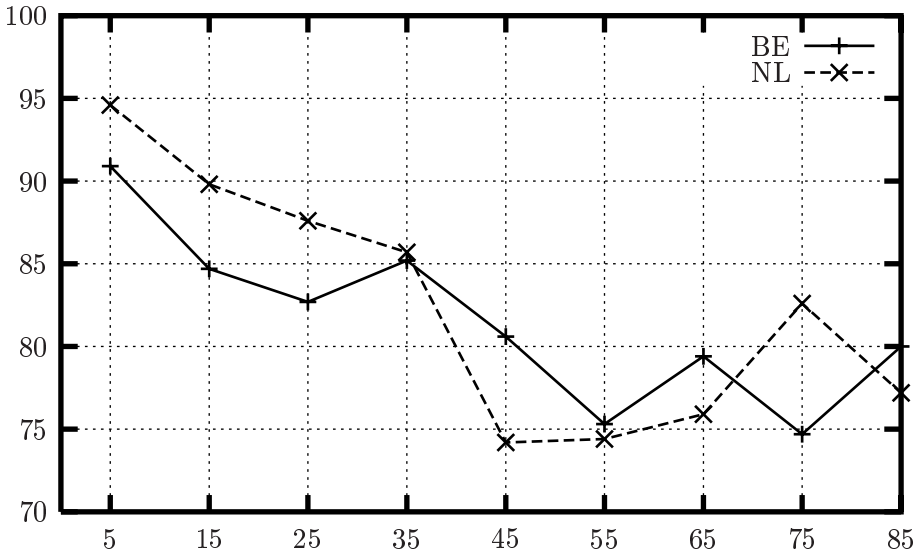


FIGURE 1. Percentage of local coins in the Netherlands and Belgium over the first 85 days

Although a general downward trend can be seen for both curves, the size of the fluctuations is rather puzzling, considering the fact that the numbers of coins counted are quite large (corresponding to hundreds or thousands of people). A similar irregular behaviour is also observed when the data is restricted to smaller areas (e.g., greater Utrecht or Brussels).

We make a couple of observations about the data:

0. Not surprisingly, the percentage of local coins at any given location tend to decrease over time. On February 1, these numbers are 91% for the Netherlands and 89% for Belgium (averaged over all denominations); the corresponding figures are 87% and 81%, respectively, for March 1.
A possibly more illuminating example is the case of Luxembourg: *72% of

coins are (still) of local origin on February 1, and *68% on March 1 (data is from the Euro Diffusion web-site, asterisks denote possible inaccuracies.) This suggests that the mixing process is likely to take several years.

1. The observed data for the Netherlands and Belgium tell us that coins of different denominations mix at different rates. At any given time, the percentage of foreign coins is higher the higher the denomination is: In the Netherlands on March 1, only 8.3% of the 1-cent coins are of foreign origin; for 10c it is 11.7%, and for 2-euro it is 19.6%.

2. Data for the Netherlands and Belgium show apparently anomalously large proportions of Austrian coins, being 2.3% and 1.7% on March 1 (well on its way to the long-time limit of 3.1%). Data from the Canary Islands (Spain) and Creta (Greece) show that foreign coins are almost exclusively German. These suggest that there are considerable variations in the travel habits of people from different countries; one may attempt to determine certain aspects of these in addition to the questions posed above.

The fluctuation in the data may be partially explained by the fact that at the Euro Diffusion web site everybody on the Internet can enter measurements of the number of different euro coins in their wallets. For this reason there is no guarantee that these measurements represent the real situation: the selection of people is probably biased and the moments at which they enter data might be strongly biased (e.g., somebody enters data when he has something to tell, i.e., he has some "rare" coins in his wallet!). The extent to which these biases corrupt the data is difficult to measure without further study, so to get some feeling for it we did some counts ourselves during the workshop. The results can be found in Table 3.

Who/where	date	NL	B	other countries
Lunch (Amsterdam)	19 Feb	293/93.1%	12/3.8%	10/3.1%
Lorna's Students (Utrecht)	20 Feb	125/98.4%	2/1.6%	0/0.0%
Check-out K.d.V. Inst. (A'dam)	20 Feb	233/94.7%	3/1.2%	10/4.1%

TABLE 3. Measurements during the SWI workshop in February 2002

As an alternative we tried to get the, in our opinion, most reliable data from the web site. We looked for unique "eurometers", with multiple participants, that have entered data three times. The results can be found in Table 4.

Eurometer	dates	percentages NL
46 (Rotterdam)	18/1, 31/1, 8/2	98, 96, 98
167 (Utrecht)	21/1, 11/2, 18/2	95, 98, 96
510 (Hengelo (O))	15/1, 24/1, 1/2	96, 92, 91

TABLE 4. Selected measurements from Euro Diffusion database

It is surprising to see that, also in the measurements considered to be reliable, there is a high variability in outcome, despite the size of the measurements. This suggests that the measurements come from different distributions.

4. A First Estimate**

To come to an estimate of the whereabouts of coins on 1st February we assume that diffusion is approximately linear during January and February. Thus the measurements at Feb 19 and 20 of Table 3 should be multiplied by roughly 3/5 to find an estimate for the end of January. For the measurements in Table 4 similar factors are used. This gives the following estimate (numbers correspond to own measurements and "reliable" eurometers, respectively):

$$(1)\quad \Big[\frac{3}{5}(6.9+1.6+5.3)+\frac{3}{2}2+4+\frac{3}{4}2+\frac{3}{2}5+\frac{3}{4}2+\frac{3}{5}4+2\times 4+\frac{5}{4}8+9\Big]/12\approx 4.6$$

This result is also unreliable, and biased by the high number of non-Dutch euros measured by eurometer 510, therefore as a first estimate we took 4% diffusion per month. This has to be verified during later months, although even the March 1 measurement will be disturbed by the influence of the February holidays.

Note that we assumed that the percentage of foreign coins equals the rate at which the coins move between the Netherlands and abroad, which is a good approximation as long as the diffusion is slow and we are near to the beginning of the process. In case of an increased rate this is no longer the case, then we have to use more sophisticated models, and estimate based upon these models. In the next section we will describe the types of models we use, and after that we will explore a number of techniques for estimating the parameters more accurately.

5. Modelling the Spread of Euros with Markov Chains

Following the grand total of 64.9 billion euro coins in some mathematical model is impossible. It is also useless: the movement of a single coin shows

the same (type of) behaviour as any other coin, although this may depend on the country of origin and denomination. Therefore we may concentrate on a single coin and follow it in a probabilistic manner its way around Europe. From this we can then draw conclusions for the entire population of euro coins.

The movement of some arbitrary coin can be followed by a *Markov chain*. The simplest Markov chain model has only two states but can already be used to give a first simple model for the movement of coins. We will describe this model in detail, discuss its shortcomings, and deal in less detail with more sophisticated models.

5.1. Two-state model****.** In the simplest model we have two states: a coin can either be in Holland or abroad. We denote these states with H and A. We neglect wastage, and do not distinguish between coins in active circulation and saved coins. We also assume that the net rate of coins into the Netherlands is equal to the net rate out. Let d be the rate at which diffusion occurs: this means the percentage of coins in Holland that is replaced by coins from abroad after one month. Note that it is *not* the rate at which non-Dutch coins arrive: foreign coins in Holland may go abroad again, and Dutch coins abroad may be taken back to Holland! As we can assume that the net flow equals 0, we know that (on average) every coin going abroad is replaced by a coin coming from abroad. Thus having d percent of foreign coins in the Netherlands after 1 month corresponds to there being a probability of d that an arbitrary coin "goes" abroad and doesn't return in one month. Thus p_{HA}, the probability of a transition from H to A in the 2-state Markov chain, is d and, assuming an active circulation of 1.6 GC in the Netherlands, the flow out in one month is $1.6d$ GC. As the flow in equal the flow out, we find $p_{AH} = \frac{1.6d}{ACA}$, with ACA the active circulation abroad. Assuming that the active circulation abroad is the proportion of coins in active circulation in the Netherlands multiplied by the total number of coins abroad, we see that $ACA = \frac{1.6}{3.3}61.6 = 29.9$ GC. Using this we find that if $d = 0.04$ (see section 3) then $p_{AH} = 0.00214$. Now that we know the transition probabilities we can start our computation. Let Q be the *transition matrix*:

$$Q = \begin{pmatrix} 1 - p_{HA} & p_{HA} \\ p_{AH} & 1 - p_{AH} \end{pmatrix} = \begin{pmatrix} 0.96 & 0.04 \\ 0.00214 & 0.99786 \end{pmatrix}.$$

The matrix Q should be interpreted as follows: The first row shows where a coin, initially in Holland, will be after 1 month. The first entry shows the probability that it will be in Holland, the second entry that it will be abroad. The second row corresponds to a coin originating from abroad. Again, the number on the diagonal, Q_{22}, is the probability that the coin is abroad after 1 month; the left hand element is the probability that the coin is in Holland. Multiplying Q with itself gives numbers with the same interpretation, but for a 2-month period, and, similarly, Q^n gives the transition probabilities

for n months. For example lets look at what happens after 15 months. Computation, with $d = 0.04$, shows that

$$Q^{15} = \begin{pmatrix} 0.54841 & 0.45159 \\ 0.02639 & 0.97361 \end{pmatrix}.$$

Thus, 15 months with a diffusion at the rate of January will lead to 45% of Dutch coins being abroad. They are replaced by foreign coins, and thus there are 45% foreign coins in the Netherlands. coins.)

We expect that the speed of diffusion is temporarily increased by the summer holidays and the February skiing holidays. Assume that 2 out 3 people go on holidays during summer to another country which uses the euro. As 15% of all coin are in our wallets, this will result in 10% diffusion. At a 4% diffusion rate, this is the equivalent of 2.5 months. Together with the effect of the skiing holidays we assume that the combined holidays count for 3 to 4 months. Thus the situation by the end of the year is equivalent to 15 to 16 months at rate 0.04. To conclude, we expect that the percentage of foreign coins will be 50% somewhere in the first months of 2003.

5.2. Savings*. We can model the effects due to people saving in a similar way. We now have another state - the piggy-bank - from which coins are released only after a long time. To compensate for this the Dutch National Bank puts new (Dutch) euros into circulation until the piggy-banks are full. Simulations show that this slows the diffusion of coins, for two reasons. The first is the addition of Dutch euros and the second is that the piggy-banks act as a reserve of Dutch euros which are released later into the money flow.

5.3. Refinements*. The simple model has $(0.05085, 0.94915)$ as equilibrium, which is is not realistic as wastage is ignored. Wastage leads to shortages of coins that lead to the production of new coins. These new coins are then (presumably) released in the home country, leading to a slightly higher percentage of national coins in each country.

In the simple model all countries are aggregated in a single state, however it is to be expected that diffusion occurs faster with certain countries than with other countries. A model with multiple states representing different countries could make predictions on the diffusion of other coins.

Finally, a regional approach, splitting countries up in regions, can be followed. This would be of particular value if we want to use the data from the Euro Diffusion web-site to the full. In this case we could use a model that has all of the regions (or groups of the regions) used by the web-site, to estimate how quickly coins move around inside the Netherlands and Belgium.

As also mentioned in Section 5.1, seasonal variations will have non-negligible effects on the transition to an equilibrium. It may be almost impossible to insert these variations into the model in a realistic way.

6. Markov Chains and Continuous Models**

In the last section we used Markov Chains to model the movement of coins around. Each coin jumped around independently with the probabilities given in the matrix Q. For example, the chance that a coin that was in the Netherlands moves outside it after one month is 0.04. However if we have a very large number of coins, as we do, roughly 4 per cent of the coins that are the Netherlands will move outside of it. Although each coin chooses to move or not independently of the other coins, there is a mathematical theorem, the Law of Large Numbers, which tells us that the actual proportion of coins moving will be very close to 4 per cent. Therefore if we assume a continuous model, in which *exactly* 4 per cent of the coins move from the Netherlands to somewhere else, we will be close to the truth. In fact the error in the model will be roughly a constant divided by the square root of the number of coins, which really is very small. Therefore in Sections 7 and 8 we consider continuous models.

7. Parameter estimation based on data from multiple months**

Suppose we have a matrix Q, like that in Section 5.1, which tells us how coins move about. The entry in the ith row and jth column tells us what proportion of coins from region i move to region j in one month. In the example in section 5.1 we had just two regions, the Netherlands (region 1) and the rest of the world (region 2), but we could instead have a model with more regions. If we wanted to make the most possible use of the data from the Euro Diffusion web-site we might end up having 70 regions, more than 60 parts of the Netherlands and Belgium and the other 10 countries. We might also want to have regions representing areas outside of Europe, or the piggy-banks (saving pots) of people in different countries, or even the place where all of the lost money goes. Suppose we have r of these regions and give each region a number from 1 to r. This makes Q an $r \times r$ matrix, with r^2 entries, none of which we know. We want to estimate Q using the data we have from the web-site.

Suppose that we consider m types of coin. These could be the 12 nationalities, in which case m would be twelve, or we could decide to group some nationalities together as in Section 5.1 in which we had two types, Dutch coins (type 1) and all others (type 2). We label each type with a number from 1 to m. We know or can find out in which region(s) each type of coin begins, on 1st January 2001. For example in Section 5.1 we knew that all of the Dutch coins began in the Netherlands while all of the non-Dutch coins began in the rest of the world. So for each type of coin we can make a vector, b, which tells us where those coins are. In the example we have $b(1) = (1, 0)$ which says that of the 1st type of coins (the Dutch ones) all are in region

1(the Netherlands) and none are in region 2 (the rest of the world), and $b(2) = (0, 1)$ which says that of the 2nd type of coins (the non-Dutch ones) none are in region 1 (the Netherlands) and all are in region 2 (the rest of the world). We then have m of these vectors $b(1), b(2), b(3), \ldots, b(m)$ (one for each type of coin) and each will have r entries.

We also have some measurements from the web-site. We call the number of measurements of coin type c in region i after t months $n(c, i, t)$. Suppose we measure the number of coins of each type in each region every month for s months. In the next sections we will discuss a number of possible estimation techniques.

Unfortunately all of these estimates inevitably run the risk of significant errors, due to the issues discussed in Section 3.

7.1. Maximum Likelihood Estimation*.** A common method in this type of problem is called *maximum likelihood estimation*. To use this we first write down the probability that we would see the measurements we have seen if we knew what Q was. If we had samples from all of the regions this turns out to be

$$p(Q) = \prod_{t=1}^{s} \prod_{i=1}^{r} N(i,t) \prod_{c=1}^{m} ((b(c)Q^t)_i)^{n(c,i,t)}$$

where

$$N(i,t) = \binom{\sum_{c=1}^{m} n(c,i,t)}{n(1,i,t)n(2,i,t)\cdots n(m,i,t)},$$

a multinomial coefficient. What we then do is find the matrix Q which maximises this probability.

This does look horrible, but fortunately we don't have to worry about $N(i,t)$. It is possible to show that the value of Q that maximises this probability also maximises:

$$\hat{p}(Q) = \prod_{t=1}^{s} \prod_{c=1}^{m} \prod_{i=1}^{r} ((b(c)Q^t)_i)^{n(c,i,t)}.$$

There are two things we do have to keep in mind when we do this. The entries of Q are proportions and are therefore all larger than or equal to zero. The entry in the ith row and jth column tells us what proportion of coins from region i move to region j, so if we sum all the entries is a row we should get the proportion of coins that go anywhere or stay in the same country, i.e. 1.

In formulas we need to

$$\text{maximise} \prod_{t=1}^{s} \prod_{c=1}^{m} \prod_{i=1}^{r} ((b(c)Q^t)_i)^{n(c,i,t)}$$

$$\text{subject to } Q(i,j) \geq 0, \text{for all } i \text{ and } j, \text{ and } \sum_{j=1}^{r} Q(i,j) = 1 \text{ for all } i$$

Generally this is a horrible problem which isn't easy to solve using mathematics and so we will turn to computers to do this numerically. This will typically be hard and slow.

7.2. An Iterative Approach*.** The "true" proportion of coins of type c in region i after t months is $(b(c)Q^t)_i$. A logical estimate for this is

$$p_t(c,i) = n(c,i,t) / \sum_{c'=1}^{m} n(c',i,t), \tag{2}$$

i.e. the proportion of coins of type c that we see in region i after t months.

One of the problems that we have with the estimation is that we do not have measurements from every region. If we did we could estimate the row vector $b(c)Q^t$ by the row vector $(p_t(c,1)\ p_t(c,2)\ p_t(c,3)\ \ldots\ p_t(c,r))$. Notice that if we stack up the row vectors $b(c)Q^t$ we get,

$$\begin{pmatrix} b(1)Q^t \\ b(2)Q^t \\ \vdots \\ b(m)Q^t \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & \ddots & \\ \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} Q^t = Q^t$$

Therefore the matrix,

$$M_t = \begin{pmatrix} p_t(1,1) & p_t(1,2) & p_t(1,3) & \cdots & p_t(1,r) \\ p_t(2,1) & p_t(2,2) & p_t(2,3) & \cdots & p_t(2,r) \\ & & & \ddots & \\ \cdots & \cdots & \cdots & & \cdots \\ p_t(m,1) & p_t(m,2) & p_t(m,3) & \cdots & p_t(m,r) \end{pmatrix}$$

this would give us a good estimate for Q^t, and we could find an estimate for Q by taking this matrix to the power of $(1/t)$. We could then combine the estimates from each timestep to give a final estimate for Q,

$$\frac{1}{s} \sum_{t=1}^{s} M_t^{1/t}.$$

One way we can deal with this missing data is to use an iterative algorithm. We begin with a guess for Q, call it Q_0, and gradually improve this. We begin with j (a counter) set to 0.

Step One: Let $p_t(c,i) = (b(c)Q_j^t)_i$ for all coin types and each of the regions i that we do not have data from.

Step Two: For each t, form M_t using Equation 2 for the regions we have data from and the estimates from Step One for the regions where we do not have data.

Step Three: Increase j by 1, let $Q_j = \frac{1}{s}\sum_{t=1}^{s} M_t^{1/t}$, and go to Step One.

This should converge to a reasonable estimate for Q, in a reasonable amount of time.

7.3. An Example*.** To illustrate the methods above we will apply them to a three state model, in which we take the Netherlands (region 1), Belgium (region 2) and everywhere else (region 3) as our regions. We also consider three types of euro coin, Dutch (type 1), Belgian (type 2) and the others (type 3). Therefore $b(1) = (1, 0, 0)$, $b(2) = (0, 1, 0)$ and $b(3) = (0, 0, 1)$, as all of the coins begin in their respective countries. For the maximum likelihood we consider two measurements, that of 1st February and that of 1st March. For the iterative approach we also consider 1st April.

The data we have from the Euro Diffusion web-site is that:

$$\begin{array}{ll} n(1,1,1) = 0.908 & n(1,2,1) = 0.041 \\ n(2,1,1) = 0.023 & n(2,2,1) = 0.888 \\ n(3,1,1) = 0.069 & n(3,2,1) = 0.071 \\ n(1,1,2) = 0.871 & n(1,2,2) = 0.06 \\ n(2,1,2) = 0.031 & n(2,2,2) = 0.807 \\ n(3,1,2) = 0.098 & n(3,2,2) = 0.133 \\ n(3,1,2) = 0.098 & n(3,2,2) = 0.133 \\ n(1,1,3) = 0.823 & n(1,2,3) = 0.064 \\ n(2,1,3) = 0.045 & n(2,2,3) = 0.807 \\ n(3,1,3) = 0.132 & n(3,2,3) = 0.129 \end{array}$$

Maximum Likelihood

In this case we don't have measurements for the numbers of coins outside the Netherlands and Belgium. Therefore for the maximum likelihood we need to

$$\text{maximise } \prod_{t=1}^{2}\prod_{c=1}^{3}\prod_{i=1}^{2}((b(c)Q^t)_i)^{n(c,i,t)}$$

$$\text{subject to } Q(i,j) \geq 0, \text{for all } i \text{ and } j, \text{ and } \sum_{j=1}^{r} Q(i,j) = 1 \text{ for all } i$$

If we write this out, filling in the values we know and writing

$$Q = \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix},$$

this becomes the problem of maximising,

$$q_{11}^{0.908} q_{12}^{0.041} q_{21}^{0.023} q_{22}^{0.888} q_{31}^{0.069} (q_1 1^2 + q_{12}q_{21} + q_{13}q_{31})^{1.742} \times$$

$$(q_{12}q_{21} + q_{21}q_{22} + q_{23}q_{31})^{0.031} q_{32}^{0.071} (q_{11}q_{12} + q_{12}q_{22} + q_{13}q_{32})^{0.06} \times$$
$$(q_{12}q_{21} + q_{22}^2 + q_{23}q_{32})^{0.807} (q_{11}q_{31} + q_{13}q_{32} + q_{31}q_{33})^{0.098} \times$$
$$(q_{12}q_{31} + q_{22}q_{32} + q_{32}q_{33})^{0.133}$$
$$\text{subject to } q_{ij} \geq 0, \text{for all } i \text{ and } j, \text{ and } \sum_{j=1}^{r} q_{ij} = 1 \text{ for all } i$$

This really hard, even for a computer, but a preliminary numerical investigation suggests that this is maximised by a matrix close to:

$$\begin{pmatrix} 0.975 & 0.025 & 0 \\ 0.02 & 0.98 & 0 \\ 0.4 & 0.6 & 0 \end{pmatrix}.$$

The Iterative Approach

The iterative approach is easily programmed and the estimates converge quickly to:

$$\begin{pmatrix} 0.924326 & 0.0172391 & 0.0584345 \\ 0.0309947 & 0.903493 & 0.0655122 \\ 0.0578321 & 0.065527 & 0.876641 \end{pmatrix}.$$

8. A Continuous Model for Euro Movement**

Coins of neighbouring countries mix at boundary cities. Coins of various origins mix with each other at tourist places, at airports, in highway restaurants, gasoline stations etc. This is a typical *continuous* diffusion phenomenon like the mixing of different gases or the spread of smoke or fragrance. The process of continuous diffusion frequently occurs naturally and plays an important role in many applications. The main cause for diffusion is gradient profile, i.e. different concentrations in different places. The coefficient of diffusion (denoted by D) is the measure of the continuous diffusion.

In the long-time limit, and in the absence of sources (which we think is quite a good approximation), it is clear coins of different origins will have roughly the same proportion everywhere within the euro-zone. The natural question to ask is, over what timescale is this going to happen?

In our problem we may be able to use the country-of-origin breakdown of euro coins (which are identical for practical purposes) to deduce information on coin transport (and hence certain aspects of people's economic behaviour) which would be very difficult to measure by other means.

8.1. Model*.** In this section we seek to design a continuous model which, with the least amount of complexity, will best fit the available data (the first three months of 2002) and will make useful predictions of its future behaviour.

As the basis of this model, we shall make the assumption that people are indifferent to origin of coin, i.e., they treat coins of different countries equally. Also, in view of observation **1** in Section 3.1, we shall consider a fixed denomination, say, one euro, in our model.

In this model, we consider two different processes that contribute to the dispersal of coins. The first process is a local one, which arises from people carrying out their daily activities: going to the market, the bank, restaurants, etc. We model this by a diffusion equation, with a diffusion constant $D(x,t)$. The second process is inherently non-local, arising from medium- and long-distance travels. We model this process, a priori, by an integral term. In addition to these, we also need to include sources (bank issues) and losses.

We note that, in reality, there is likely a continuous spectrum of "coin transport scales"; nevertheless, for conceptual purposes and modelling feasibility, we assume that there is a "separation of scale" between local and non-local processes.

Let $m^{(n)}(x,t)$ denote the (normalised; see below) density of 1-euro coins of country n at location x at time t, where $x \in \Omega =$ euro-land and where we take 2002 January 1 to be $t = 0$. A general governing equation for $m^{(n)}(x,t)$ then reads

(3)
$$\frac{\partial}{\partial t} m^{(n)}(x,t) = f^{(n)}(x,t) - \nu(x) \cdot m^{(n)}(x,t) + \nabla \cdot (D(x,t) \nabla m^{(n)}(x,t)) + \int_\Omega K(x,z,t)\, m^{(n)}(z,t)\, \mathrm{d}z.$$

Here the terms are

$f^{(n)}(x,t) =$ source density such as bank,

$\nu(x) =$ rate of loss of coins, assumed independent of t,

$D(x,t) =$ diffusion coefficient,

$K(x,z,t) =$ integral kernel for non-local transport.

We note that if the loss term $\nu(x)$ is independent of x and t, it can be dropped provided that one rescale: $m^{(n)}(x,t) \mapsto \mathrm{e}^{-\nu t} m^{(n)}(x,t)$, this being irrelevant if one only considers the *ratios* (such as percentages) of coins.

Let $e(x)$ denote the population density at point x, assumed to be constant in time. To a very reasonable approximation, we may take the initial

conditions to be

$$m^{(n)}(x,0) = \begin{cases} e(x) & x \text{ is in country } n \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The possibility (not considered further here) that the coin/population ratio may vary by country can be taken care of by normalising $e(x)$ to account for this factor.

Let us now construct a general model for the integral kernel $K(x,z,t)$. As before, we assume that long-distance travel does not change the population density $e(x)$. Let $p(x,z,t)$ be the probability that an inhabitant of x visits z. Then the integrand $K(x,z,t)m^{(n)}(z,t)$ in (3) can be modelled as,

$$\begin{aligned} K(x,z,t)\, m^{(n)}(z,t) = {} & \varepsilon_0(x)p(x,z,t)e(x)[m^{(n)}(z,t) - m^{(n)}(x,t)] \\ & + \varepsilon_0(z)p(z,x,t)e(z)[m^{(n)}(z,t) - m^{(n)}(x,t)]. \end{aligned} \tag{5}$$

Here the first term corresponds to residents of x travelling to z and the second to residents of z travelling to x; ε_0 is a numerical coefficient related to the number of coins people carry on long-distance trips. To simplify this further, we may assume (somewhat unrealistically) that "tourists" of different origins are distributed equally among all popular destinations; this would allow us to write $p(z,t)$ on the first line, measuring how popular z is, and $p(x,t)$ on the second.

In what follows, we take $D(x,t)$ to be inversely proportional to the population density $e(x)$, which is effectively constant over the timescales considered here. The reason for this is that people tend to travel farther to the "local shop" in sparsely populated areas than in urban centres, and, assuming that city and rural shops have equal number of customers, this distance is proportional to $e(x)^{-1/2}$. Thus,

$$D(x,t) = D_0/e(x), \tag{6}$$

where the constant D_0 is to be determined from a fit with the observed data.

Instead of the full integral kernel $K(x,z,t)$ in (3), we shall model the non-local effects by "connecting" a small number (10 in this work) of "major airports"; by this we mean to account for all types of long-distance travels, not only by air. We denote each airport by $L_i \subset \Omega$, having a fixed area δ (one element in the numerical model below). For each airport L_i, we assign the number of people served by it, $\rho(L_i)$, and we assume that the number of passengers travelling from airport L_i to L_j is a constant ε' times $\rho(L_i)\rho(L_j)$. The coin transport due to this process can then be modelled by adding the following term to (3) above: For $x^* \in L_i$ and $z^* \in L_j \neq L_i$,

$$\left.\frac{\partial}{\partial t}\right|_{\text{nl}} m^{(n)}(x^*,t) = \frac{\varepsilon}{\delta} \sum_{L_j} \int \rho(x^*)\, \rho(z^*) \left[m^{(n)}(z^*,t) - m^{(n)}(x^*,t)\right] \mathrm{d}z^* \tag{7}$$

where the constant ε (which like D_0 is to be determined from a fit with the data) is related to ε'. The value of δ is irrelevant if it is sufficiently small; one can take the limit $\delta \to 0$ if desired.

Since the population density data is readily accessible (e.g., from a school atlas), our model so far only has two parameters, D_0 and ε, that need to be determined. Noting that we can scale time by D_0, as far as the *qualitative* behaviour of our model so far is concerned, only one parameter matters: the ratio ε/D_0.

The model can also be easily modified (by changing the diffusion constant) to investigate the economic importance of national borders, e.g, whether people living close to the border are more or less likely to cross the border to shop.

8.2. Numerics and Fit****.** We ran numerical simulations of a "finite-element" version of the model described above, consisting of 100 cells, each representing 3 million people, which are connected in a rough approximation to the actual geography. The variables are $m_i^{(n)}$, with $n \in \{\text{NL, BE, DE, } \cdots\}$ and $i = 1, \cdots, 100$. The diffusion term

$$\nabla \cdot (D_0 e(x) \nabla m^{(n)}(x,t))$$

is approximated for cell i by $\alpha D_0 \sum_j (m_j^{(n)} - m_i^{(n)})$ where j ranges over the neighbouring cells and α is a numerical constant.

Due to the discreteness of the numerical model, the long-time limits of Dutch and Belgian coins are 5% and 4%, respectively. Turning to the Figure 2 below, let us first consider the percentage of Dutch coins in the Netherlands (dotted line, right scale) in the diffusion-only scenario. As may be expected, the fraction of the local coins decrease monotonically over time towards its long-time equilibrium; this is also true for the airport-assisted cases (with faster decay rate) not shown here.

Next, we consider the percentages of Belgian coins in the Netherlands for three values of ε/D_0: 0 (solid line, left scale), 3×10^{-3} (long dashed line) and 0.1 (short dashed line). In all of these cases we see that the fraction of Belgian coins first increases before decaying to its long-time limit, with the "overshoot" being greater in the absence of long-distance travel. These are both to be expected, intuitively, due to the proximity of Belgium to the Netherlands.

To actually determine which of these qualitative curves best describe the reality, as well as to determine the actual timescales, we need to turn to the observed data. This turns out to be quite problematic, as we mentioned in Section 3, and for the above plot we have made a *very tentative* estimate for the timescale (in years).

This leaves us with several possibilities: First, assuming that our data accurately reflects the actual distribution of coins, the source terms over

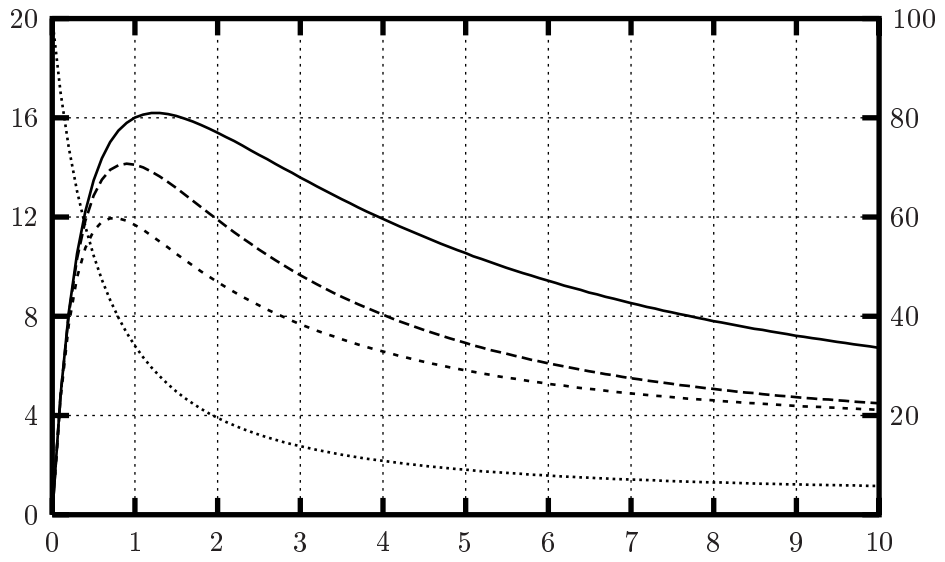


FIGURE 2. Percentages of Dutch and Belgian coins in the Netherlands (see text for details)

the timescale concerned may be important; this scenario can presumably be easily resolved by a simple check with the central banks. Second and still assuming the accuracy of our data, we may be forced to consider processes other than the mixing of coins, as no mixing model (regardless of the details) would produce an increase in the fraction of local coins; this possibility is difficult and, in our opinion, unappealing. Finally, the fluctuating tendencies may be caused by systematic bias in the data collection procedure; this possibility, too, is difficult to establish.

9. Conclusions*

In roughly a year half our wallet will be filled with foreign coins!

Acknowledgements

We thank the Euro Diffusion web site group for setting up the web site that is indispensable for the collection of data. We thank Jeanine Kippers from DNB (the Dutch national bank) for supplying additional information. We thank Richard Gill and Andreas Kyprianou for advice on the statistical part of the project the models respectively.

pp. 59 – 76

CHAPTER 4

Roses are unselfish: a greenhouse growth model to predict harvest rates

Onno Bokhove, Johan Dubbeldam, Philipp Getto, Bas van 't Hof, Nick Ovenden, Derk Pik, Georg Prokert, Vivi Rottschäfer, Dick van der Sar.

Abstract. We consider the question of how rose production in a greenhouse can be optimised. Based on realistic assumptions, a rose growth model is derived that can be used to predict the rose harvest. The model is made up of two constituent parts: (i) a local model that calculates the photosynthetic rate per area of leaf and (ii) a global model of the greenhouse that transforms the photosynthesis of the leaves into an increase in mass of the rose crop. The growth rate of the rose stems depends not only on the time-dependent ambient conditions within the greenhouse, which include temperature, relative humidity, CO_2 concentration and light intensity, but also on the location and age distribution of the leaves and the form of the underlying rose bush supporting the crop.

Keywords: Rose production model, advection equation, stem density function, global and local leaf photosynthesis

1. Introduction

The production of roses has become more competitive and commercialised over the last few decades. While the rose grower's own experience remains the key to producing a large rose harvest, qualitative and quantitative modelling of the biochemical processes in rose plants is becoming increasingly important in optimising rose production even further.

In this article, we develop a simplified mathematical model for rose production to predict the total mass of rose crop produced per square metre of greenhouse per week, depending on the climatic conditions inside the greenhouse. The goal of our model is to tune these conditions in such a way that the harvest of roses is maximised.

Rose stems grow by assimilating CO_2 from the air. This is done in the leaves and is called photosynthesis. In the greenhouse, rose stems are cut once they have reached a certain length and, when a rose is harvested, it (obviously) stops assimilating. The CO_2-assimilation and, therefore, the growth of the roses is influenced by several environmental factors. Some of

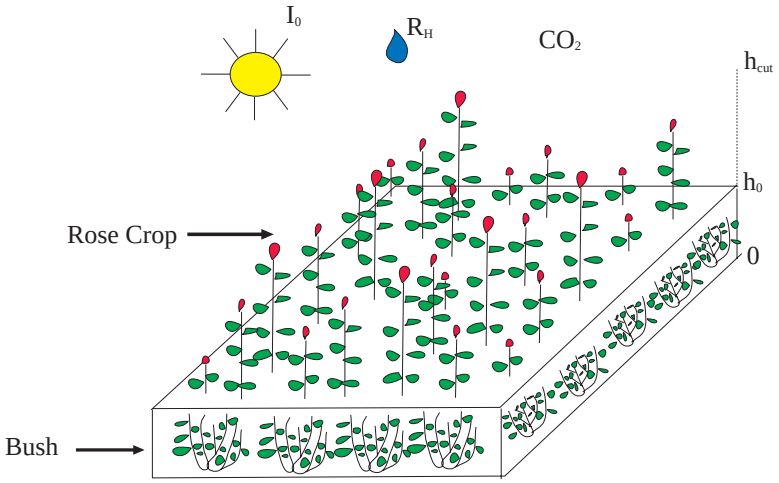


FIGURE 1. A rose plant is divided into a bush part below supporting the crop above, which consists of the stems to be harvested when they reach height h_{cut}.

these can be controlled by the rose-grower, for example by using heaters, opening or closing the windows and putting up blinds for shade. These actions, in turn, alter the CO_2-concentration in the air C_a (by ventilation), the relative humidity R_H, the temperature in the greenhouse T_a, and the light intensity I_0 (see figure 1).

To model the rose plants in the greenhouse, we assume that the plants can be divided into two constituent parts. The lower part is the 'bush' that supports the upper part or the 'crop', see figure 1. We assume that the bush has height h_0 and that it is not harvested but has leaves that assimilate. The crop on the other hand consists of stems, each with a rose bud on top, that are harvested once they reach a certain height h_{cut}. They are then cut at the level $h = h_0$ so that the harvested stems all have length $h_{cut} - h_0$. As mature plants are cut, new stems begin to grow from the top of the bush appearing at a rate proportional to the total photosynthetic rate in the greenhouse. We ignore at present the part of the acquired photosynthetic energy that is used for maintenance and storage, and assume that the photosynthesis in the crop and bush is entirely used to increase the mass of stems in the crop.

The model can be split into two distinct levels. The first level is concerned with the biological process inside a leaf, in other words the local photosynthesis. The other level handles modelling the greenhouse as a whole and here the global CO_2-assimilation of all the rose plants and the resultant harvest are taken into account. We make realistic and sometimes simplifying assumptions based on biological observations. As the photosynthesis in a leaf depends on the age of the leaf, we have to know where the young and

old leaves are positioned on a rose plant. For this reason we assume that stems grow vertically, and that new leaves grow at the top. Thus, we find the older leaves on the lower part of the stem and the younger ones near the bud. We also suppose that the leaves, and therefore the leaf area, are distributed uniformly along the stem. In other words, the leaf area of each stem is proportional the stem's length.

One of the essential assumptions on which the global model is built is the so-called 'unselfishness principle'. The principle says that any energy gained by photosynthesis of a single leaf, either located on a stem or within the bush, contributes equally to the growth of all the stems, large and small. Hence, a taller stem, which has more leaves, will assimilate more CO_2 and produce more energy than a shorter one but their combined energy will be shared equally between them. As a result, every stem grows at the same speed, independent of its own photosynthetic rate. This assumption reflects both real data and the observation that a single rose plant, possessing a number of rose stems of differing heights, acts as a single entity; in this way young stems can develop quickly, even though they do not possess a large leaf area.

Based on the principles stated above, a global rose production model has been constructed that resolves the rate of change in height distribution of rose stems (see section 2). The state of the crop at any given instant of time is uniquely determined by a stem density function $d(h, t)$, describing the number of stems per area of greenhouse as a function of height h and time t. The dynamics of d are given by a linear advection equation and the unselfishness principle implies that the relevant advection speed is a function of time only. The growth or advection rate is found by calculating the total net photosynthesis of a square metre of rose plants. This is determined by adding the local photosynthetic contribution from each leaf in the rose crop and rose bush. As the leaf's local photosynthetic rate depends both on its age and on the amount of light it receives (affected by shading from higher leaves), an ability to model the age and height distribution of leaves is important. The total photosynthesis produced per square metre follows, in turn, by integration of the local photosynthesis rate over all the leaf ages and heights in both the rose crop and the bush, weighted by the leaf area distribution. In our model, the leaf distributions of the rose crop and the bush are treated separately. Indeed, two different approaches to model the total photosynthesis of the bush are given with their respective advantages and disadvantages.

To close the global rose production model, we must also model the local photosynthesis to obtain the photosynthetic rate of a single leaf as a function of height and leaf age; this is done in section 3. The local model used is a simplified version of the models developed by Harley *et al* (1992) and Kim

and Lieth (2001). Of course, the global model can be equally well coupled to other local models of leaf photosynthesis.

Given the necessary simplifications, several proportionality constants appear as parameters in the global model. These parameters must be determined either by direct measurement of the rose plants or by fitting them to given harvest data. In section 4, we describe how the estimation of these parameters can be accomplished. The model should then, in principle, be able to aid the rose grower to optimise the weekly amount of harvested roses. However, adequate testing of the model by numerically fitting the parameters to real data is still in progress.

The outline of the article is the following. In section 2, the global mathematical model for rose growth is developed. The simplified local leaf model used for photosynthesis is then described in section 3. The combination of the local and the global model contains seven unknown parameters. In section 4, we argue how these parameters can be estimated by direct measurements on a rose plant, and also from the harvest data. Finally, we summarize the theory and discuss directions for future work in section 5.

2. Global rose production model

The rose plants growing in a greenhouse can be separated into two parts: the rose stems, which are harvested, and the rose 'bush' below that contains the body of the rose plants supporting each individual stem (see figure 1). The rose bush is not harvested and lies between $h = 0$ and $h = h_0$. Its leaves assimilate energy that contributes to the growth of the crop. The rose stems growing vertically out of the rose bush are, at a given time, of different heights. As each rose plant consists of a mixture of mature and young rose stems, rose stems of different heights are taken to be distributed evenly throughout the greenhouse.

2.1. A representation of the greenhouse. The mathematical model for rose production is based on the following assumptions, which agree with the rose growers experience and simplify the mathematical modelling:

i. The principle of unselfishness: roses are unselfish, meaning that any biomass gained through photosynthesis of a single leaf is equally distributed among all the stems, large and small. In other words at any fixed time, every stem grows at the same speed $v(t; d)$, independent of its own photosynthesis production.
ii. All stems grow vertically.
iii. All stems start at height h_0. The appearance (sprouting) of new stems depends on the climatic conditions and is therefore assumed to be proportional to the photosynthetic rate.
iv. Stems are cut at a height h_0 and harvested when they exceed height $h > h_{cut} > h_0$.

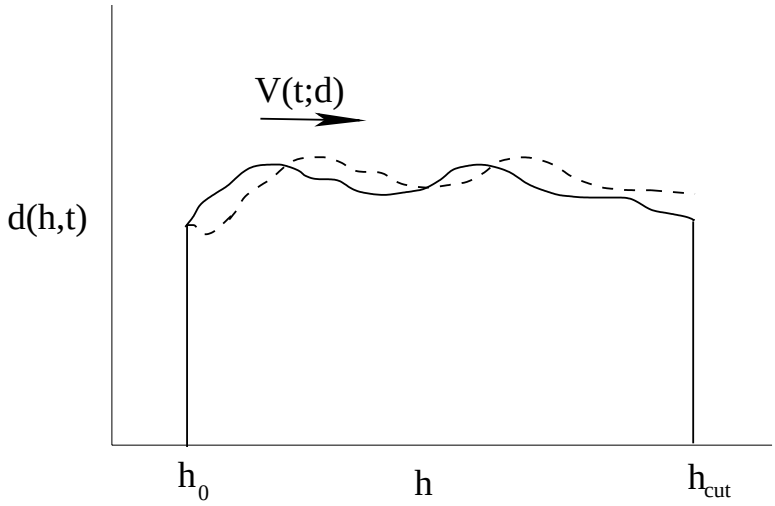


FIGURE 2. The state of the rose crop can be expressed by a stem density function $d(h,t)$ representing the distribution of stems of differing heights per square metre of greenhouse area. The dynamics of $d(h,t)$ is governed by an advection equation and the unselfishness principle implies that the advection speed is independent of h.

v. All new leaves appear at the top of a stem, implying that the leaves closest to the rose bud are the youngest.
vi. Both mass and leaf area of a stem are proportional to the length of the stem and uniformly distributed along it.

For the dimensions of all occurring quantities and constants we refer to tables 1 and 2.

The state of the greenhouse is given by a stem density function $d = d(h,t)$ for $h > h_0$ such that the number of stems of lengths between h and $h + \mathrm{d}h$ per square metre of greenhouse is $d(h,t)\mathrm{d}h$ (see figure 2).

2.2. The advection equation for the stem density function d. The unselfishness principle (assumption i) implies that this density function is advected by a growth rate $v = v(t;d)$, which is independent of h and will be determined later. Hence, we obtain

$$\partial_t d + v\,\partial_h d = 0, \tag{1}$$

where $\partial_t = \partial/\partial t, \partial_h = \partial/\partial h$ denote partial derivatives with respect to time t and height h.

The boundary condition at $h = h_0$ represents the creation of new stems from the rose bush (provided there is enough light such that $v(t;d) > 0$). By assumption iii, the appearance of new stems at h_0 is proportional to the

Symbol	*Quantity*	*unit*
h	height	m
h_0	starting height of stems	m
h_{cut}	cutting height	m
d	stem density distribution	m^{-3}
v	growth velocity	$m\ s^{-1}$
P_{net}	total net photosynthesis rate	$\mu mol\ m^{-2}s^{-1}$
H	harvest rate	$kg\ m^{-2}s^{-1}$
M	crop mass	$kg\ m^{-2}$
N	number of stems	m^{-2}
ρ	leaf area density	m^{-1}
q	age density distribution	$m^{-1}s^{-1}$
a	leaf age	s
I	photosynthetic photon flux density	$\mu mol\ m^{-2}s^{-1}$
T_{max}	average growth time of rose (6 to 8 weeks)	s
τ	Length of growing season (6 months)	s

TABLE 1. Dimensional quantities used in the global greenhouse model

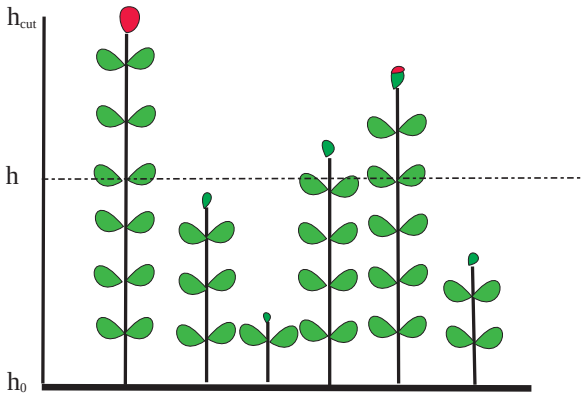


FIGURE 3. At a certain height h all rose stems of heights greater than h contribute to the leaf area density $\rho(h)$. Smaller rose stems do not.

rate of photosynthesis $P_{net}(t;d)$:

$$d(h_0,t) = k_2\, P_{net}(t;d). \tag{2}$$

This net photosynthetic rate P_{net} represents the biochemical intake or loss of CO_2 per square metre of greenhouse and will be determined later.

Assumption iv implies that the harvest rate $H(t)$ per square metre of greenhouse is given by

$$H(t) = k_3\, v(t;d)\,(h_{cut} - h_0)\, d(h_{cut},t), \tag{3}$$

where k_3 is the mass of a rose per unit length. It is straightforward to alter the model to a situation where the roses larger than h_{cut} are harvested at discrete times and not continuously (see Appendix A).

2.3. Determining the growth speed v. The mass of crop in the greenhouse is again proportional to k_3 and the first moment in h of the stem density function

$$M(t) = k_3 \int_{h_0}^{h_{cut}} (h - h_0)\, d(h,t)\, \mathrm{d}h, \tag{4}$$

where the stem density is properly weighed by the stem length $(h - h_0)$ in order to obtain the mass (that is, following assumption vi). Differentiating (4) and integrating the r.h.s. by parts gives

$$\begin{aligned}\frac{dM}{dt} &= -k_3\, v(t;d) \int_{h_0}^{h_{cut}} (h - h_0)\, \partial_h d(h,t)\, \mathrm{d}h \\ &= k_3\, v(t;d) \int_{h_0}^{h_{cut}} d(h,t)\, \mathrm{d}h - k_3\, v(t;d)\, (h_{cut} - h_0)\, d(h_{cut},t) \\ &= k_3\, v(t;d)\, N(t) - H(t), \qquad (5)\end{aligned}$$

where the integral over $d(h,t)$ is the total number of stems per square metre and given here as

$$N(t) = \int_{h_0}^{h_{cut}} d(h,t)\, \mathrm{d}h. \tag{6}$$

The net photosynthetic rate P_{net} is proportional to the change in productive mass plus the harvest rate, and by using (5) we can subsequently obtain the growth rate $v(t;d)$:

$$k_1\, P_{net}(t;d) = \frac{dM}{dt} + H(t) = k_3\, v(t;d)\, N(t) \tag{7}$$

$$\iff \quad v(t;d) = \frac{k_1\, P_{net}(t;d)}{k_3\, N(t)}. \tag{8}$$

2.4. The leaf density functions. In order to calculate P_{net} from the local photosynthesis model, we require information about the distribution of leaf area, leaf age and light intensity (photon flux density) with respect to height.

The density function $\rho(h,t)$ is defined so that $\rho(h,t)\,\mathrm{d}h$ yields the area of leaves with stem lengths between h and $h + \mathrm{d}h$ per square metre of greenhouse. It is related to $d(h,t)$ by

$$\rho(h,t) = k_4 \int_{h}^{h_{cut}} d(\zeta,t)\, \mathrm{d}\zeta. \tag{9}$$

Symbol	*Constant*	*unit*
k_1	mass production per CO_2 intake	$kg\ \mu mol^{-1}$
k_2	birth rate of stems	$s\ \mu mol^{-1}\ m^{-1}$
k_3	stem mass per unit length	$kg\ m^{-1}$
k_4	leaf area of a stem per unit length	m
k_5	inverse average growth velocity	$s\ m^{-1}$
k_6	light absorption coefficient	–
k_7	ratio of leaf area in bush to leaf area in crop (rose stems)	
k_8	contribution of bush second model	–

TABLE 2. Parameters in the global greenhouse model

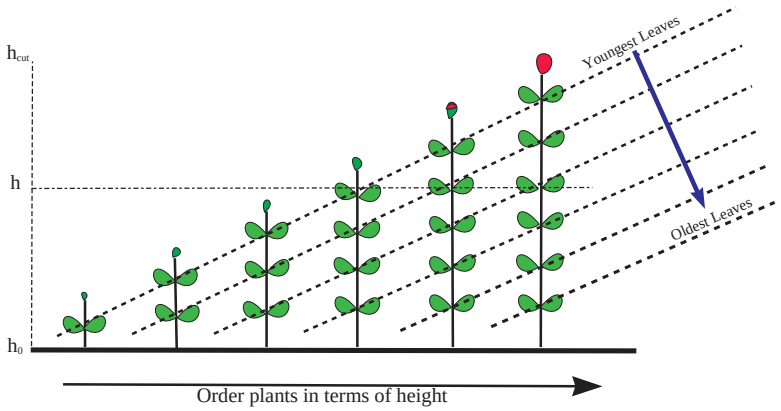


FIGURE 4. Placing the rose stems in order of size is a useful guide to calculate the age distribution of leaves at a given height h. The sketch depicts the situation where the growth velocity is approximated to be constant to simplify the age density distribution.

The integration limits are chosen as h and h_{cut} because all the rose stems of heights greater than h contribute to the leaf area density at h; figure 3 provides a more graphical explanation why.

The age density distribution $q(t, a, h)$ is defined so that $q(t, a, h)\,\mathrm{d}h\,\mathrm{d}a$ yields the leaf area of age between a and $a+\mathrm{d}a$ located between the heights h and $h+\mathrm{d}h$ per square metre of greenhouse. Under the simplifying assumption that the age of a leaf is proportional to its distance from the top of the stem we find that the leaves of age a at height h belong to stems of height $h + \frac{a}{k_5}$ (see figure 4). The parameter $k_5 = T_{max}/(h_{cut} - h_0)$ is the inverse average growth rate of a typical stem. The time T_{max} indicates the average total growth time of a stem from its first appearance on the bush to harvest; T_{max} differs for each type of rose and each growth season. Thus

$$q(t, a, h) = c\,d\left(h + \frac{a}{k_5}, t\right).$$

Using the fact that $\rho(h,t) = \int_0^{T_{max}} q(t,a,h)\mathrm{d}a$ we can determine the proportionality constant c to obtain

$$q(t,a,h) = \frac{k_4}{k_5}\, d\left(h + \frac{a}{k_5}, t\right). \tag{10}$$

Of course, the assumption of a constant growth velocity for the branches is in contradiction with our model assumptions. We use it, however, for the determination of the age distribution in order to avoid the highly complex nonlinear integrodifferential equation for v which would result if v itself would be used there. The assumption can be justified by the fact that T_{max} is large compared to the time scale on which the photosynthesis varies.

2.5. The light penetration. The top leaves of the tallest rose stems receive all the light available. However, the amount of light reaching the lower leaves of the mature plants and of the newer stems is diminished by the amount of leaf coverage above. The isotropic nature of the greenhouse means that all leaves at the same height have approximately the same amount of shade. The change in light intensity $I(h)$ as function of h is thus taken to be proportional to $\rho(h)$ and $I(h)$, leading to

$$\frac{dI(h)}{dh} = k_6\,\rho(h)\,I(h), \quad I(h_{cut}) = I_0, \quad \Longleftrightarrow \quad I(h) = I_0\, e^{-k_6 \int_h^{h_{cut}} \rho(\zeta)\,\mathrm{d}\zeta}. \tag{11}$$

The assumed age distribution of the leaves of the rose stems and the change in light intensity at each height now enable us to calculate the net photosynthesis produced by the rose crop per square metre of greenhouse, as follows

$$P_{crop}(t;d) = \int_{h_0}^{h_{cut}} \int_0^{T_{max}} q(t,a,h)\,P(t,a,h)\,\mathrm{d}a\,\mathrm{d}h. \tag{12}$$

Here $P(t,a,h)$ is the local photosynthesis rate per unit area for a leaf at height h and age a under given exterior climatic conditions (that is, temperature, relative air humidity, light intensity, and CO_2-concentration) at time t, as predicted by the local leaf model described in section 3.

2.6. The photosynthesis in the bush. While P_{crop} represents the major source of biomass for the roses in the greenhouse, the rose bush below h_0 also contains leaves and produces an additional seasonally-varying contribution to the net growth rate. We take two different approaches in modelling the rose bush, although other models are possible.

For the first approach it is assumed that the number of leaves in the rose bush, $h < h_0$, is some given ratio k_7 of the number of leaves in the crop above, for a given type of rose in a given season. These leaves within the bush are taken to be uniformly distributed between $h = 0$ and $h = h_0$. Furthermore, it is assumed that the leaves' ages are distributed uniformly

throughout the bush from newly created leaves to leaves roughly as old as the length of an entire growing season τ; here the season is assumed to last roughly six months (winter and summer). From these assumptions, the leaf area density within the bush can be written as

$$\rho(h < h_0) = \frac{k_7}{h_0} \int_{h_0}^{h_{cut}} \rho(\zeta) \mathrm{d}\zeta. \tag{13}$$

The amount of light reaching these bush leaves can then be determined from the solution to equation (11) for $0 < h < h_{cut}$. Furthermore, the uniform age and height distribution of leaves in the bush leads to the expression

$$q_{bush}(t, a, h) = \frac{k_7 \int_{h_0}^{h_{cut}} \rho(\zeta) \mathrm{d}\zeta}{h_0 \ \tau}, \tag{14}$$

for the age density distribution, which is constant in both a and h. The net rate of photosynthesis of the bush can subsequently be found, as for the crop, by integration over all leaf ages and heights,

$$\begin{aligned} P^{(1)}_{bush}(t; d) &= \int_0^{h_0} \int_0^{\tau} q_{bush}(t, a, h)\, P(t, a, h)\, \mathrm{d}a\, \mathrm{d}h \\ &= \frac{k_7 \int_{h_0}^{h_{cut}} \rho(\zeta) \mathrm{d}\zeta}{h_0 \ \tau} \int_0^{h_0} \int_0^{\tau} P(t, a, h)\, \mathrm{d}a\, \mathrm{d}h. \end{aligned} \tag{15}$$

In the second approach, the above model of the bush is simplified. Now, we assume that the leaves in the bush all have a mean age $\tau/2$ and a mean height $h_0/2$; again τ is the length of an entire growing season. Introducing another constant k_8 to represent the leaf area in the bush, we obtain

$$P^{(2)}_{bush}(t) = k_8\, P(t, \tau/2, h_0/2) \tag{16}$$

for the net rate of photosynthesis of the bush.

Both of these approaches to model the bush have their advantages and disadvantages. The first approach is more realistic compared to the second one since the bush is taken to have a similar leaf distribution as the crop. However, this model breaks down when all of the crop is harvested at the same time. In that case, the parameter k_7 limits to infinity and renders the bush model invalid. This scenario is realistic for certain rose types, like the variety "Sweet Unique", where all the roses are harvested at once particular instant. In addition, determining the parameter k_7 from the greenhouse data or relating it to the other parameters is non-trivial, so that implementation of the model may be more difficult compared to the second approach.

In the second approach on the other hand, the structure of the bush is simplified too much. More realistically, the bush consists of leaves with different ages and height like in the first approach. An advantage of this approach is, however, that k_8 can be estimated more readily from the harvest data, see section 4.

Finally, the total net photosynthetic rate is the sum of the net crop photosynthesis from (12) and the net bush photosynthesis from either (15) or (16), depending on the approach taken for modelling the bush, leading to

$$P_{net}(t;d) = P_{crop}(t;d) + P_{bush}(t;d). \tag{17}$$

In order to use our model to predict rose production, seven parameters k_i with $i = 1, \ldots, 6$ and either k_7 or k_8 must be estimated, for each rose type and for each season, using real harvest data with corresponding climate information measured inside the greenhouse.

3. Local leaf model for photosynthesis

It is important to emphasize that the production model described in section 2 is closed once only we have a model for the local photosynthesis in a leaf. Presently, we will use a version of the photosynthesis rose leaf model of Harley *et al.* (1992) and Kim and Lieth (2001).

Following Harley *et al.* (1992) and Kim and Lieth (2001), the photosynthetic rate in a unit area of leaf at height h and with age a, is given by

$$P(t,a,h) = min\{A_v, A_j\} - R_d. \tag{18}$$

Here A_v and A_j are the rate of Rubisco limited photosynthesis and the rate limited by RuBP regeneration respectively, while R_d is a threshold CO_2 consumption or dark respiration, for example due to losses at night, which we take constant at $R_d = 0.82\,\mu mol/(m^2\,s)$ (cf. Hartley *et al.*, 1992). The existence of the dark respiration term R_d implies that $P(t,a,h)$ can be less than zero at height h and a. However, any local losses can be compensated by a positive global photosynthesis rate elsewhere due to the unselfishness principle. Both P_{net} and $P(t,a,h)$ are expressed in terms of a CO_2-rate per square metre of greenhouse, which is $\mu mol\ CO_2/(m^2\,s)$.

The photosynthesis rates A_v and A_j depend on the intercellular CO_2-concentration C_i, as is shown in figure 5: the photosynthesis stops in conditions of too little intercellular CO_2 (i.e. if $C_i < \Gamma_*$). For increasing values of C_i, the photosynthetic rate allowed by RuBP regeneration increases faster than the Rubisco limited photosynthesis rate, but the latter attains a higher value V_{cmax} than the former.

The formula for Rubisco limited photosynthesis A_v is given by (Kim and Lieth, 2001)

$$A_v = V_{cmax}\frac{C_i - \Gamma_*}{C_i + \kappa} \qquad \text{with} \quad V_{cmax} = V_m\, g(T)\, f(a), \tag{19}$$

where V_{cmax} is the maximum rate of carboxylation and C_i the intercellular CO_2-concentration, and $g(T)$ and $f(a)$ represent the dependence on leaf temperature T and leaf age a. The remaining unknowns are constants,

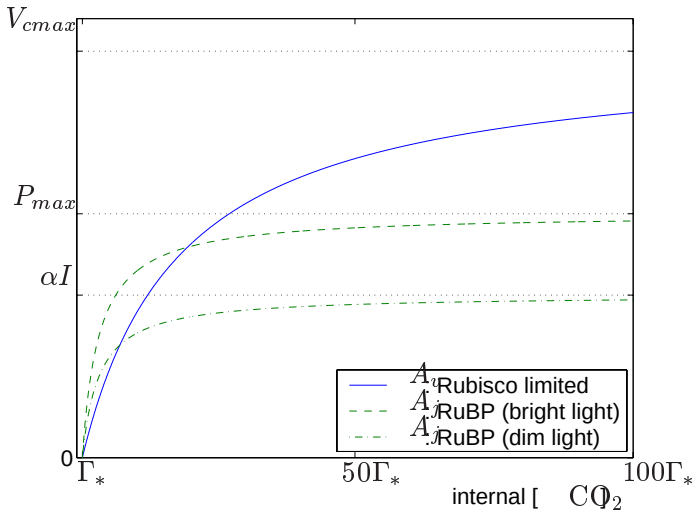


FIGURE 5. The C_i-dependence of photosynthesis rates A_v and A_j.

defined and given in table 3; see also Harley *et al.* (1992) and Kim and Lieth (2001). The RuBP limited photosynthetic rate A_j is

$$A_j = \frac{C_i - \Gamma_*}{4\,(C_i + 2\,\Gamma_*)}\,J \tag{20}$$

with the potential electron transport rate J given by

$$J = \frac{8\,\alpha\,I\,P_{max}}{\alpha\,I + P_{max} + \sqrt{(\alpha\,I + P_{max})^2 - 4\,\alpha\,I\,P_{max}\,\theta}} \tag{21}$$

with

$$P_{max} = P_m\,g(T)\,f(a). \tag{22}$$

Here $I = I(h)$ is the photosynthetic flux density given in (11) at height h above the ground. We note that when $P_{max} \gg \alpha\,I$ the potential rate $J \approx 4\,\alpha\,I$, which means that when $\alpha\,I$ is sufficiently small the production is limited by the lack of light. Inversely, when $P_{max} \ll \alpha\,I$ the potential rate $J \approx 4\,P_{max}$, which means that when $\alpha\,I$ is sufficiently large, any increase in the amount of light has no additional influence on the rate of photosynthesis.

The temperature dependence $g(T)$ and age dependence $f(a)$ of the photosynthesis rates V_{cmax} and P_{max} are shown in figure 6. These dependencies are described by the formula

$$g(T) = \frac{4\,(T - T_o)\,(T_d - T)}{(T_d - T_o)^2} \quad \text{and} \quad f(a) = (a/a_{opt})\,e^{(1-a/a_{opt})}. \tag{23}$$

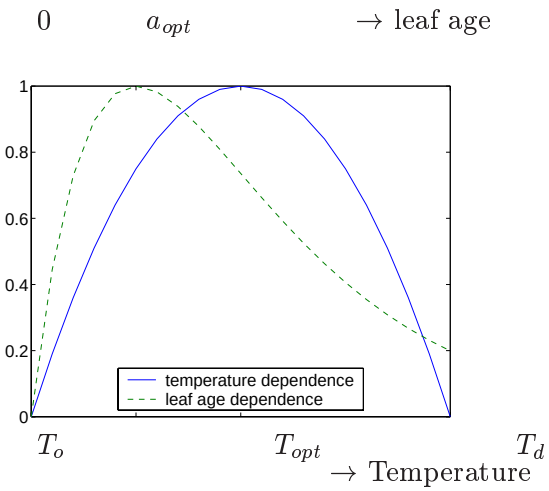


FIGURE 6. The temperature and leaf age dependence of local photosynthetic rates V_{cmax} and P_{max}

This temperature dependence $g(T)$ is chosen, instead of the one in Harley *et al.* (1992) and Kim and Lieth (2001), because it includes a minimum and maximum temperature T_o and T_d, respectively, below and above which the photosynthesis production is zero, respectively. The age dependence $f(a)$ follows from Lieth and Pasian (1990).

The intercellular CO_2-concentration is

$$C_i = C_a - \beta \frac{P(t,a,h)}{g_s}, \qquad g_s = g_0 + g_1 P(t,a,h) R_H / C_a \tag{24}$$

where the ambient CO_2-concentration is given in $\mu molCO_2/(mol\ air)$, R_H is the relative humidity, g_s is the stomatal conductance to H_2O in $mol\ H_2O/(m^2\, s)$, g_0 is the minimal stomatal conductance to H_2O in $mol\ H_2O/(m^2\, s)$, and $\beta = 1.6$. Note that (24) is the quasi steady-state solution of

$$\frac{\partial C_i}{\partial t} = g_s(C_a - C_i) - \beta P(t,a,h), \tag{25}$$

expressing the effects of consumption of CO_2 by photosynthesis and conduction of CO_2 by the leaf stomata. Note also that the intercellular concentration C_i is lower than the ambient one, provided $P(t,a,h) > 0$. For an increasing production $P(t,a,h)$ the concentration C_i is decreasing, while for increasing ambient humidity R_H the concentration C_i is increasing.

Constant	*definition*	*value*
R_d	dark respiration	$0.82\,\mu mol\,CO_2/(m^2\,s)$
V_m	-	$94.3\,\mu mol\,CO_2/(m^2\,s)$
Γ_*	CO_2 compensation point	$44\,\mu mol\,CO_2/mol$
κ	-	$730\,\mu mol\,CO_2/mol$
P_m	-	$56.6\,\mu mol\,CO_2/(m^2\,s)$
α	quantum efficiency	$0.055\,mol\,CO_2/(mol\,photon)$
θ	curvature factor	0.7
T_o	lower temperature bound	$10\,^oC$
T_d	upper temperature bound	$48.6\,^oC$
a_{opt}	optimum age	$28.01\,days$
g_0	minimum stomatal conductance	$0.18\,mol\,H_2O/(m^2\,s)$
g_1	-	6.71
β	conversion factor CO_2/H_2O	1.6

TABLE 3. Definitions and given values of the constants used in the leaf photosynthesis model.

The leaf temperature T and the ambient temperature T_a are related by linearizing the expression on page 232 of Jones (1992)

$$T = T_a + [4 - (2/45)\,T_a]\,I_0/(1380) - [1 + (6/45)\,T_a]\,(1 - R_H)/0.7 \tag{26}$$

with I_0 expressed as $\mu mol\,photons/(m^2 s)$ and R_H taking some value between 0 and 1.

The calculation of the photosynthesis rate $P(t, a, h)$ requires the solution of a quadratic equation, since the intercellular CO_2-concentration C_i and the stomatal conductance g_s both depend on $P(t, a, h)$. However, in various asymptotic limits, the equation for $P(t, a, h)$ can be linearised, simplifying the calculations (see Appendix A.2).

4. Parameter estimation from harvest data

In the model, we have introduced seven parameters: $k_1, \cdots, k_6$ and either k_7 or k_8. These parameters need to be estimated.

The rose grower can directly estimate k_4 by measuring the average leaf area per metre of stem. Similarly, the average growth speed $V = (h_{cut} - h_0)/T_{max}$ can be found from the stem height desired $(h_{cut} - h_0)$ and the average length of the growth cycle of such a rose stem T_{max} for the current season; the reciprocal of $V = 1/k_5$ determines k_5 as required.

We determine k_3 by averaging (3) in time. After simplifying and rearranging, we find

$$k_3 = \frac{\bar{H}}{V\,(h_{cut} - h_0)\,D_{cut}} \tag{27}$$

where $\bar{H}$ is the average harvest of stems per square metre of greenhouse, V is the average of $v(t; d)$ and D_{cut} is the average of $d(h_{cut}, t)$. A relation

between k_1 and k_2 is obtained by averaging $d(h_0, t) = k_2\, P_{net}$ and $k_1\, P_{net} = k_3\, v(t; d)\, N(t)$. Using (27), our simplification gives

$$k_2 = k_1\, \frac{D_0\, D_{cut}\, (h_{cut} - h_0)}{\bar{H}\, \bar{N}} \tag{28}$$

with D_0 the average of $d(h_0, t)$ and $\bar{N}$ the average of $N(t)$.

In the first approach to model the bush, a rough estimate of k_7, the ratio of the leaf area in the bush to that of the crop, needs to be provided by the rose grower.

In the second approach, we take $k_8 = \kappa\, k_4/k_5$. It now turns out that it is only necessary to obtain the combinations $\kappa_1 = k_1\, k_4, \kappa_2 = k_2\, k_4$ and $\kappa_6 = k_6\, k_4$.

Finally, the parameters k_1 and k_6 (or κ_1 and κ_6) are obtained by fitting them to the weekly harvest data, given the measured time series for the ambient climate.

The roses in the greenhouse considered are of the variety "Red Berlin", planted in May 1999 on a total surface of $8480\, m^2$. The data consists of the number of harvested stems per square metre of greenhouse and the harvested grams per stem over 56 weeks, from week one 2001 until week four in 2002. In addition, time series data of the ambient conditions are provided over the same period. These quantities are all measured and recorded at irregular times, ranging from a few minutes to one hour or more. The resulting ambient light intensity can also be determined from the data using the information on the incoming sunlight, the intensity of any additional artificial light sources, and the screen settings (screens are used to shield roses from too much sunlight). By averaging we can subsequently find $\bar{H}$ and $\bar{N}$.

5. Conclusion and discussion

We have considered the question of optimising rose production in a greenhouse. A rose production model has been constructed that consists of a local and a global model coupled together. In this model, rose growth depends naturally on the time-dependent ambient conditions given by the temperature, the relative humidity, the CO_2-concentration, and the light intensity.

The global model is governed by an advection equation for the stem density function $d(h, t)$. The key assumption used is the unselfishness principle, which implies that the photosynthetic energy produced in the leaves is distributed evenly among the stems, and hence that the advection speed $v(t; d)$ is an explicit function of time only. Other simplifying assumptions used are that any new leaves appear at the top of the stem and that the mass and leaf area are uniformly distributed along the stem. Consequently, the leaf area distribution is directly proportional to the stem density function. It is shown that v can be determined from the net photosynthetic rate, which is

the sum of the photosynthetic rate in the bush below height $h < h_0$ and the photosynthetic rate in the rose crop between heights h_0 and h_{cut}.

The net photosynthetic rate depends on the local photosynthesis in a leaf as well as the ambient climate. A local model adapted from the biological literature (Harley *et al.*, 1992, Kim and Lieth, 2001) is then used to determine this local photosynthetic rate which is a function of leaf age and height.

The total model contains seven unknown parameters that can be estimated from direct measurements on the rose plants, the average harvest and weekly harvest data, as well as from the time series data of the ambient climate in the greenhouse.

This article describes the theory behind our model. Future research is required to test and validate the model. The first necessary step is to estimate the model's parameters by using the greenhouse data provided for the rose variety "Red Berlin" over 56 weeks in the years 2001 and 2002. Subsequently, we can attempt to optimise the rose production by running the model in forecasting mode. During this comparison between model and data, we anticipate further model improvements may be required, such as solving the nonlinear integral equation for the advection speed and the inclusion of a storage mechanism for photosynthetic energy.

It is quite clear that some of our modelling assumptions are an oversimplification of the real situation. One major drawback in our approach seems to be the fact that the only measure for the development of a rose stem is given by its accumulation of biomass due to photosynthesis. This is not very realistic, as can be seen for example from the seasonal differences in the thickness of the harvested rose stems. In particular, the process of blossoming, which is a crucial guide to when the rose stem should be cut, is not modelled at all. Moreover, various bush models are possible, corresponding to the different ways the rose bush supporting the stems can be allowed to grow (or not) in the greenhouse. However, we hope that our approach via stem, leaf, and age density functions will prove flexible enough to be used as a basis for more complex and precise models. Such improvements, together with the continuation of the work on parameter estimation, form an intriguing challenge for further research in optimising rose production.

Acknowledgements

We thoroughly enjoyed this challenging 42nd European Study Group with Industry and wish to thank the organizers for all their efforts. Djoko Wirosoetisno was an invaluable help in converting the supplied Windows data to a readable format. We also thank Odo Diekmann for his useful comments.

A. Appendix

A.1. Numerical methods. In this appendix, we describe how we (numerically) integrate the rose production model forward given the ambient climate. We illustrate the method for a slightly different scenario to the one described in section 2, where all the rose stems with a length greater then $h_{cut} - h_0$ are harvested at certain specific times (discrete) instead of continuously.

A.1.1. Advection equation. In summary, the complete mathematical model is given by

$$\partial_t d + \frac{k_1\, P_{net}(t;d)}{k_3 \int_{h_0}^{h_{max}} d(\zeta,t)\,\mathrm{d}\zeta}\,\partial_h d = 0 \tag{29}$$

with boundary conditions

$$d(h_0,t) = k_2\, P_{net}(t;d) \qquad \text{when} \quad v(t;d) > 0, \quad \text{and} \tag{30}$$

$$h_{max}(t) = max(h_{cut}, \int_{h_{max}(0)}^{t} v(\gamma;d)\,\mathrm{d}\gamma). \tag{31}$$

Irrespective of the up- and downwind cases $v(t;d)$ remains similar, but when $v(t;d) = k_1\, P_{net}/(k_3\, N(t)) < 0$ with $N(t) = \int_{h_0}^{h_{max}} d(h,t)\,\mathrm{d}h$ we have $h_{max} < h_{cut}$ and the crop length decreases. Since (29) is an advection equation, it requires a boundary condition at $h = h_0$ when $v > 0$, and it has a moving boundary condition at $h = h_{max}$ when $v < 0$. We assume that the conditions are such that $h_{max} > h_0$.

Define $x = (h - h_0)/(h_{max} - h_0)$, then (29) becomes

$$\partial_t d + \left(\frac{v - x\,\dot{h}_{max}}{h_{max} - h_0}\right)\partial_x d = 0 \quad \Longrightarrow \quad \partial_t d + u\,\partial_x d = 0 \tag{32}$$

with $x \in [0,1]$ and $u = (v - x\,\dot{h}_{max})/(h_{max} - h_0)$. When $v < 0$, we note that $d(x = 1, t)$ is the last value $d(1, t')$ at time t' when v became zero. When $h_{max} = h_{cut}$ we have $dh_{max}/dt \equiv \dot{h}_{max} = 0$ and (32) is just a rescaled version of (29).

We used a second-order up- or downwind scheme depending on the sign of u to spatially discretise x in the interior and use a first-order up- or downwind scheme at the left-or right boundary, respectively. Using $N + 1$ grid points from $x \in [0,1]$ we arrive at a system of ordinary differential equations. A variable time stepping scheme for ordinary equations from Matlab is used (i.e. `ode15`).

A.2. A linear expression for the local photosynthetic rate. In the limits where $C_i \gg \Gamma_*$ and $C_i - \Gamma_* = \epsilon$ with $\epsilon \ll 1$, the expression for the photosynthetic rate becomes linear. This can be seen by considering the limiting behaviour of A_v in (19) and A_j in (20) and by approximating

$C_i \geq C_{ia} = \max[C_a - \beta\, P(h,a)/g_0, C_a - \beta\, C_a/(g_1\, R_h)]$. Then, we can simplify (18) to give

$$P(t,a,h) = \min\left\{ V_m\, g\, f, V_m, \frac{C_{ia} - \Gamma_*}{\Gamma_* + \kappa}, J/4, J\,\frac{C_{ia} - \Gamma_*}{12\,\Gamma_*} \right\} - R_d, \tag{33}$$

which is linear in $P(t,a,h)$. The net photosynthesis $P_{net}(t;d)$ as a function of time t follows from (17) and (33) by using the climate data. Linear interpolation is used to find the climate values at times t between given data points.

Bibliography

[1] P.C. Harley, R.B. Thomas, J.F. Reynolds, & B.R. Strain, *Plant, Cell, and Environment* **15**, 271–282, 1992.
[2] H.G. Jones, *Plant and Microclimate*, CUP, 232–235, 1992.
[3] S.-H. Kim & J.L. Lieth, *Proc. III IS Rose Research*, 111–119, 2001.
[4] J.L. Lieth & C.C. Pasian, *J. Amer. Soc. Hort. Sci.* **115**(3), 486–491, 1990.

swi 2002

pp. 77 – 90

CHAPTER 5

Magma Design Automation: Component placement on chips; the "holey cheese" problem

Rachel Brouwer, Thijs Brouwer, Cor Hurkens, Martijn van Manen, Carolynne Montijn, Jan Schreuder, JF Williams.

Abstract. The costs of the fabrication of a chip is partly determined by the wire length needed by the transistors to respect the wiring scheme. The transistors have to be placed without overlap into a prescribed configuration of blockades, i.e. parts of the chip that are beforehand excluded from positioning by for example some other functional component, and holes, i.e. the remaining free area on the chip. A method to minimize the wire length when the free area is a simply connected domain has already been implemented by Magma, but the placement problem becomes much more complex when the free area is not a simply connected domain anymore, forming a "holey cheese".
One of the approaches of the problem in this case is to first cluster the transistors into so-called macro's in such a way that closely interconnected transistors stay together, and that the macro's can be fit into the holes.
One way to carry out the clustering is to use a graph clustering algorithm, the so-called Markov Cluster algorithm. Another way is to combine the placement method of Magma on a rectangular area of the same size as the total size of the holes, and a min cut-max flow algorithm to divide that rectangle into more or less rectangular macro's in such a way that as little wires as possible are cut.
It is now possible to formulate the Quadratic Assignment Problem that remains after clustering the original problem to one with 100 up to 1000 macros. There exists a lot of literature on finding the global minimum of the costs, but nowadays computational possibilities are still too restrictive to find an optimal solution within a reasonable amount of time and computational memory. However, we believe it is possible to find a solution that leads to a acceptable local minimum of the costs.

1. Introduction

One of the steps in the design process of chips is the positioning of every single transistor or "cell" on the chip. This means that, given a certain wiring scheme, i.e. the scheme describing the connections between the cells, and taking into account the - relatively few - cells with a prescribed position, the positioning of the various cells has to be determined while under the following conditions. First, the cell must be placed within a certain rectangle, the so-called core area; next, the cells are not allowed to overlap; and finally, the

total wire length must be minimized, as the cost of a chip is proportional to the total wire length. For this problem, many algorithms are known, each one with its specific pros and cons.
The problem becomes more difficult when large parts of the core area are excluded from positioning, often due to large, functional components that were placed beforehand (e.g. memory, or components designed by other companies), creating so-called blockades. The remaining "free area" within the core area is usually comparable to a "holey cheese". Obviously, the cells cannot be placed on the blockades, and this additional requirement makes the positioning problem significantly harder. Figure 1 shows an example of a typical "holey cheese". The filled areas are allowed and form the so-called "holes", the white ones are blockades. Notice that the free area is strongly disconnected.

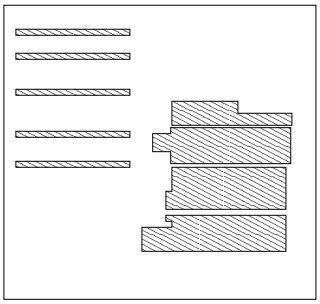


FIGURE 1. A typical example of a "holey cheese" configuration: the transistors may be placed in the filled zones, the white remaining areas are blockades.

The current algorithms of Magma suffice for simply connected domains. However, in "holey cheese" cases they often end up in local minima for the costs that are far from optimal. The purpose of our group during the study group week was to find an algorithm enabling Magma to find more optimal placements of the cells in the case of a "holey cheese".

To this end we decided to make the following approach of the problem: first regroup the strongly connected cells in more or less equally sized "macros", then place these macros in the holes in such a way that the wire length is minimized. We present two possible approaches for the regrouping of the cells. The first one, dealt with in section 2.1, departs from the wiring scheme and uses a clustering algorithm. The second one, subject of section 2.2, is a combination of a preprocessing step using the original Magma software, followed by a repeated application of a min cut-max flow algorithm. Finally,

in section 3, we discuss the method to minimize the wire length, formulating a Quadratic Assignment Problem (QAP).

2. Clustering the cells

2.1. The Markov Cluster Algorithm. We believe that the Markov Cluster Algorithm (MCL) provides a good method to group the cells in macros. This algorithm uses the notion of random walk for the retrieval of cluster structure in a graph. In a random walk at each cell the direction to be followed is given by chance. Imagine a vast collection of random walks, all starting from the same cell. Walkers will in general follow different paths. An observer floating high above them will see a flow: the crowd slowly swirles and disperses, much as if a drop of ink is spilled into a water-filled tray.
The aim of a cluster method is to dissect a graph into regions with many interconnections inside, and with only a few interconnections between regions. Once inside such a region, a random walker has little chance to get out. The idea behind MCL is very simple. Simulate many random walks (or flow) within the whole graph, and strengthen flow where it is already strong, and weaken it where it is weak. By repeating the process an underlying cluster structure will gradually become visible. The process ends up with a number of regions with strong internal flow (clusters), separated by 'dry' boundaries with hardly any flow.
We refer to the PhD-thesis of Stijn Van Dongen[6] for a detailed review of this algorithm.

2.2. Constructing macros with the min cut-max flow algorithm. The method of constructing macros with a min cut-max flow algorithm departs from a square S with surface A_s over which all the cells have been positioned in such a way that they do not overlap and that their connectivity has already been taken into account, meaning that the highly interconnected cells are already put together. This positioning of the cells within a square is the result of a preprocessing method implemented by Magma.
Schematically the procedure is as follows: put a grid over square S of which the grid lines may slightly be deformed. Then use a min cut-max flow algorithm to deform the grid lines in such a way that they cut as little connections as possible. The cells contained in the resulting grid cells then form the macros.

2.2.1. *Restrictions on the macros.* Assume the number of macros we want to make is n, and let A_h and A_m be the total surface of the holes and the average size of the macros, respectively. There are some restrictions on the number and size of the macros. First, as will be shown in section 3, the computationnal hardness of the QAP imposes the number of macros to be no larger then 1000. Also creating a large number of macros might result in breaking up highly connected parts, which looks inefficient since

it is probable that the QAP routine will put them together again. But the number of macros shouldn't be too small either since then the QAP might have no effect. Second, we want the average macro to fit at least twice in the smallest hole; this constraint is not very strict when $A_s \ll A_h$ since we then may choose to ignore very small holes, however, if $A_s \approx A_h$, it is necessary to use all possible space.

2.2.2. *Defining the adjustable grid.* Figure 2 shows how the adjustable grid can be defined: each grid cell should contain exactly one striped rectangle and the grid lines (dash-dotted lines) can be moved freely within the white areas.

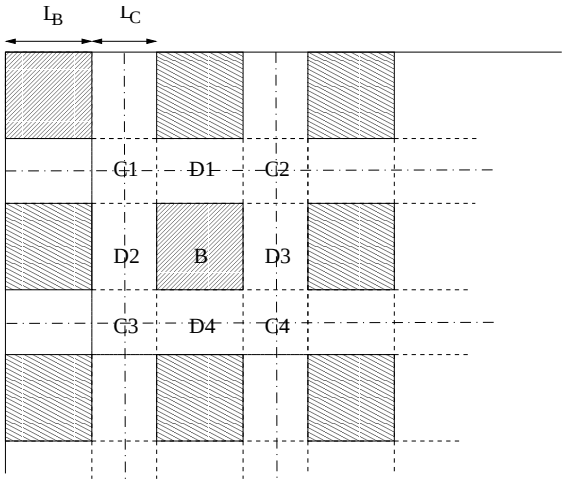


FIGURE 2. The definition of the adjustable grid: each grid cell should contain exactly one striped rectangle and the grid lines (dash-dotted lines) can be moved freely within the white areas.

Define squares B_{ij} (the striped ones in figure 2) with edges of length L_B, the upper left and right corner points given by the points $\{(i-1)(L_B+L_C), (j-1)(L_B+L_C)\}$ and $\{(i-1)(L_B+L_C), (j-1)(L_B+L_C)+L_B\}$, respectively, and the lower left and right corner points being $\{(i-1)(L_B+L_C)+L_B, (j-1)(L_B+L_C)\}$ and $\{(i-1)(L_B+L_C)+L_B, (j-1)(L_B+L_C)+L_B\}$, respectively, for $i = 1...\lceil\sqrt{n}\rceil$, for $j = 1...\lceil\sqrt{n}\rceil$. Each square B_{ij} is now separated of a neighboring square by a distance L_C. We assign all the cells contained in B_{ij} to macro A_{ij}, and we still have to assign the cells contained in the surrounding areas to one of the neighboring macros. This we will be done with help of a min cut-max flow algorithm.

For the lengths L_B and L_C we suggest:

$$L_B = (A_s/2n)^{1/2},$$

$$L_C = (\sqrt{2} - 1)L_B.$$

That is, exactly half of the surface is assigned to the macros beforehand. With $L_C = (\sqrt{2}-1)L_B$ we obtain $n \cdot (L_B + L_C)^2 = A_s$. It is of course possible to adjust these numbers. If we take L_C bigger and L_B smaller, we have less surface assigned beforehand, hence it seems reasonable to assume that we will obtain a better assignment. On the other hand, this has also disadvantages : the size of a macro will vary more, and hence we may encounter problems with the placing of these macros in the holes if A_s is close to A_h. Moreover, the speed of the min cut-max flow algorithm depends on the size of the graph. However, the algorithm can be executed in polynomial time, so this does not have to lead to considerable delay.

2.2.3. *The assigning procedure with a min cut-max flow method.* For all cells c, we make a list of possible assignments to the macros. Notice, from figure 2 that the cells in the areas $C_1, ...C_4$ belong to the surrounding areas of four squares, whereas the cells in the areas $D_1...D_4$ belong to the surrounding area of only two squares. Notice also that connections that reach over the borders of a square B_{ij} and its surrounding area will automatically be cut. We then assign all cells c with their midpoints inside B_{ij} to macro M_{ij}. For each cells c in the surrounding area of B_{ij} we make a list with all the possible macros it can be assigned to.
We now construct a graph consisting of all cells in B_{ij} itself and in its surrounding area. We contract all cells inside the square B_{ij} to one point S, and we contract all cells outside the surrounding area to one point T, while keeping their connecting wires as they are (See Figure 3.) We now apply a min cut-max flow algorithm to determine a minimal cut set and assign cells to M_{ij} according to this cut set. The min cut-max flow theorem and algorithm have first been investigated in 1956 by Ford and Fulkerson. All textbooks on graphs and flows contain the theorem and often also an algorithm. For some recent versions of the theorem and the algorithm we refer to Diestel[5], Gross[6] and Jungnickel[7]. For a cell that is not assigned to M_{ij}, we remove this macro from its list as a possible assignment. Notice that it is also possible that we have cells in the surrounding area that are not connected to either S or T. For these (small clusters of) cells we can choose an arbitrary macro. In our opinion, it is advisable to assign these cells in such way that it will lead to macros that do not vary much in size.

The min cut-max flow algorithm assures that for the graphs as we have constructed above, a minimum of connections will be cut. This does not take into account the length of these connections. It does keep clusters of highly connected cells in one macro. Note that "long" connections, that reach over two or more squares B_{ij} will be cut automatically. These connections will probably belong to nets that reach over big distances and that therefore

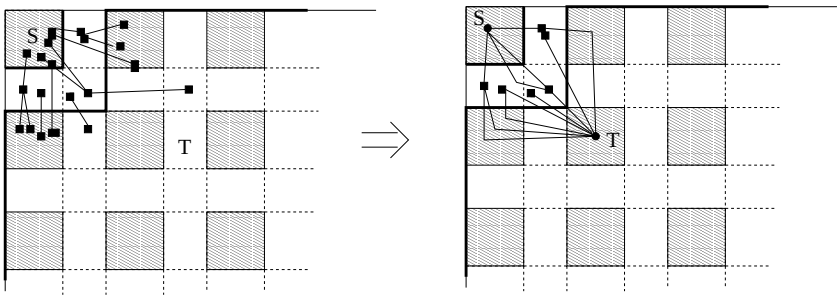


FIGURE 3. Sketch of the graph used in the min cut-max flow algorithm. The last figure shows the cluster (filled) obtained with the algorithm

must always be cut, no matter what way we choose to construct the macros. Note that the procedure above depends heavily on the ability of the Magma-procedure to keeping highly connected clusters together.

3. Formulating a new macro placement problem

After the preprocessing stage, the set of all movable cells c has been partitioned into n macros. For each movable cell c, let $M(c)$ denote the macro to which it belongs. Each macro μ has an area requirement $A(\mu)$, which is equal to the total area of the transistors in the macro: $A(\mu) = \sum_{c \in \mu} A(c)$. Whenever two macros contain transistors that appear in the same net, these macros are connected. Similarly, when a macro contains a transistor which is connected by a net with a fixed pin f, the macro is also considered to be connected to pin f. We elaborate on the connectivity structure between macros and fixed pins in the following section.

Besides the macros we are given m holes in the placement area, with total area not less than the total required transistor area. We consider the problem of assigning the macros to the holes in such a way that the expected resulting wire length — after refined placement of the transistors within each hole — will be as low as possible.

In order to properly define the problem, we first have to define the exact connectivity requirements, and make up our mind how to assign macros to holes in such a way that the resulting wire length between the macros is accounted for as precisely as possible.

3.1. Connectivity between macros. The connectivity between transistors and fixed pins has originally been defined in terms of *nets*, where a net is simply a subset of the collection of placeable and fixed pins. It is evident that after placeable transistors have been clustered into macros these connectivity requirements carry over.

Let N be an original net, that is, $N = \{c_1, c_2, \ldots, c_k\} \cup \{f_1, \ldots, f_l\}$ containing k transistors and l fixed pins, $k \geq 0, l \geq 0$. These pins have to be connected by some wiring network. This implies that this wiring network covers macros $M(c_1), \ldots, M(c_k)$ and fixed pins $f_1, \ldots, f_l$. See Figure 4 for an example.

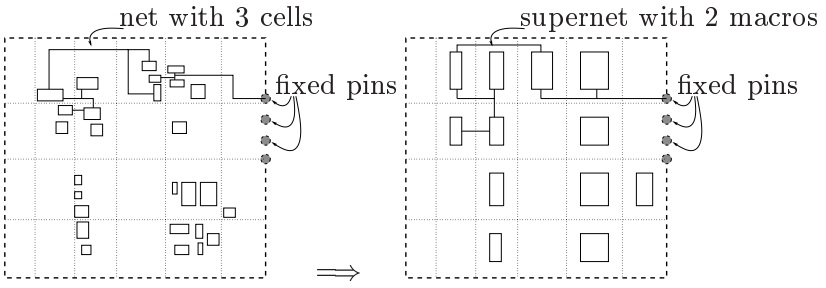


FIGURE 4. Partitioning into macros, connected by supernets

Note that in the macro formulation some of the requirements will be lost, in particular when $l = 0$. It may be the case that $M(c_1) = \ldots = M(c_k)$. In this case we have no information about the resulting wiring length for net N, other than that it will not be big, since it will not be stretching over two or more holes.

What matters is, how many **distinct** macros are covered by each supernet. Removing duplicates from $M(c_1), \ldots, M(c_k)$ we say that the *net* N induces a *supernet* $N' = \{M(c_1), M(c_2), \ldots, M(c_k)\} \cup \{f_1, \ldots, f_l\}$.

It is of interest to consider the numbers of nets and the connectivity of transistors, and to see how this carries over to macros. To play around with the problem we were given several instances of the Magma-problem, and for one of these we were actually given a layout of the transistors with a relative low wiring length, not taking into account that transistors can be placed only inside the prescribed holes.

The toy problem contains 310 fixed pins, and 2099 movable cells, and the wiring structure consists of 2234 nets. The total number of pin-net combinations is 8614. The net size varies from 2 to 288 pins, with an average of 3.58 pins per net. There are 1359 nets of size two, and 318 nets of size three. Each fixed pin was contained in a single net, and each movable cell was contained in between 2 and 7 nets, with an average of 3.96 nets per movable cell.

From the given layout in which the transistors were laid out more or less uniformly over a square we constructed a partition into 100 macros by subdividing this area in a 10 by 10 grid. Now the number of macro-supernet combinations was 4505, where one should note that as many as 1173 supernets cover only one macro, so there are 1061 supernets that cover

two or more macros, with an average of 3.15 macros per supernet. Among these 1061 supernets, 772 cover only two macros. The maximum number of macros covered by a supernet is 78. So, on average, each macro is contained in 45 supernets, of which 11.7 are singleton supernets. The number of macro pairs that have at least one supernet of size less than 10 in common, is 651. Hence, on average, each macro is connected to 13 other macros.

3.2. Subdividing holes into smaller areas. In order to get a better estimate of the ultimate wire length it seems appropriate to subdivide the holes into areas that are more or less equally sized, and such that the wire length within a hole can be neglected without introducing a too large error. The idea is to subdivide the target holes into such smaller areas, taking the midpoint of each area as the virtual placement position. The choice for M, the number of sub-holes, should depend on m, the number of original holes, as well as on the sizes of the holes, and on the number of macros n. M should preferably divide n, and the area of each sub-hole should be an integral multiple of the average macro area. Since the wire length estimate is more precise when more target holes are taken, we have chosen to take $M = n$, for the toy problem. See Figure 5 for an example.

A problem that arises is to split the target holes into the required number of approximately equally sized sub-holes. First one determines the average sub-hole area by $\hat{A} = \frac{1}{M}\sum_{h=1}^{m} \text{area}(H_h)$. Then hole H_h initially gets assigned $\lfloor \text{area}(H_h)/\hat{A} \rfloor$ sub-holes. Next the remaining $M - \sum_h \lfloor \text{area}(H_h)/\hat{A} \rfloor$ sub-holes are "evenly" distributed over the holes, in the same way as rest votes after an election have to be distributed. Once it is decided that hole H_h is subdivided into M_h sub-holes, the question is where to put these sub-holes so as get an even distribution. This depends on the aspect ratio (ratio of longer side over shorter side) as well as the number. For instance, it is not so obvious to subdivide an area of 20 by 40 into 5 more or less equal areas. The average area within the hole is $800/5 = 160$. One can take five rectangles of 8 by 20, or one of 8 by 20, and four of 16 by 10, or two of 8 by 20 and three of 12 by $\frac{40}{3}$.

Once it is decided how the sub-holes are defined we take the midpoints of these sub-holes as our target positions. Let (X_p, Y_p), denote the midpoint of sub-hole p, for $p = 1, \ldots, M$.

3.3. Assignment cost. Next we will try to assign macros to positions in such a way that each macro is assigned one position, and each position is assigned only one macro. When the clusters are more or less equally sized, and when there is enough slack area in the system such an assignment yields a more or less feasible layout. The cost of such an assignment should reflect the final wire length incurred. Now the actual wiring is done later so we can only estimate it. MAGMA is faced with the same problem and has chosen to

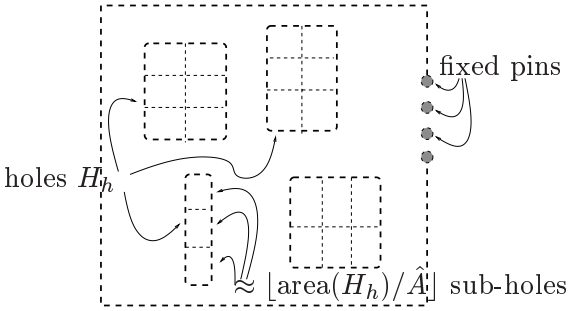


FIGURE 5. Subdivision of holes into subholes

estimate the wire length for a net as (half of) the perimeter of the bounding box of the pins inside a net. Adopting the same approach we could take as the wire length for each *supernet* half the perimeter of the bounding box of the fixed pins of the supernet and the midpoints of the sub-holes covered by the supernet.

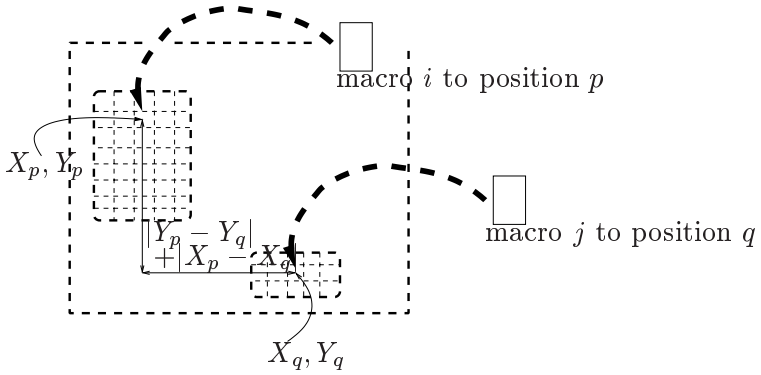


FIGURE 6. Assignment of macros to subholes

Again, this is cumbersome in the sense that such a cost function can only be evaluated once the total assignment has been made. We prefer to formulate a cost function that is of a more local nature. Now recall that in the toy problem many of the supernets only cover two or three macros. For a net of two pins at positions P_1 and P_2, the Manhattan distance $d(P_1, P_2)$ equals the (shortest possible) wire length and equals half the perimeter of the bounding box. For a net of three pins placed at positions P_1, P_2, and P_3, the perimeter of the bounding box is exactly equal to $d(P_1, P_2) + d(P_2, P_3) + d(P_3, P_1)$. Hence, in the estimate of the wire length each distance $d(P_i, P_j)$ contributes with a factor 0.5. For (super)nets of more than three pins it is impossible to tell a priori how much the distance $d(P_i, P_j)$ will contribute

to the perimeter of the bounding box. One may estimate the contribution of the distance between positions P_i and P_j when covered by a net of size $S > 3$ by something like $\frac{2}{S}d(P_i, P_j)$ or by $\frac{3}{S(S-1)}d(P_i, P_j)$. The latter one yields a true lower bound. Another choice could be to neglect wire length with regard to large supernets.

3.4. Formulation. Let $D_{pq} := |X_p - X_q| + |Y_p - Y_q|$ denote the Manhattan distance between positions p and q. Let macros i and j have connectivity C_{ij}, defined by $C_{ij} := \sum_\nu \gamma_\nu$, where the sum is taken over all supernets ν that cover both macro i and j, and where $\gamma_\nu = 1, 0.5$ or $\frac{3}{S(S-1)}$, if supernet ν has size $2, 3$ or $S > 3$, respectively. Note that the size of a supernet is the number of distinct macros and fixed pins that it covers. Then the contribution of the connectivity between macros i and j, when placed at positions p and q respectively, to the estimated wire length will be $C_{ij}D_{pq}$. Let $E_{ip} := \sum_f \gamma_{\nu(f)} d(P_p, f)$ denote the fixed cost associated with assigning macro i to position p. Here the sum is taken over all fixed pins f connected to macro i by a net $\nu(f)$.

We now set the problem in variables x_{ip} with x_{ip} equal to 1, if macro i is placed at position p, and 0, otherwise. The final mathematical problem to 'solve' is the following quadratic assignment problem

$$
\begin{array}{lll}
 & \text{Minimize} & \sum_{ip} \sum_{jq, j>i} C_{ij} D_{pq} x_{ip} x_{jq} + \sum_{ip} E_{ip} x_{ip} \\
\text{(QAP)} & \text{subject to} & \\
(1) & \sum_p x_{ip} = 1 & \forall i \\
 & \sum_i x_{ip} = 1 & \forall p \\
 & x_{ip} \in \{0,1\} & \forall i, p
\end{array}
$$

4. Attempting to solve the QAP

As the title of this section suggests, it is possible to formulate the global placement problem in an appropriate model, but it is not that easy to actually solve this problem to optimality. From literature [1], [2], [3], [4] we have found that even problems of moderate size (with $n = M = 30$) are very notorious. Quite recently a paper [2] has been published announcing the optimal solution of `nugent30`, by years of CPU, using a distributed computing network. This is a quadratic assignment problem with a background similar to that of the MAGMA problem.

One should keep in mind that it is not necessary to solve our problem to optimality, since it is already an approximation of an approximation. We have investigated methods to find a proper lower bound on the QAP-value. The problem is that such a lower bound is found by relaxing the original problem to something that is solvable. The actual solution of the relaxed problem may be far away from a decent solution of the original problem, as

we will see. However, even the *value* of the relaxed problem can be of use, since, if it is a proper lower bound, it may indicate the quality of any solution obtained by whatever way. For instance, it is possible to find assignments of macros to positions, by means of local search: simply start with any solution, apply small changes by exchanging the positions of two or more macros, and proceed until no such change leads to improvement. We have not implemented this approach but find it a true possibility.

Next we will describe some of the insights we found in the literature, and show how some of the easiest lower bounds can be effectively computed. We will give the reference and provide MAGMA with copies thereof.

4.1. Lower bounds. Each method to solve the QAP to optimality needs, at some point in time, a way of proving that the achieved result is best possible. In order to compute a lower bound for the general quadratic assignment problem of the form

$$
\begin{array}{rll}
 & \text{Minimize} & \sum_{ip}\sum_{jq,j\neq i} Q_{ijpq}x_{ip}x_{jq} \\
(\text{GenQAP}) & \text{subject to} & \\
 & \sum_p x_{ip} = 1 & \forall i \\
 & \sum_i x_{ip} = 1 & \forall p \\
 & x_{ip} \in \{0,1\} & \forall i,p
\end{array} \tag{2}
$$

one can rewrite the objective to $\sum_{ip} x_{ip}\sum_{jq,j\neq i} Q_{ijpq}x_{jq}$ and replace the last part by the solution of

$$
\begin{array}{rll}
 & \text{Minimize} & \sum_{jq,j\neq i} Q_{ijpq}x_{jq} \\
(\text{LAP})_{ip} & \text{subject to} & \\
 & \sum_q x_{jq} = 1 & \forall j \\
 & \sum_j x_{jq} = 1 & \forall q \\
 & x_{ip} = 1 & \\
 & x_{jq} \in \{0,1\} & \forall j,q
\end{array} \tag{3}
$$

To finally compute the lower bound, in literature known as the Gilmore-Lawler Bound (GLB), one has to solve one additional Linear Assignment Problem:

$$
\begin{array}{rll}
 & \text{Minimize} & \sum_{ip}(\text{LAP})_{ip}x_{ip} \\
(\text{GLB}) & \text{subject to} & \\
 & \sum_p x_{ip} = 1 & \forall i \\
 & \sum_i x_{ip} = 1 & \forall p \\
 & x_{ip} \in \{0,1\} & \forall i,p
\end{array} \tag{4}
$$

All these linear assignment problems can be solved efficiently using standard network algorithms, even for $n = M = 100$, which yields a 10,000 by 10,000 assignment problem.

Other lower bounds can be computed based on eigenvalue decomposition and projection methods. The main problem with these methods is that the matrix Q is not positive semi-definite. This means that the values one gets by relaxing the constraints $x_{ip} \in \{0, 1\}$ are very low, even negative.

4.2. The QAP in Koopmans-Beckmann form. Note that the cost coefficients have a special form, as the main part of the cost is the product of connectivity and distance. This special form allows for a fast computation of GLB. That is, in order to solve GLB one first has to compute the value $(\mathrm{LAP})_{ip}$, for each ip. When the Koopmans-Beckmann form applies, solution of $(\mathrm{LAP})_{ip}$ amounts to sorting the values C_{ij} (for fixed i, and for $j \neq i$) in non-increasing order, sorting values D_{pq} (for fixed p, for $q \neq p$) in non-decreasing order and computing the inner product of the two arrays. Let $< C_{i*}, D_{p*} >$ denote the value of this inner product. Then the Gilmore-Lawler lower bound for the MAGMA problem is given by

$$
\begin{array}{lll}
 & \text{Minimize} & \sum_{ip}(E_{ip} + \frac{1}{2} < C_{i*}, D_{p*} >)x_{ip} \\
(\mathrm{GLB-MAGMA}) & \text{subject to} & \\
 & \sum_p x_{ip} = 1 & \forall i \\
 & \sum_i x_{ip} = 1 & \forall p \\
 & x_{ip} \in \{0,1\} & \forall i, p
\end{array} \tag{5}
$$

4.3. Constructing a true solution. The GLB value is a true lower bound on the global location problem. The solution of GLB-MAGMA will actually give an assignment. However, this will probably only be good in the sense that the allocation of macros to 'bad positions' does select those macros that have a limited connectivity. But the assignment does not discriminate too much between macros with limited connectivity. So the main contribution of the GLB-solution is its value. Furthermore the GLB-solution could be used as the starting point for an exchange algorithm that can be set up in a local search frame work. This local search approach should be using the true quadratic cost function. By exchanging the assignment of a limited number of macros at a time, say two, three or four, one can effectively compute the change in the objective of a tentative exchange, and perform such an exchange as long as an improvement is made. It is mentioned in literature, that GLB gives a poor bound when the number n is high. In view of this observation, it may be wise to experiment with the number $n = M$, with values in the range from 4 up to 100.

5. Conclusions and recommendations

The problem of positioning transistors in a "holey cheese" configuration in such a way that the costs are minimized can be approached by first clustering the cells with respect to their interconnectivity, and then positioning the resulting macros into the holes so that the resulting wire length is minimized.

One method to perform the clustering is the so-called Markov Clustering Algorithm. Another method is to use the algorithm developed by Magma to position the cells on a rectangular plane in such a way that the wire length is minimized can be used to carry out a further clustering of the cells. The macros can be obtained by putting a grid with movable grid lines over the rectangle that resulted from the Magma procedure. The exact positioning of the grid lines in such a way that as little wires as possible are cut can be found with the help of a min cut-max flow algorithm, and the resulting grid cells then form the macros.

The positioning of the macros in the holes in such a way that the costs are minimal can now can be translated into a Quadratical Assignment Problem. The problem is then not to find an optimum, i.e. a global minimum of the cost function, but an acceptable local minimum with respect to the computation time. We thought for example of partitioning the macros in the holes randomly, and then make small changes to the partitioning to reduce the wire length until these changes have no more effect. There are numerous approaches to reach the minimal costs with a reasonable computation time. We believe that the method of exchanging successively the positions of two or more macros, starting from a random placement, until no improvement is obtained anymore, is an option worth being looked at. A lot of literature exists both on the hardness of QAPs in general and on the wiring problem as a special case. In particular papers by Anstreicher and Brixius [1],[2],[3] are of interest. They deal with finding true optima, and give references to a host of related material. Most of these are results of the PhD-thesis work of Nathan Brixius [3].

Bibliography

[1] K.M. Anstreicher, N.W., Brixius, (2001), A new bound for the quadratic assignment problem based on convex quadratic programming. *Mathematical Programming* **89**, 341–357.

[2] K.M. Anstreicher, N.W. Brixius, J.-P. Goux, J. Linderoth (2002, published online 2001), Solving large quadratic assignment problems on computational grids *Mathematical programming* **91** 563–588.

[3] N.W. Brixius, K.M. Anstreicher, The Steinberg Wiring Problem, to appear in *The Sharpest Cut*, M. Grötschel Ed., SIAM.

[4] R.E. Burkard, E. Çela, P.M. Pardalos, L.S. Pitsoulis (1998), *The Quadratic Assignment Problem* Technical Report no. SFB-126, Technische Universität Graz, Mathematik-B.
[5] R. Diestek, *Graph Theory*, 2nd edition, Graduate texts in Mathematics 173, Springer-Verlag, New York, 2000.
[6] S. van Dongen, Graph clustering by flow simulation, PhD. thesis, Universiteit Twente, The Netherlands, 2000.
[7] J. Gross, J. Yellen, *Graph theory and its applications*, CRC Press, Boca Raton, 1999.
[8] D. Jungnickel, *Graphs, Networks and Algorithms, Algorithms and Computation in Mathematics*, Volume 5, Springer, Berlin, 1999.

swi 2002

pp. 91 – 120

CHAPTER 6

Reconstruction of sea surface temperatures from the oxygen isotope composition of fossil planktic foraminifera

Jan Bouwe van den Berg, Natalia Davydova, Barbera van de Fliert, Frank Peeters, Bob Planqué, Harmen van der Ploeg, Guido Terra.

Abstract. Knowledge of the historic surface temperature of sea water is of importance for the calibration of climate models. The oxygen isotope composition of the shells of several species of planktic foraminifera can be used as a measure for this sea surface temperature. In this paper we investigate how mathematical models can contribute to the process of extracting information about the temperature at which the foraminifera lived from measurement of the oxygen isotope composition of their shells. A simple model is proposed which captures both the average and the variability of the temperature. Preliminary findings suggest that this model forms a solid basis for future research.

Keywords: Historic sea surface temperature, foraminifera, oxygen isotope composition, mathematical models

1. Introduction

With the ongoing debate on global warming of the last decades, the need for a solid understanding of the variability of the world's climate is a hot issue. To assess such changes researchers often use climate models to predict the climate of the future. Predictions of future climate conditions are based upon knowledge of current and past conditions. The more detailed and accurate this knowledge, the better the prediction. One of the many components of climate models is the surface temperature of sea water. To be able to predict sea surface temperatures for the future, knowledge of sea surface temperatures of the past is thus essential. Here we propose to use the chemical composition of fossil shells of planktic foraminifera to reconstruct past sea surface temperatures and their variability. We investigate several mathematical models that can be used to perform this task.

In sections 1.1–1.5 we discuss various aspects of the background information on planktic foraminifera and their dependence on the sea surface temperature. In section 2 a basic model is presented and the reconstruction via this model of the sea surface temperature from the experimental data is

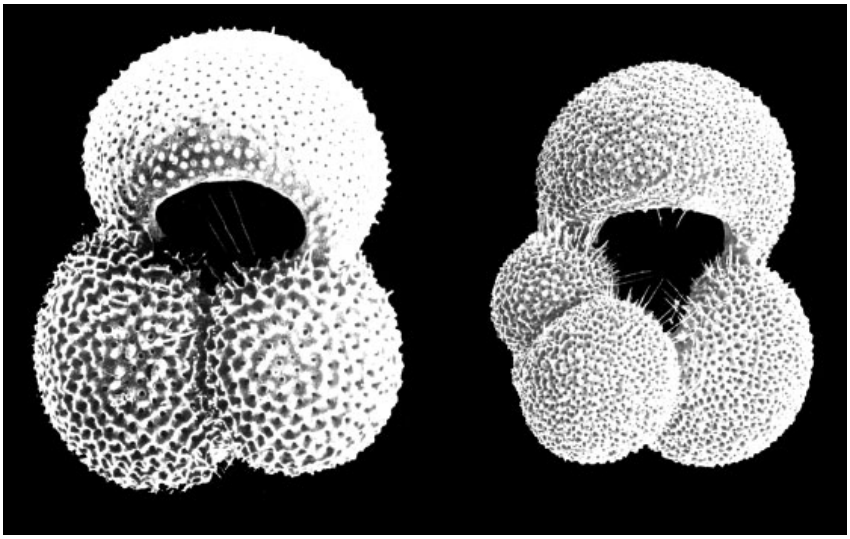


FIGURE 1. Scanning Electron Microscope image of the shells of two species of planktic foraminifera: *Globigerinoides ruber* (left) and *Globigerina bulloides*. The shells were collected in the surface waters of the Arabian Sea. The shells have a diameter of about 300 μm.

carried out in detail. A more sophisticated model describing the population dynamics in terms of a system of ordinary differential equations is discussed in section 3 and it is explained how this in principle can be used to analyse the experimental data. Finally, in section 4 we discuss our findings and give perspectives on possible further research.

1.1. Planktic foraminifera. Planktic foraminifera are small marine unicellular organisms with a shell made of calcite ($CaCO_3$). As opposed to benthic foraminifera which live on the sea floor, planktic foraminifera float in the upper water column. The shells of planktic foraminifera are between $50-500$ μm in diameter and function as their skeleton, see Figure 1. Depending on the taxonomic perspective (it is not always easy to decide whether differences are sufficient to justify a division into separate species), this group of marine microzooplankton comprises about 40 currently living species [**2**]. At the end of their life cycle of about four weeks the organisms reproduce, die, and subsequently sink to the sea floor. Consequently, the sediments found on the ocean floor contain a large number of fossil shells of different species. These sediments serve as a geological archive that can be used to extract environmental information of the ancient upper water column such as, for example, the sea surface temperature. This is possible

since the temperature of the ambient sea water is recorded in the oxygen isotopic composition of the shells during the life cycle of a foraminifer.

A number of methods have been developed to determine the ancient temperatures of the upper ocean. Such models do not use the chemical composition of the shells, but only make use of the relative abundance of different species. These methods assume that there is a relationship between fossil faunas found in the uppermost part of the ocean floor sediments and present-day physical conditions in the ocean (e.g. see [**6**] and references therein). On a wide range of observations multivariate statistical techniques are used to establish a relationship between the relative abundance of different species and the sea surface temperature. This statistical relationship is then used for older sediment layers to reconstruct climatic changes over time. They rely on empirical statistical correlation and do not include ecological information of the organisms considered. We try to improve on these investigations by developing a model that encompasses some basic ecology of the foraminifera and uses the oxygen composition of their shells to reconstruct sea surface temperatures.

1.2. Oxygen isotope composition of $CaCO_3$ and its relationship to temperature. There are several naturally occurring isotopes of oxygen. The main stable isotope is ^{16}O, which has a natural abundance of 99.76%; the next most abundant isotope is ^{18}O. The oxygen isotope composition of a substance is given in conventional delta-notation as ‰ deviation from a given standard, the so-called PDB standard (that is the isotopic composition of CO_2 gas produced with phosphoric acid from a Cretaceous belemnite (*Belemnitella americana*) of the Peedee Formation of South Carolina [**14**]):

$$\delta^{18}\mathrm{O}_{sample} = \frac{(^{18}\mathrm{O}/^{16}\mathrm{O})_{sample} - (^{18}\mathrm{O}/^{16}\mathrm{O})_{standard}}{(^{18}\mathrm{O}/^{16}\mathrm{O})_{standard}} \cdot 1000\ ‰.$$

The isotopic composition of the calcite shells of foraminifera, $\delta^{18}\mathrm{O}_C$, depends on the isotopic composition of the water in which they live, $\delta^{18}\mathrm{O}_W$, the sea surface temperature T, and a so-called vital effect $\delta^{18}\mathrm{O}_{VE}$ (modified after [**4**]):

$$\delta^{18}\mathrm{O}_C = 25.778 - 3.333 \cdot (43.704 + T)^{0.5} + \delta^{18}\mathrm{O}_W + \delta^{18}\mathrm{O}_{VE}. \tag{1}$$

The vital effect is a species-specific correction term. Although the mechanims causing this offset is not well-understood, the vital effect is relatively well-known from field observations [**10**]. In this study we assume that the vital effect for a given species is constant over the sea water temperature range, i.e. $0 - -30^{\circ}C$. The composition of the sea water varies on a much larger time scale than its temperature. Thus, if one performs experiments with currently living organisms, one takes the $\delta^{18}\mathrm{O}_W$ to be a known constant. The change in the oxygen isotopic ratio of the calcite shell is about a 0.2‰

decrease for each degree of temperature increase. This is sufficient for us to be able to estimate the temperature of the water in which the organisms lived. Thus, by measuring the oxygen isotope composition of a foraminifer shell $\delta^{18}O_C$, one can determine the temperature of the water in which the organism built its shell, given the oxygen isotope composition of the sea water $\delta^{18}O_W$.

In this study, we make use of the oxygen isotope composition of three species that live in the uppermost layers of the ocean: *Globigerina bulloides*, *Globigerinoides ruber* and *Neogloboquadrina dutertrei*. Based on field observations the vital effect of these species are: $-0.41‰$ for *G. bulloides*, $-0.45‰$ for *G. ruber* and $-0.01‰$ for *N. dutertrei*. Although these species often occur simultaneously in tropical waters, their ecology differs considerably. For example, *G. bulloides* prefers relatively cool food-rich water, between 3 and 19 °C, such as found at high latitudes (Arctic and Sub-arctic) and in (tropical) upwelling areas. The species *G. ruber*, however, favours relatively food-poor and warmer water between 13 and 32 °C, while the temperature range of *N. dutertrei* is estimated to be between 15 and 25 °C.

1.3. Towards an Isotopic Transfer Function. The main objective of this study is to reconstruct sea surface temperatures of the past using data on the oxygen isotope composition of different species of foraminifer shells. Equation (1) would be a good first candidate, but cannot be applied directly because no accurate information on the oxygen composition of the sea water from past times is available. This makes a direct inference of the temperature from the calcite composition of an individual shell theoretically impossible.

Naively, this problem seems to be overcome when we subtract two such equations by comparing the isotope composition of shells of two species: this would eliminate the $\delta^{18}O_W$ from the equation. But, apart from the vital effect $\delta^{18}O_{VE}$, the relation (1) is the same for each species. Hence one does not expect this difference to contain any information. However, due to differences in the ecological preferences of the different species, the temperature recorded in their shells according to formula (1) is not the same for all species. Since the different species of planktic foraminifera do not prefer the same environmental conditions, the production patterns of different species will vary in the course of the year. For example, it can be expected that "cold loving" species will produce their shells preferably during the cool period of the year, whereas the shell production of "warm loving" species will be biased towards the warm period of the year. Consequently, the isotopic composition of a given species reflects the temperature of the sea water during that part of the year for which environmental conditions were optimal.

A single sample from the cores represents about 100 years, whereas the life span of a foraminifer is a few weeks. For an isotopic analysis of a species about 20 shells are needed. This means that one only obtains (yearly) averaged values. For each species these averages will be biased towards that part of the year during which the environmental conditions were most favourable to them. In an open oceanic setting the $\delta^{18}O_W$-value is relatively constant on seasonal to decadal time scales. We therefore may assume that the different species have experienced the same $\delta^{18}O_W$. Hence by subtracting two such averages the $\delta^{18}O_W$-value drops out. This idea has first been proposed by [**8, 9**]. The differences between the isotopic composition of different species, corrected for their vital effect, must be explained by their difference in calcification temperature. This difference is determined by: 1) the ecology of the species under consideration and 2) the environmental conditions in the upper ocean throughout the year. The relation between the environmental conditions throughout the year and the recorded $\delta^{18}O_C$-values is called the Isotopic Transfer Function. The production of foraminifer shells largely depends on the seasonal temperature distribution (annual mean and variability [**10**]), but other ecological factors, such as food availability, may have to be considered as well. It will be easy to make this model very complicated, considering the number of side effects involved. Hence the ecology incorporated in our models has to be simplified.

1.4. The Arabian Sea. In order to test our models, we make use of data from two sediment cores collected from the Arabian Sea [**3**]. Core 905P is located in the upwelling area off Somalia (to be precise, 10°46"N ; 51°57"E) and core 929P is located north of the island of Socotra (13°42"N ; 53°15"E). We focus on the last 30,000 years, thus covering what is known as the last Glacial-Interglacial Cycle. Three intervals can be recognised. Measuring time in units of a thousand years, a kilo-year (ky), and choosing the origin in the present, these are 1) the time span from 30–18 ky representing the conditions of the Last Glacial, 2) from 18–10 ky representing the transition from Glacial to Interglacial conditions (also known as Termination I) and, 3) the time span from 10–0 ky representing the interglacial conditions of the Holocene. In both cores the $\delta^{18}O_C$ of three species, *G. bulloides*, *G. ruber* and *N. dutertrei*, were measured. In addition, both cores have an independent estimate of past sea surface temperature, derived from the so-called $U_{37}^{k'}$ ratio. It is beyond the scope of this paper to discuss this temperature proxy (an indirect measure for a variable which is not observable directly) in detail, but it is important to remember that it provides a reliable independent sea surface temperature estimate, that is based on microfossils other than planktic foraminifera (in this case coccolithophorids). The $U_{37}^{k'}$ temperature proxy cannot be related to a certain time of the year or season, because it is poorly known when these fossils are produced most. For further information

on this temperature proxy we refer to the work of Ivanova [**3**] and references therein.

The present research is in part aimed at obtaining a method of measuring the surface sea temperature which is independent from the $U_{37}^{k'}$ temperature determination. Another goal is to obtain a measure for the variability of the surface sea temperature during the year.

The hydrography in the western Arabian Sea is controlled by the monsoon system. During summer (June–September) the SW (southwest) monsoon winds blow over the Arabian Sea and cause upwelling in the area off Somalia. This results in lower sea surface temperatures and higher biological productivity since cold and nutrient-rich waters are brought to the sea surface from deeper regions. In winter (December–March), the NE-monsoon winds blow from the continent to the sea and do not cause upwelling, but result in convective mixing of the upper layers of the ocean. This also causes increased biological productivity, but generally does not lower the sea surface temperature as much as during the SW-monsoon period. The inter-monsoon periods are characterised by a relatively high sea surface temperature and low biological productivity. Based on present-day observations [**1**], it is evident that the planktic foraminifer shells are mainly produced during the two monsoon seasons and much less during the inter-monsoon periods. Consequently, it can be expected that the fossil shells found in the sedimentary record were produced during the SW- or NE-monsoon period. It therefore makes sense to reconstruct sea surface temperatures of the two monsoon periods only.

1.5. Secondary calcification. The shells of planktic foraminifera are mainly composed of so-called primary calcite, a type of calcium carbonate that is formed during their life in the surface waters of the oceans. It is known, however, that shells of equal size found in and on the sediments on the sea-floor often have a higher shell mass compared to the shells found in the surface layers of the oceans. The main reason for this increase in shell mass is the formation of an extra calcite layer, also known as secondary calcite, see e.g. [**5**]. This crust is formed at the end of the foraminifer life cycle, at a depth in the ocean where reproduction takes place. For "shallow dwelling" species, such as considered in this study, this depth level is found roughly between 50 and 100 meter. At these depths the water temperature is lower than the sea surface temperature, which thus results in an increase of the $\delta^{18}O_C$ of the shell. Secondary calcite therefore may mask the $\delta^{18}O_C$ of primary calcite. The amount of secondary calcification differs for different species: *N. dutertrei* and *G. ruber*, for example, have more secondary calcite than *G. bulloides*. This process obviously obstructs the straightforward use of the $\delta^{18}O_C$ of fossil shells in the calculation of sea surface temperatures. In this study, we assume that the amount of secondary calcification is a

constant fraction of the total shell weight, and that this fraction is species dependent. We also assume that the temperature at which the secondary calcite is formed has a constant offset from the sea surface temperature. These two assumptions allow us to correct for the effect of secondary calcification on the $\delta^{18}O_C$ of the shells, by subtracting a constant value from each observation (section 2.2).

2. A simple mathematical model for the production of fossil layers

In this section two models describing how sea surface temperatures are reflected in fossil foraminifer skeletons are discussed. The simplest one takes into account the influence of temperatures only, the other one addresses the importance of food availability as well. Rather than discussing each model separately, they will be dealt with simultaneously. Both models consist of the same components, differing only in the way each component is filled in:

1. Modelling the dependence of foraminifera on environmental conditions such as sea surface temperature and food availability.
2. Modelling seasonal variations of the environmental conditions.
3. The relation between temperature and δ^{18}O-values.
4. Prediction of the resulting measurements in fossil cores.
5. Reconstruction of paleo-temperature from the fossil core data.

The following sections mirror these parts of the models. The first component is discussed in section 2.1.1, the second in 2.1.2. Parts 3 and 4 are combined in section 2.1.3 and the reconstruction process for sea surface temperatures is finally described in section 2.1.4. The results of the analysis of the experimental data using this model are presented in section 2.2.

2.1. Description of the model.

2.1.1. *Dependence of foraminifera on environmental conditions.* In population dynamics many different ways exist to model the dependence of the population size of a certain species on environmental conditions. In general these models, of which one possible model is discussed in section 3, may lead to complicated ((quasi-)periodic, chaotic) behaviour. Even if the system simply tends to a steady state (depending on temperature), the population needs time to adjust to changing environmental conditions. In fact, the life cycle of an individual foraminifer lasts about two to four weeks, so the biological response of the whole population on a changing environment takes place on this time scale, or slower. This may well be the reason for the large amount of scatter in the data [**13**] showing the relation between population size and temperature. However, in this simplified model we will assume that the population adapts instantaneously and is always at its equilibrium size corresponding to the conditions present. Namely, we assume there is a direct

Species A	$T_{\min,A}$ $(^\circ C)$	$T_{\max,A}$ $(^\circ C)$	$\bar{T}_A$ $(^\circ C)$	σ_A $(^\circ C)$
G. ruber	13	32	32	14
G. bulloides	3	19	11	4
N. dutertrei	15	25	20	2.5

TABLE 1. Parameters describing the temperature dependence of the three different species. The values are taken from [**13**].

relation between the population size and temperature given by a Gaussian distribution

$$P_A(T) = \alpha_A \, e^{-\frac{(T-\bar{T}_A)^2}{2\sigma_A^2}} \tag{2}$$

for species A, where T is the sea surface temperature. The temperature range $[T_{\min,A}, T_{\max,A}]$, the optimal temperature T_A and standard deviation σ_A for the three species we consider can be found in Table 1. The value of the constant factor α_A depends on the exact definition and units chosen for P_A (population size, density, flux of skeletons, calcite deposited). It does not seem possible to estimate α_A accurately. Fortunately, its value is not important in our analysis since this factor scales out of the calculations.

Note that all of these foraminifera die and settle on the bottom of the ocean, so we can also refer to this $P_A(T)$ as the *production* of fossil shells of species A. The parameters tabulated in Table 1 stem from present-day measurements [**13**]. Although it is hard to quantify this due to the large amount of scatter, it is clear that *G. ruber* favours high temperatures, whereas *G. bulloides* prefers colder waters. The *N. dutertrei* species flourishes under moderate conditions. For each species a temperature range, optimal temperature (for which the population size is maximal) and standard deviation is estimated from the data [**13**]. The maximum and minimum temperatures are set to have a distance of two standard deviations from the mean. Outside the ranges $[T_{\min,A}, T_{\max,A}]$, the respective species are barely present at all. We experimented with truncating the production function (2) to zero outside these temperature ranges, but this did not change the results significantly. The results in this paper are obtained by using the Gaussian formula (2) for all T.

The production function (2) describes the temperature dependence of the different species. Another important factor is the food supply. Two main sources of food can be distinguished on which the foraminifera feed, phytoplankton and zooplankton. The species *G. ruber* feeds mainly on zooplankton, *G. bulloides* feeds mainly on phytoplankton, whereas *N. dutertrei* can feed on both, making it less sensitive to the food supply. Sufficient data quantifying this are not available at the moment. For the moment, the

Species A	μ_A	κ_A	ν_A	λ_A
G. ruber	0.4	0.03	1	0.8
G. bulloides	0.2	0.6	0.004	1.9
N. dutertrei	0.6	1.9	1.6	1.6

TABLE 2. Parameters describing the dependence of the three different species on food availability. The values are estimated from Peeters, unpublished data.

dependence of the populations on the nutrients is modelled simply by multiplying the temperature production function (2) with a food factor, resulting in

$$P_A(M,N,T) = (\mu_A M^{\kappa_A} + \nu_A N^{\lambda_A})\, e^{-\frac{(T-\bar{T}_A)^2}{2\sigma_A^2}} \tag{3}$$

where M and N denote the concentration of phyto- and zooplankton respectively (mg·m^{-3}) and κ_A, λ_A, μ_A and ν_A are coefficients describing the sensitivity of species A to the different food sources. In this study we use the values shown in Table 2 for the parameters κ_A, λ_A, μ_A and ν_A. They are rough estimates from experimental data.

A note of caution with regard to the production function (3): to our knowledge it has not yet been attempted to describe the abundance of foraminifera in terms of temperature and food availability, and it may be possible to improve upon (3) on the basis of further research. For instance, the data show a correlation between the abundance of the foraminifera and the phosphate concentration (which is the main food source for phytoplankton) and total biomass (essentially the sum of phyto- and zooplankton).

We conclude this section about the behaviour of the foraminifera under different circumstances by summarising our main assumptions: the populations are always in steady state and adapt to changes in environmental conditions instantaneously (a so-called "quasi-steady" model); either the population density does not depend on food availability and the production function (2) is used or the population density is influenced by two different food sources and a food factor is included, see equation (3).

2.1.2. *Modelling the environmental conditions.* This section deals with the way we model the seasonal variation of sea surface temperature and food availability, namely we will consider temperature and food availability as a function of time throughout the year.

In the introduction it was described already that the Arabian Sea, for which this model is considered, is strongly influenced by the monsoon system. It thus makes sense to divide the year into two distinct seasons, the SW-monsoon period in the summer and the NE-monsoon period in the winter. A simple temperature model can be constructed by assuming a constant

temperature in each season. We fix the duration of each monsoon season at four months. During the inter-monsoon periods, the abundance of the foraminifera is very low due to the lack of food. Therefore we will neglect the production of fossils during those periods. The SW-monsoon is stronger than the NE-monsoon, hence the upwelling induces lower sea surface temperature in the summer whereas the winter temperature is higher. This is illustrated in Figure 2; notice that T_{NE} and T_{SW} are not chosen a priori but will be deduced from the data. Moreover, we interpret (2) to be valid during the two monsoon periods while the production $P_A(T) = 0$ in between the monsoons due to lack of nutrients (hence the temperature in the inter-monsoon periods is (left) unspecified). Also, only the ratio of the duration of the SW- and NE-monsoon is relevant for the outcome of the model.

In principle, alternative temperature models are also possible, e.g. [**9**]. But we have one restriction: for our procedure to work a temperature curve should be completely described by no more than two parameters. For example, a sinusoidal cycle with a mean temperature and an amplitude could be specified as well. The temperature curve used in section 3 (equation (13)) also depends on two parameters. Here we remark only that this is connected with the fact that we have data on three species. If we consider k species we can allow temperature to depend on $k-1$ parameters. The reason for this is discussed in more detail in section 2.1.4. Although the model can be used with other descriptions of the yearly temperature cycle as well, the calculations are simplified considerably by assuming two seasons of constant temperature. We therefore limit ourselves to this description in the current section:

$$
(4)\qquad \begin{array}{lcll} T & = & T_{SW}, & \text{during the SW-monsoon (summer),} \\ T & = & T_{NE}, & \text{during the NE-monsoon (winter),} \\ T & = & \text{unspecified}, & \text{during the inter-monsoon periods.} \end{array}
$$

Having described the temperature dependence on time, the environmental conditions are sufficiently specified for the simplest model (2) in which temperature only is considered. For model (3) we need to specify time dependence of the food availability as well.

An accurate description of the food supply is much harder because phyto- and zooplankton have complicated dynamics of their own, see [**7**]. In fact, the phytoplankton consumes inorganic nutrients, like phosphate, which are brought to the sea surface during an upwelling phase. Zooplankton feeds, though not exclusively, on phytoplankton. Therefore its bloom occurs only after the phytoplankton population has begun to develop. We will not try to model this here. In order to investigate the influence of the food availability on recorded δ^{18}O-values we simply try to specify the values of M and N through time, like we did for the temperature. Note that the two-dimensional food availability function $(M(t), N(t))$ should depend on at most

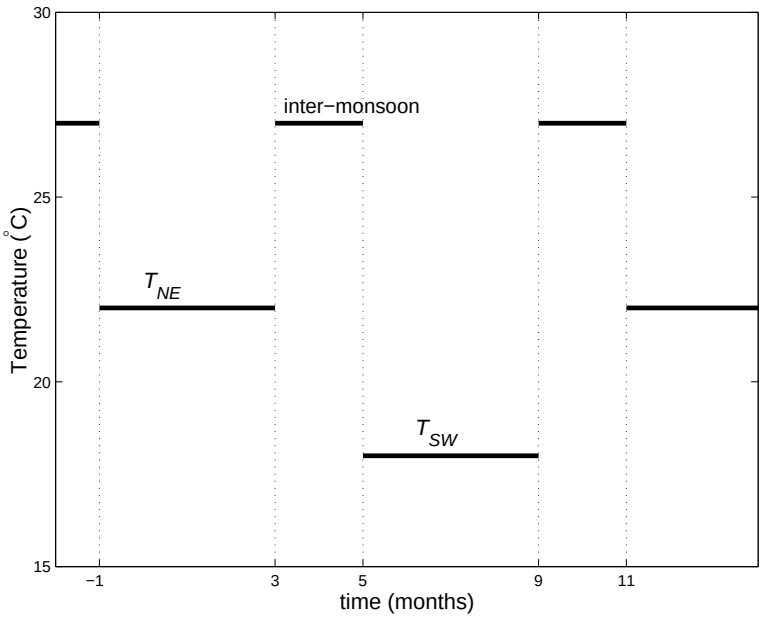


FIGURE 2. The description of the yearly signal of sea surface temperature in the Arabian Sea; two monsoon-seasons at constant temperature, the SW-monsoon during summer and the NE-monsoon during winter.

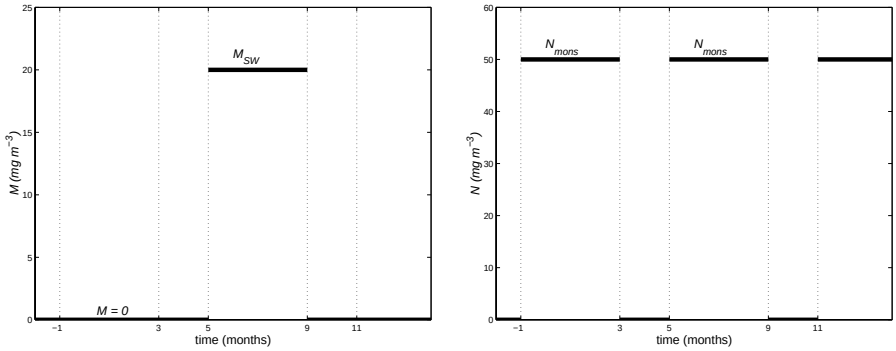
two parameters, so each of the components $M(t)$ and $N(t)$ may depend on one parameter only. To simplify the calculations we use the same two seasons as before, assuming that the phytoplankton is present in the summer season only, whereas the zooplankton is present the whole year round (Fig. 3). This is based on field observations showing that the abundance of phytoplankton is much lower during the winter monsoon than during the summer monsoon period.

This concludes the modelling of the environmental conditions. For the model without nutrients, only the temperature dependence (4) (Figure 2) is used. The other model uses the same temperature model and incorporates the food availability description as in Figure 3:

(5)

$$\begin{array}{rcl@{\quad}rcl@{\quad}l} M & = & M_{SW}, & N & = & N_{mons} & \text{during the SW-monsoon (summer),} \\ M & = & 0, & N & = & N_{mons} & \text{during the NE-monsoon (winter),} \\ M, N & = & 0, & & & & \text{during the inter-monsoon periods.} \end{array}$$

2.1.3. *The resulting core data.* This section deals with the main ingredient of the model: how do the previous two sections relate to the data we measure in the cores? Two aspects of foraminifera skeletons are measured in



(a) Phytoplankton availability (b) Zooplankton availability

FIGURE 3. Description of the yearly cycle of food availability for the foraminifera: (a) The phytoplankton is assumed to be present only during the SW-monsoon (summer). (b) The zooplankton is assumed to be equally available during all seasons.

the cores, the relative abundances of the different species and the mean δ^{18}O-values for each species. In principle it is possible to obtain absolute fluxes for each species instead of relative abundances by measuring the amount of deposited skeletons per period of time. However, these absolute fluxes are less reliable because to determine them one also needs to estimate the age of the core (for example by radioactive carbon measurements). Therefore, relative abundances are more commonly used to express the results of measurements.

In our model the production functions P_A from equations (2) or (3) describe the number of skeletons that settle on the bottom of the ocean. The total production of one species during a year will be the integral of P_A over this period. In the cores, separate years, let alone separate seasons, can not be distinguished. So the measurements in the core show averaged productions over several years. From the model the relative abundances are found by dividing the yearly production of one species by the total production of all species together. This leads to

$$\chi_A = \frac{\int P_A(M(t), N(t), T(t))\, dt}{\int (P_{rub} + P_{bul} + P_{dut})\, dt} \tag{6}$$

for the relative abundance of species A, with the obvious need to suppress $M(t)$ and $N(t)$ from the notation if (2) is used.

Moreover, the $\delta^{18}O_C$-values of the fossil foraminifera can be measured. As was explained in section 1.2, the oxygen isotope composition in their skeleton reflects the temperature at which they lived, according to formula (1). It can be measured in the core as well, for each species separately.

Because it is not possible to measure it for a single foraminifer, only averaged values over several years can be recovered. This average $\delta^{18}O_C$-value is weighed by the number of skeletons produced in different periods (and hence reflects the conditions under which the respective species flourishes). This average for species A reads

$$\overline{\delta^{18}O_A} = \frac{\int \delta^{18}O_A(T(t), \delta^{18}O_W)\ P_A(M(t), N(t), T(t))\, dt}{\int P_A(M(t), N(t), T(t))\, dt} \tag{7}$$

again with the obvious change of notation if (2) is used. Because $\delta^{18}O_W$ appears in equation (1) in a linear way, it can be pulled out of the averaging. Hence

$$\begin{aligned} \overline{\delta^{18}O_A} &= \langle \delta^{18}O_A(T(t), \delta^{18}O_W)\rangle_A \\ &= \langle f(T)\rangle_A + \langle \delta^{18}O_W\rangle_A \\ &= \langle f(T)\rangle_A + \delta^{18}O_W \end{aligned} \tag{8}$$

where we define the (species dependent) averaging $\langle\cdot\rangle_A$ abbreviating the integrals and

$$f(T) = 25.778 - 3.333 \cdot (43.704 + T)^{0.5}. \tag{9}$$

Note that we incorporated the correction for the vital effect $\delta^{18}O_{VE}$ in $\delta^{18}O_A$ without a change in notation. This will be done throughout the remainder of this section, except in Figure 4. The final equality is based on the fact that $\delta^{18}O_W$ does not change much over years so we can assume that it was the same in all seasons.

2.1.4. *Reconstruction of sea surface temperatures.* Equations (6) and (8) predict the outcome of core measurements according to our models. In this section all ingredients of the two models discussed in sections 2.1.1 to 2.1.3 are put together. By working out the equations (6) and (8), we find the Isotopic Transfer Function, i.e. the influence of the environmental conditions on the oxygen isotope composition and the relative abundances.

First consider equation (8). It contains the unknown background value $\delta^{18}O_W$, which is the same for all species. Therefore the $\delta^{18}O_W$-value can be eliminated by subtracting the $\delta^{18}O$-values of two species. For the three species *G. ruber*, *G. bulloides* and *N. dutertrei* this leads to two independent differences

$$\begin{aligned} \overline{\delta^{18}O_{bul}} - \overline{\delta^{18}O_{dut}} &= \langle f(T)\rangle_{bul} - \langle f(T)\rangle_{dut}, \\ \overline{\delta^{18}O_{bul}} - \overline{\delta^{18}O_{rub}} &= \langle f(T)\rangle_{bul} - \langle f(T)\rangle_{rub}. \end{aligned} \tag{10}$$

The right-hand sides of (10) depend on T_{SW}, T_{NE}, M_{SW} and N_{mons}. The left-hand sides of (10) can be found from the core measurements. Hence this provides us with two equations for T_{SW}, T_{NE}, M_{SW} and N_{mons}.

The three relative abundances sum up to one ($\chi_{rub} + \chi_{bul} + \chi_{dut} = 1$), so only two of them are independent. Hence, when we fit the outcome of

the model to the measured relative abundances we have again two independent equations. This implies that if only δ^{18}O-values are used the maximum number of independent parameters that can be recovered is two, while if relative abundances are taken into account as well, four independent parameters can be determined. This is the reason why relative abundances have to be considered as well when using model (3), incorporating the influence of two types of food sources and why no more than a single parameter per food source could be used in section 2.1.2.

Working out the model without food availability leads to

$$\langle f(T)\rangle_A = \frac{\frac{4}{12} f(T_{SW}) e^{-(T_{SW}-\bar{T}_A)^2/2\sigma_A^2} + \frac{4}{12} f(T_{NE}) e^{-(T_{NE}-\bar{T}_A)^2/2\sigma_A^2}}{\frac{4}{12} e^{-(T_{SW}-\bar{T}_A)^2/2\sigma_A^2} + \frac{4}{12} e^{-(T_{NE}-\bar{T}_A)^2/2\sigma_A^2}}$$

for species A. The integrals have been evaluated easily because we assumed two seasons of constant temperature. The parameters $\bar{T}_A$ and σ_A of the Gaussian temperature distribution are taken from Table 1, the factor α_A from equation (2) drops out and $f(T)$ is given by equation (9). The two unknowns are T_{SW} and T_{NE}, which of course are the same for each species. Hence substituting this in equation (10) leads to two equations for two unknowns. They must be solved numerically in order to obtain T_{SW} and T_{NE} from the measured δ^{18}O-differences.

The equations for the model incorporating the effect of food availability are more elaborate. The integrals simplify again due to the assumption of two seasons with constant environmental conditions. There are four unknowns, T_{SW}, T_{NE}, M_{SW} and N_{mons}. Four equations are obtained by considering the relative abundances (6) as well, for two species. It will come as no surprise that these equations (10) and (6) have to be dealt with numerically as well.

We implemented these equations in *Mathematica* and in MATLAB, the latter appearing faster and more robust. The results are discussed in section 2.2. Note that, for reasons of numerical robustness, we did not use standard root finding routines to solve equations (10) and (6). Instead, we used a least squares method: we have taken the differences between left- and right-hand sides of the system of equations (10) and (6) and considered the sum of squares of these differences. The resulting function was minimised with respect to the variables to be solved for: T_{SW}, T_{NE}, M_{SW} and N_{mons}. Of course, if the system has a solution, the minimum is zero.

Finally, note that in principle it is possible to work out the model with other temporal behaviour of the environmental conditions. The resulting integrals will be more difficult to evaluate and have to be calculated numerically as well. However, as long as the number of parameters describing them equals the number of equations, it will generally be possible to recover them in the same manner as described above.

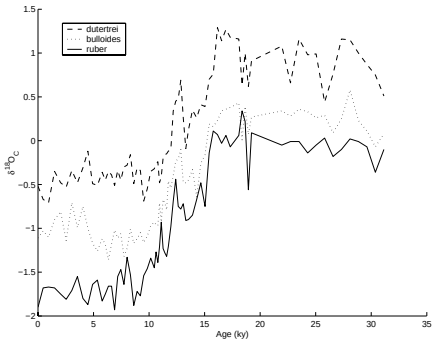
2.2. Results of the simple model. In this section the results of the models are discussed. As input data from two cores are used, core 905P and core 929P, see section 1.4. The $\delta^{18}O_C$-values measured for the three different species *G. bulloides*, *N. dutertrei* and *G. ruber* are shown in Figure 4(a) and 4(b).

A striking aspect in these figures is the clear decrease in $\delta^{18}O_C$-values at the transition between the glacial and interglacial period, around 15 thousand years ago. First this was interpreted as the result of increasing temperature, but it appeared to be strongly influenced by the change in the background value $\delta^{18}O_W$, see equation (1). The melting of polar ice-sheets caused $\delta^{18}O_W$ to decrease significantly.

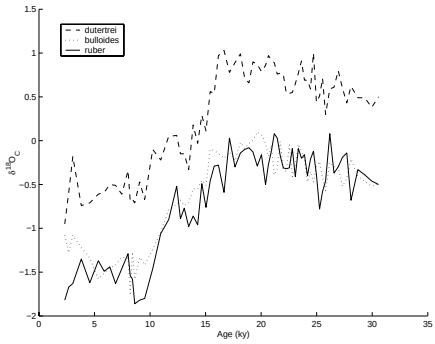
Another aspect is that although the difference between the different species is partly due to the vital effect $\delta^{18}O_{VE}$, differences between the species remain present after correcting for it. We precisely aim to describe these differences with our models.

The first results using the model (2) without food availability, are shown in Figure 4(c) and 4(d). As an independent reference $U_{37}^{k'}$-temperature (see section 1.4) is shown as well. These results should be rated as unreliable. It is highly unlikely that such sudden and vigorous variations in summer and winter temperature have occurred. The changes in summer and winter temperature are expected to be comparable to the changes of the $U_{37}^{k'}$-temperature. Moreover, there seems to be a preference for a temperature of about 20.5 $^\circ C$, with exactly coinciding summer and winter temperatures, hence no temperature variation. If there would be no temperature variation during the year, all species would have recorded the same temperature, hence the same $\delta^{18}O$ (after correction for vital effect). This is not consistent with the data, which do show differences between species. The reason for these inconceivable results is illustrated in Figure 5(a); the measured $\delta^{18}O$-differences lie outside the range that can be explained by our model. The problem is that the data suggest that *G. ruber* and *N. dutertrei* record a lower temperature in their oxygen isotope compositions than *G. bulloides*, whereas in our model the optimal temperature $\bar{T}_{bul}$ is smaller than both $\bar{T}_{rub}$ and $\bar{T}_{dut}$.

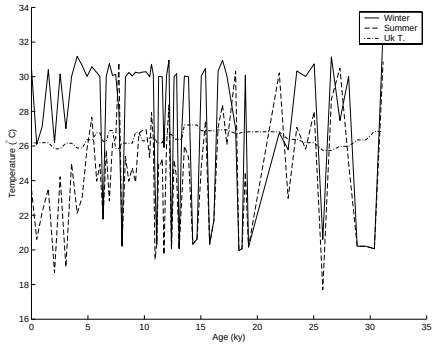
Figure 5(a) shows the space of $\delta^{18}O$-differences. The lines show the model results where the temperature of one of the seasons is fixed. The same model results are shown in Figure 5(b) with lines of constant mean temperature $T_{mean} = \frac{1}{2}(T_{SW} + T_{NE})$ and variation $\Delta T = T_{NE} - T_{SW}$. From these figures it would be possible to read the temperatures responsible for certain core measurements. A measurement on the intersection of two lines would, according to the model, have been produced by the corresponding temperatures. The lines for constant variation ΔT are almost horizontal.



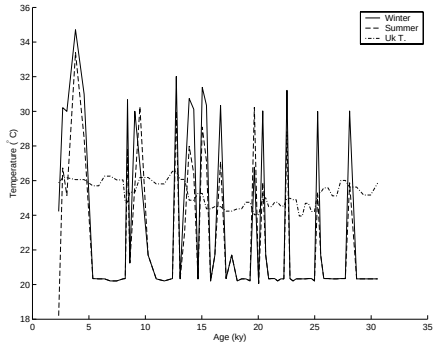
(a) Measured $\delta^{18}O_C$ for core 905P



(b) Measured $\delta^{18}O_C$ for core 929P



(c) First temperature results for core 905P



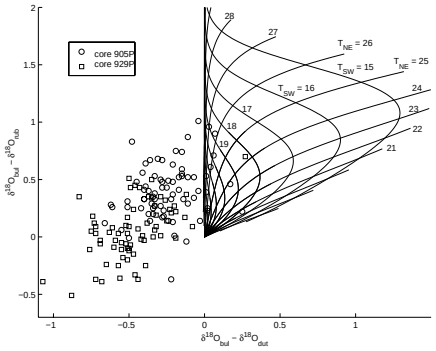
(d) First temperature results for core 929P

FIGURE 4. The raw $\delta^{18}O_C$-data after [**3**] for (a) core 905P and (b) core 929P, and the resulting temperatures for model (2) without food availability for (c) core 905P and (d) core 929P.

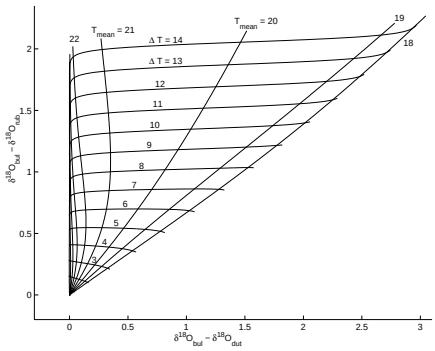
This means that the temperature variation is recorded mainly by the vertical difference $\delta^{18}O_{bul} - \delta^{18}O_{rub}$ with only a moderate correction for mean temperature, confer [**11**]. The δ^{18}O-differences do not depend on mean temperature much, except in the range 19–21 $^\circ C$, which is around the optimal temperature of *N. dutertrei*.

The fact that most measurements are not within the range of the model explains the bad results that were shown in Figure 4(c) and 4(d). In particular the points where the temperatures T_{SW} and T_{NE} were found to be equal, is related to this and with our use of least squares: the origin $(0, 0)$ without difference between species is an extreme point and is the closest the model can get to many of the measurements.

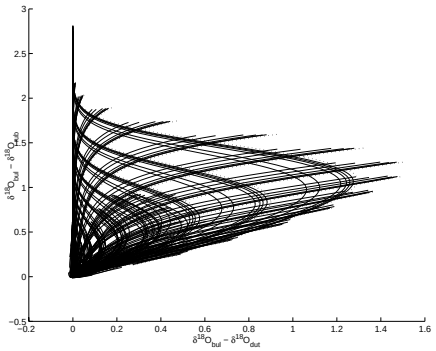
We propose to expand the model to be able to explain the observed measurements. Let us see what happens when we include food availability in the model. This might cause the production patterns to be changed



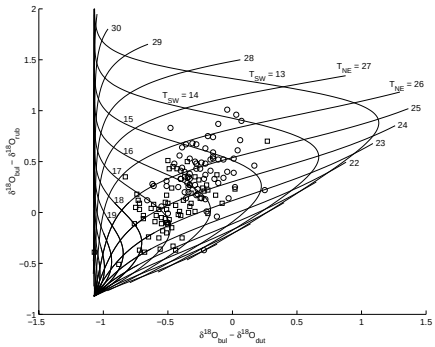
(a) Measurements and model predictions



(b) Mean temperature and variation



(c) The effect of food availability



(d) Measurements and model predictions with translation

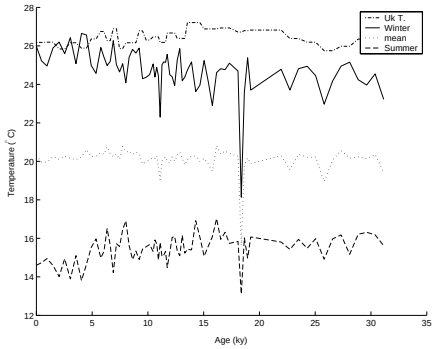
FIGURE 5. Space of δ^{18}O-differences, corrected for Vital Effect. The measurements from core 905P are indicated by circles, from core 929P by squares. (a) shows the measurements and model predictions for temperatures T_{NE} ranging from 15–25 °C, T_{SW} from 18–35 °C. (b) shows the same model predictions but as a function of mean temperature from 18–26 °C and seasonal difference from 0–14 °C. (c) shows the influence of food availability showing the same as (a) for $M_{SW} = 10, 20, 50, 10, 20, 50$ and $N_{mean} = 20, 20, 50, 50, 50, 100$ respectively. (d) shows the same as (a) with a translation (attributed to secondary calcification) by which all measurements are within the range of the model predictions.

such that they record other temperatures than was to be expected when temperature only would play a role. However, this seems to be of minor influence. In order to illustrate this, Figure 5(c) shows the model predictions for δ^{18}O-differences under several conditions. In fact, the same

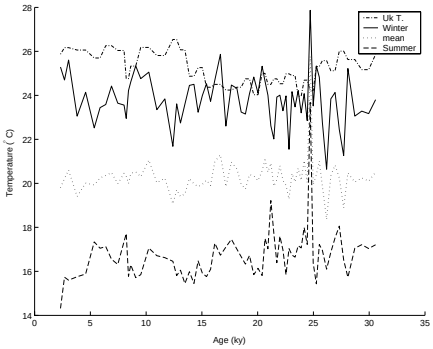
temperature ranges T_{SW} =15–25°C, T_{NE} =18–35°C as in Figure 5(a) are shown for six different values for the food parameters: $(M_{SW}, N_{mean}) = (10, 20), (20, 20), (50, 50), (10, 50), (20, 50)$ and $(50, 100)$. Dotted lines showing the predictions without food availability are also included. This figure is barely readable, but the important thing which should be noted from Figure 5(c) is that the region of possible δ^{18}O-differences covered by the model is hardly enlarged by the incorporation of food availability. Although the inclusion of food availability in the model is at present fairly rudimentary and could be made more sophisticated, the first indications are that the nutrients do not solve the problem!

Another explanation for the difference between our model results and the measurements is the fact that secondary calcification may occur, as is explained in the introduction, section 1.5. The model described so far, predicts the $\delta^{18}O_C$-value of shells that are free of secondary calcite. At the end of their life cycle, they settle down on the bottom of the ocean. In this period their oxygen isotope composition will increase due to secondary calcification. Therefore, the δ^{18}O-differences differ from the ones predicted by our model. To compensate for this, quantification of the process of secondary calcification for each species is necessary. Although accurate data are not available yet, a rough estimate of the correction terms needed is in the order of 0.5–1.0‰ for *G. ruber* and about 1‰ for *N. dutertrei*. The species *G. bulloides* however hardly experiences secondary calcification, so it can be expected that no correction term is needed for this species.

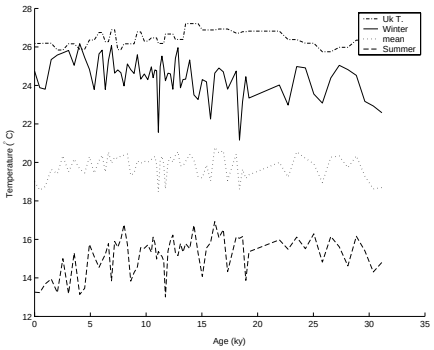
Comparing the model predictions to the observed measurements, Figure 5(a), we found that a minimum correction of 1.07‰ for the difference $\delta^{18}O_{bul} - \delta^{18}O_{dut}$ and 0.82‰ for $\delta^{18}O_{bul} - \delta^{18}O_{rub}$ is needed to ensure that all measurements are contained within the region of model predictions. These values compare well to the roughly estimated values. Because it is complicated to quantify the effect of secondary calcification with great accuracy, we will use these pragmatic correction terms for now. The result is shown in Figure 5(d). Of course, all measurements can be reached by the model now. The results for the model after correcting for secondary calcification are shown in Figure 6, for core 905P in the left column, for core 929P in the right one. The upper two plots 6(a) and 6(b) contain the results using model (2), including temperature effects only. Figure 6(c) and 6(d) show the results using model (3), incorporating food availability as well. In that case, we also solve for the parameters describing the food sources, M_{SW} and N_{mean}, which are shown in the lower two figures. For three measurements in core 929P the routine searching for the least squares minimum failed to converge. These points are marked by stars in Figure 6(d). Though marked as suspect, they do not seem to be very different from the other points.



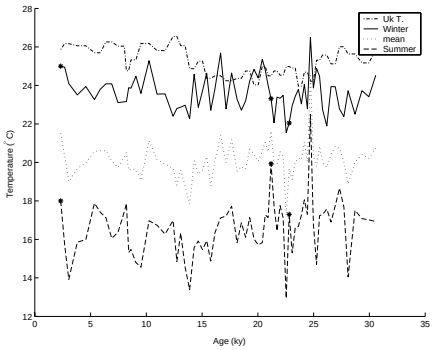
(a) Temperature results without food availability for core 905P



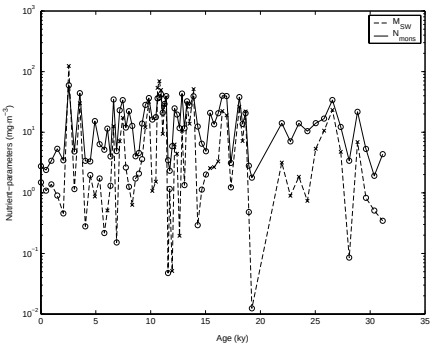
(b) Temperature results without food availability for core 929P



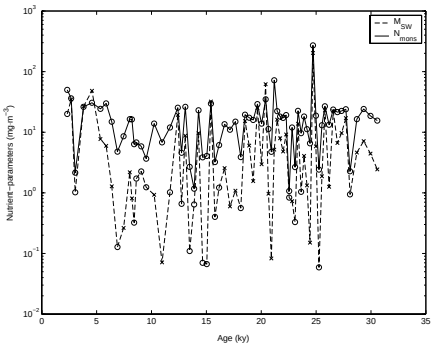
(c) Temperature results with food availability for core 905P



(d) Temperature results with food availability for core 929P



(e) Food availability, results for core 905P



(f) Food availability, results for core 929P

FIGURE 6. Results of the models after correcting for secondary calcification by a translation. The recovered sea surface temperatures using model (2), without food availability, are shown in (a) for core 905P and in (b) for core 929P. The temperatures using model (3), with food availability included, are shown in (c) for core 905P and in (d) for core 929P. The corresponding concentrations of phyto- and zooplankton are shown in (e) for core 905P and in (f) for core 929P.

The results shown in Figure 6 are much more plausible than those in Figure 4. The summer and winter temperatures T_{SW} and T_{NE} change a few degrees only and rapid oscillations as in Figure 4 do not occur. This is what should be expected. The results show an increase in seasonal variability from the glacial to the interglacial (present) period. This can be explained by the fact that there are indications that upwelling has increased in this period. This causes the sea surface temperature to be lower during the SW-monsoon, in the summer, because of enhanced upwelling. Moreover, the solar radiation is higher during the interglacial period. Therefore the temperature in the NE-monsoon period has increased. Both effects are clearly visible in core 905P. Core 929P is taken at a different spot in the Arabian Sea, where upwelling is less important. Cold water reaches this spot by advection from other regions. As a result, the decrease in SW-monsoon temperature is less pronounced.

For both cores the sea surface temperatures we find are lower than $U_{37}^{k'}$-temperatures. There may be several explanations for this. Foraminifera live at about 30 m depth, hence in a colder environment than immediately at the surface where $U_{37}^{k'}$-temperature is recorded. According to [**11**] about 1.3–1.7 $^{\circ}C$ should be added to the temperatures recorded by foraminifera. In that case, the $U_{37}^{k'}$-temperatures would be included in the yearly temperature range calculated through our model. Another explanation might be that the coccolithophorids—the microfossils used for $U_{37}^{k'}$-temperature—live in another season. If their ecological preference would be during the inter-monsoon period, they would certainly record higher temperatures than the foraminifera, which live mainly during the monsoon periods. Finally, it may well be necessary to improve our implementation of secondary calcification.

The sharp decrease in (winter) temperature in Figure 6(a) about 18 ky ago is diminished considerably in Figure 6(c). Apparently, this peak can be attributed to a nutrient effect not taken into account in Figure 6(a). The positive peak in both winter and summer temperature at about 25 ky ago in core 929P probably is an outlier.

The results in Figure 6(e) and 6(f) are fairly consistent with present-day measurements, which show the concentration of phytoplankton M to vary between 2–60 mg·m^{-3} whereas the concentration of zooplankton N is somewhat higher between 9–90 mg·m^{-3}. The model results show N_{mean} to be larger than M_{SW} as well.

Although the absolute value of M_{SW} and N_{mean} seems to be quite plausible, there is one aspect which should make the results in the figures 6(c), 6(d), 6(e) and 6(f) completely unreliable. The crosses in Figure 6(e) and 6(f) indicate that the corresponding value for M_{SW} found by the model is negative (the plot shows its absolute value)! This is clearly problematic from the

biological point of view and demonstrates that the model should be investigated further before we can draw conclusions from it. Finally, a quick glance at Figure 6 suggests that there is a correlation between the $U_{37}^{k'}$-temperature and the temperature T_{SW} during the SW-monsoon season. This needs further investigation as well.

3. A population dynamics model

The reconstruction of the sea surface temperature has so far been based on the assumption that the population adapts instantaneously to the changes in its environment. In this section we take a different approach and formulate a dynamic model consisting of a system of ordinary differential equations (ODEs) describing the populations of the three types of foraminifera and their food sources. Such an ODE model is very flexible and can easily be amended to include many different phenomena if desired. Here we restrict to the effects caused by changes in temperature and food availability.

In particular, during the southwest monsoon there is a strong growth in the populations of the foraminifera (and a corresponding increase in the flux of foraminifera shells to the sea floor [**1**]). This growth is caused by the upwelling of cold nutrient-rich water near the coast. The flux of some species depends strongly on the availability of these nutrients, while other species seem to respond mainly to the temperature variation. This information can be used to obtain a more detailed picture of the life cycle of the foraminifera.

The objective is to construct a population dynamics model for the various species of foraminifera and food sources. As is known in case of one single food source, there will be a survival of the fittest so that only one species will remain after some time ("competitive exclusion", see e.g. [**12**]). It is therefore crucial that the model encompasses at least two different food sources for the populations. We propose a population model for the three species *G. ruber* (R), *G. bulloides* (B), and *N. dutertrei* (D), feeding on two food sources (phyto- and zooplankton) which we call M and N as before. The aim of the model is to see what the effect of the upwelling during the southwest monsoon can be. As mentioned before, the ODE model also provides the opportunity to analyse the influence of additional effects, such as the northeast monsoon, the role of plankton as a primary food source or external effects such as the moon cycle.

We give here the simplest model with only growth and death terms for the populations R, B and D and their food sources M and N. We assume that R and D feed off N, while B and D live off M (see section 2.1.1). The growth of the species is determined by availability of food, by the carrying capacity of the species, and by the preferred temperature. We assume, as before, that each species has a preferred temperature, denoted $\bar{T}_R$ for *G. ruber*,

and a corresponding spread in temperature modelled by a standard deviation σ_R. The growth term of the food sources N and M is dominated by the upwelling, and we will model this as a function of the temperature change. We denote the growth functions by G_M and G_N for the moment and choose a specific function later. The ODE model then reads:

$$
\begin{aligned}
\dot{R} &= \alpha_R N \ell(R) \tfrac{1}{\sigma_R} e^{-(T-\bar{T}_R)^2/2\sigma_R^2} - \beta_R R, \\
\dot{B} &= \alpha_B M \ell(B) \tfrac{1}{\sigma_B} e^{-(T-\bar{T}_B)^2/2\sigma_B^2} - \beta_B B, \\
\dot{D} &= \tfrac{1}{2}(\alpha_{DM} M + \alpha_{DN} N) \ell(D) \tfrac{1}{\sigma_D} e^{-(T-\bar{T}_D)^2/2\sigma_D^2} - \beta_D D, \\
\dot{M} &= G_M(\dot{T}) - \mu_B M \ell(B) \tfrac{1}{\sigma_B} e^{-(T-\bar{T}_B)^2/2\sigma_B^2} \\
&\quad - \mu_{DM} M \ell(D) \tfrac{1}{\sigma_D} e^{-(T-\bar{T}_D)^2/2\sigma_D^2}, \\
\dot{N} &= G_N(\dot{T}) - \mu_R N \ell(R) \tfrac{1}{\sigma_R} e^{-(T-\bar{T}_R)^2/2\sigma_R^2} \\
&\quad - \mu_{DN} N \ell(D) \tfrac{1}{\sigma_D} e^{-(T-\bar{T}_D)^2/2\sigma_D^2}.
\end{aligned} \tag{11}
$$

The function ℓ in this model is usually assumed to be a simple linear term $\ell(R) = R$, or logistic $\ell(R) = R(k-R)$, but in order to control the population without restraining it entirely, we have chosen $\ell(R) = \frac{R}{1+R/k_R}$ with k_R the carrying capacity.

The model contains a large number of parameters, but some of the parameters can be scaled out while for others a typical physical value can be chosen. Since the life cycle of the populations is around 4 weeks, we choose all the β-values to be equal, $\beta = 13$ (since our unit of time is a year and 4 weeks is roughly one thirteenth of a year). The population of *G. bulloides* is around four times as large as that of *G. ruber* as well as *N. dutertrei*, under similar circumstances, so we choose $\alpha_R = \alpha_{DN}$ and $\alpha_B = 4\alpha_{DM}$. The data for the preferred temperatures and the sensitivities to the temperature are known from experiments, see the columns $\bar{T}$ and σ in Table 3. So now when scaling R with μ_R, B with μ_B, D with μ_{DN}, M with α_{DM}, N with α_{DN}, then this reduces the model to the following scaled version:

$$
\begin{aligned}
\dot{R} &= N \ell(R) \tfrac{1}{\sigma_R} e^{-(T-\bar{T}_R)^2/2\sigma_R^2} - \beta R, \\
\dot{B} &= 4M \ell(B) \tfrac{1}{\sigma_B} e^{-(T-\bar{T}_B)^2/2\sigma_B^2} - \beta B, \\
\dot{D} &= \tfrac{1}{2}(M+N) \ell(D) \tfrac{1}{\sigma_D} e^{-(T-\bar{T}_D)^2/2\sigma_D^2} - \beta D, \\
\dot{M} &= G_M(T) - M \ell(B) \tfrac{1}{\sigma_B} e^{-(T-\bar{T}_B)^2/2\sigma_B^2} - \tilde{\mu} M \ell(D) \tfrac{1}{\sigma_D} e^{-(T-\bar{T}_D)^2/2\sigma_D^2}, \\
\dot{N} &= G_N(T) - N \ell(R) \tfrac{1}{\sigma_R} e^{-(T-\bar{T}_R)^2/2\sigma_R^2} - N \ell(D) \tfrac{1}{\sigma_D} e^{-(T-\bar{T}_D)^2/2\sigma_D^2}.
\end{aligned} \tag{12}
$$

Here $\tilde{\mu} = \frac{\mu_{DM}}{\mu_{DN}}$ and the new carrying capacities are scaled by the corresponding μ-values.

Now the only terms that need to be specified are the temperature and the nutrient increase due to the upwelling. We couple this nutrient increase directly to a rise in food availability (because the inorganic nutrients are

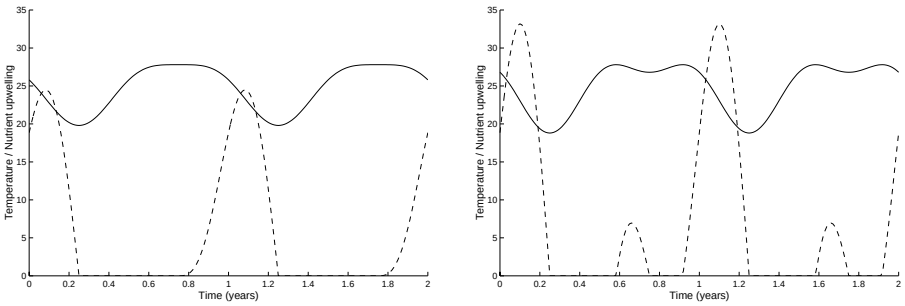


FIGURE 7. Temperature profile and nutrient upwelling for $\gamma = 0.5$ (left) and $\gamma = 1$. In both graphs $T_0 = 24.8°C$ and $T_h = 4°C$.

consumed by the plankton). Since the upwelling causes a decrease of the sea surface temperature, we choose to simplify the modelling of the upwelling by specifying a temperature profile and by making the nutrient growth a function of temperature decrease only. To be specific, we take

$$\begin{aligned} G_N(T) &= -\nu_N H(-\dot{T})\dot{T} \\ G_M(T) &= -\nu_M H(-\dot{T})\dot{T} \\ T(t) &= T_0 + \tfrac{T_h\gamma}{2} - T_h \sin 2\pi t(1 + \gamma \sin 2\pi t) \end{aligned} \tag{13}$$

with H the Heaviside function, ensuring that only temperature decrease is related to changes in the nutrients while temperature increase has no effect. The parameter γ is a measure for the relevance of the northeast monsoon. Its effect on the nutrient growth is shown in Figure 7.

We study the behaviour of the populations and try to find the mean temperature T_0 and the temperature variation T_h by matching the results from the ODE model with the data from the oxygen isotope compositions. The remaining parameters in the model are the three carrying capacities k, the value of $\tilde{\mu}$ (which we arbitrarily fix at 1) and the food growth parameters ν_N and ν_M (the coupling constants between nutrient upwelling and plankton increase in (13)). We describe various test results in the next section.

The sensitivities σ to the temperature can be obtained from measurements on their temperature ranges, see Table 3. The values of the other parameters are much harder to determine. Therefore we will follow a different approach. We will set these parameters in such a way that the model (12) and (13) describes the measurements of the current foraminifera levels as accurately as possible.

3.1. Dynamics. We are looking for solutions to (12) that are periodic with a period of one year. This means that we can derive all relevant quantities from this solution by integration over one year.

A priori one cannot expect that the system converges to such a periodic solution. However from the numerical simulations we performed it becomes

Species A	$T_{\min}$ $(^\circ C)$	$T_{\max}$ $(^\circ C)$	$\bar{T}_A$ $(^\circ C)$	σ_A $(^\circ C)$
G. ruber	15	30	25	2
G. bulloides	0	24	12	6
N. dutertrei	15	23	20	2.5

TABLE 3. Parameters describing the temperature dependence of the three different species under consideration. We note that these values are slightly different from those in Table 1 due to an a posteriori change in the value of the constants, a typical phenomenon for a studygroup problem indeed.

clear that for most parameter values there exist a periodic solution to (12). Furthermore the convergence towards this periodic solution is in general quite fast, see Figure 8. Note that in that simulation we started with initial conditions far from the final periodic solution.

We assume that the climate only changes over long periods. Over the relatively short period of the simulations there will be small variations in the climate. These variations however average out, so we will take the a fixed yearly temperature cycle. We remark that the fast convergence to the periodic solution means that the system adapts almost instantly to the changes in the climate.

Apart from the periodicity of the solution, we impose just one other condition which is derived from the present-day situation. Under the current climatological conditions ($T_0 = 24.8^\circ$C, $T_h = 4^\circ$C) we know that all three species are present. Therefore we adapt the food growth parameters ν in such a way that for this temperature profile, the three species coexist. This leads to the choice $\nu_N = \nu_M = 20$.

In Figure 9 we see that the three species react in different ways to the environmental conditions: *G. bulloides* and *N. dutertrei* have a single growth peak during the SW monsoon when the temperature is closest to their preferred temperatures. On the other hand the preferred temperature of *G. ruber* is met twice halfway the transition between the two monsoon periods. Therefore this species has two periods of growth. The second peak is smaller because there is less food available.

3.2. Getting temperature information from the model. During its life the plankton records the temperature of the surrounding sea water in its shell. We can find the average temperature that is recorded by integration over the population over time. Since the populations are periodic we only need to integrate over one year.

There are two different ways to describe the way in which the temperature is recorded in the shells. On the one hand we can assume that the shell is built over the full life span of the creature. We can find then the recorded

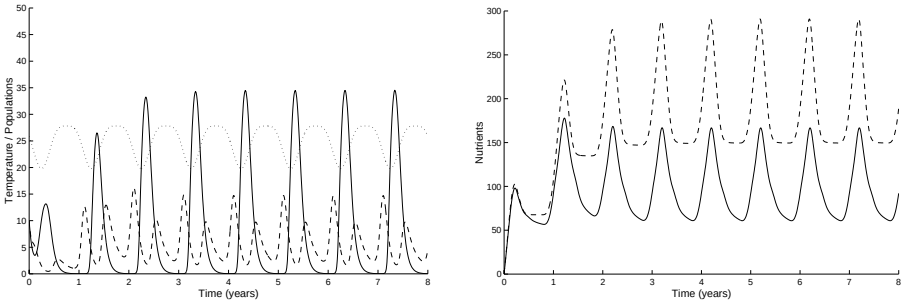


FIGURE 8. The convergence of the system towards a periodic solution, starting with arbitrary initial conditions. The left graph shows the population of *N. dutertrei* (solid) and of *G. ruber* (dashed). The dotted line shows the temperature profile. The graph on the right shows the food sources M (dashed) and N (solid) in the same simulation. Parameters: $\bar{T}_A$ and σ_A as in Table 3, $k = 10$, $\nu_N = 20$, $\nu_M = 20$. Initial values: $N = M = 1$, $R = B = D = 10$.

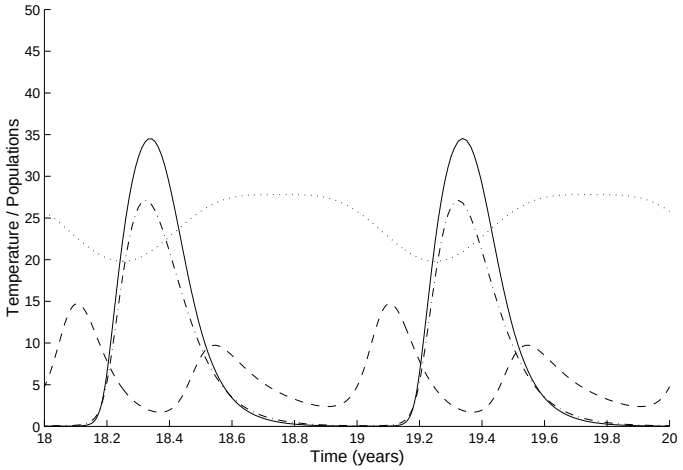


FIGURE 9. The final populations of *N. dutertrei* (solid), of *G. ruber* (dashed) and of *G. bulloides* (dash-dot) of the simulation in Figure 8. The dotted line shows the temperature profile.

temperature by calculating the average living temperature:

$$T_{rec} = \langle T \rangle_R = \frac{\int_{1year} R(t)T(t)dt}{\int_{1year} R(t)dt}. \tag{14}$$

On the other hand we can assume that the shell is only produced during a short period after the birth of the animal. To calculate the recorded

		using (14)		using (15)	
Species A	$\bar{T}_A$ ($^\circ C$)	T_{rec} ($^\circ C$)	$\overline{\delta_R} - \overline{\delta_A}$	T_{rec} ($^\circ C$)	$\overline{\delta_R} - \overline{\delta_A}$
G. ruber	25	24.56	-	25.01	-
G. bulloides	12	22.35	-0.48	20.94	-0.89
N. dutertrei	20	22.24	-0.51	20.75	-0.94

TABLE 4. Temperatures recorded by the three species. The differences in $\delta^{18}O$ are calculated using (1).

temperature we need to find the average reproduction temperature

$$T_{rec} = \langle T \rangle_R = \frac{\int_{1year} P_R(t)T(t)dt}{\int_{1year} P_R(t)dt}, \tag{15}$$

where

$$P_R(t) = N\ell(R)\,\frac{1}{\sigma_R} e^{-\frac{(T-\bar{T}_R)^2}{2\sigma_R^2}}.$$

The result of the first evaluation is less sensitive to temperature effects, because the slow decay ($e^{-\beta t}$) after a growth peak leads to an averaging out effect. In Table 4 we list the calculated temperatures for the simulation in Figure 9.

In Figure 10 we plot the recorded temperature as a function of the average sea water temperature T_0. The behaviour of the *G. ruber* population can be explained as follows. At low temperatures *G. ruber* will primarily grow during the NE monsoon when the temperatures are relatively high. Therefore at low temperatures *G. ruber* will record the maximum temperatures. At very high temperatures *G. ruber* will prefer the SW monsoon with its relatively low temperatures in combination with the large amount of nutrients present. In between these two regimes there is a transition around the optimal reproduction temperature of the species at 25°C. The decrease in the recorded temperature in the transition region is explained by the availability of food. These effect are less pronounced in the left graph because of the averaging out effect.

3.3. Reconstruction of historic temperatures. To use this model for reconstruction of historic temperatures we perform a series of simulations on a grid of values for T_0 and T_h. From each of the calculated populations we can then calculate the $\delta^{18}O$ values recorded by these populations. These values are matched to the measured values. Note that since we know only two values from the measurements, we can only reconstruct two of the three variables we used to describe the temperature profile (13). Therefore we need to fix the value of γ and we cannot reconstruct the relative strength of the NE monsoon.

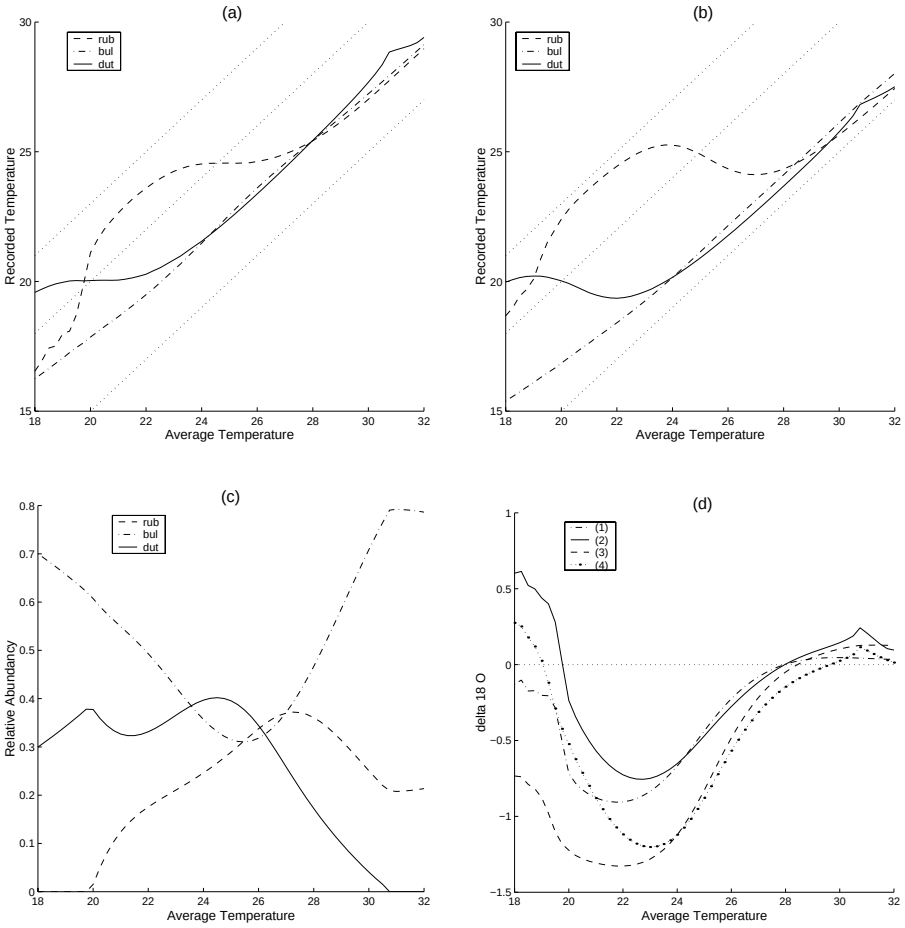


FIGURE 10. Response of the system to changes in the average temperature T_0. All other parameters are kept constant. (a) and (b) shows the average temperatures recorded by the three species, (a) is calculated using (14) and (b) is calculated using (15). The dotted lines in these figures show the average temperature and the minimum and maximum temperatures. (c) shows the relative abundance of the species and (d) shows the resulting differences in δ^{18}O-values: (1) $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_B$ using (14); (2) $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_D$ using (14); (3) $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_B$ using (15); (4) $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_D$ using (15).

As a demonstration of this method we calculated the two $\delta^{18}O$-values on a uniform rectangular grid in (T_0, T_h) space. In Figure 11 we draw the contour graphs of the two resulting δ^{18}O differences. So, if for example measurements give values of $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_B = -0.75$ and $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_D = -0.50$ we can reconstruct the historic temperature at the intersection of the corresponding contours in Figure 11. This approximately results in $T_0 =$

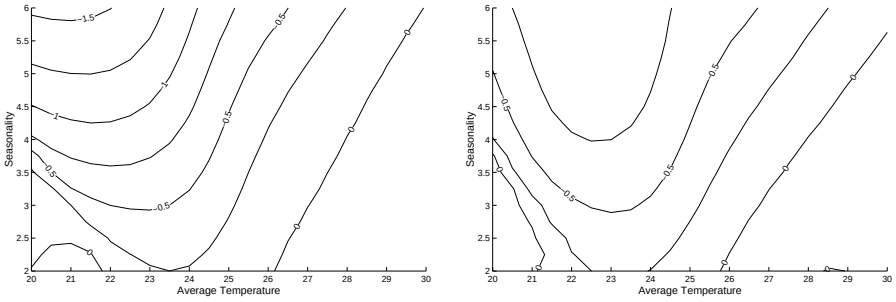


FIGURE 11. Contour graphs showing the effect of the average temperature and seasonality on the $\delta^{18}O$ values. All other parameters are kept constant. The left graph shows $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_B$, the graph on the right shows $\delta^{18}\mathrm{O}_R - \delta^{18}\mathrm{O}_D$. Notice that these are related to the δ^{18}O-differences in Figure 5(a) via a linear transformation.

21°C and $T_h = 3.7$°C. Note that in this demonstration it is possible to estimate such intersection points only in a small part of the (T_0, T_h) domain. In most of the domain the contour lines are practically parallel, which makes it difficult to determine where they intersect, yielding inaccurate results.

It is possible to use the dynamic model to reconstruct the historic sea surface temperature analogous to the detailed results obtained for the basic model in section 2 (see Figures 6(a) and 6(b)), but we have not been able to pursue that here because of limited time resources.

4. Discussion

Several detailed suggestions for further research are spread throughout the paper. Here we draw some more general conclusions.

- The interdisciplinary approach has been very beneficial in gathering new viewpoints. The use of even a fairly simple mathematical model can make a significant contribution in the determination of the sea surface temperature on the basis of the oxygen isotope composition of foraminifera.
- The model proposed in section 2, which incorporates seasonal effects, is a substantial improvement on models in which only the average temperature is considered.
- In order to obtain reliable quantitative information from the mathematical model(s) it is essential that the ecology of the foraminifera is understood in more detail.
- The incorporation in the model of effects related to nutrients and food availability does not lead to improvements; in fact the outcomes give unrealistic (negative) amounts of plankton. This may be due to the way the food availability is incorporated in the model; further investigation is necessary.

- The process of secondary calcification (see section 1.5) can explain some of the present discrepancies between the experimental data and the outcomes of the model. However, quantitative information about this process is needed before a final conclusion can be drawn on this matter.
- Overall, the basic model of section 2 seems reasonable in its simplicity and the results agree (at least) qualitatively with what is known and expected. As a next step we suggest that an attempt is made to obtain more reliable values for the parameters in the model.
- The dynamic model, a system of coupled ordinary differential equations (see section 3), is a nice mathematical toy to investigate the influence of a variety of effects. The model is robust in the sense that the solution converges quickly to a yearly periodic cycle for a large range of parameter values. In principle it can be used to estimate the temperature and its variability from the experimental data. However, at present the level of sophistication of the ODE model is not in line with the relatively poorly constrained parameters derived from field observations and experimental data.
- From a scientific point of view it is essential that the sizes of the errors in both the experimental data and the values of the ecological parameters are determined. They should be taken into account when judging the reliability of the results of the analysis. This is also related to the amount of information that can be extracted reliably from the data.

Acknowledgement
We would like to thank J. Hulshof and D. Nitzpon for their contributions to the discussions. We thank G.-J.A. Brummer for comments and helpful suggestions on a draft version of this manuscript.

Bibliography

[1] S.M.-H. Conan and G.-J.A. Brummer. *Fluxes of planktic foraminifera in response to monsoonal upwelling on the Somalia Basin margin.* Deep Sea Research, part II, 47: 2207–2227, 2000.

[2] Ch. Hemleben, M. Spindler and O.R. Anderson. *Modern planktonic foraminifera.* Springer-Verlag, Berlin, 363 pp., 1989.

[3] E.M. Ivanova. *Late Quaternary monsoon history and paleoproductivity of the western Arabian Sea.* Ph.D.-thesis, Vrije Universiteit, Amsterdam, The Netherlands, 172 pp., 1999.

[4] S.-T. Kim and J.R. O'Neil. *Equilibrium and non-equilibrium oxygen isotope effects in synthetic carbonates.* Geochimica Cosmochimica Acta, 61, 3461–3475, 1997.

[5] G.P. Lohmann. *A model for variation in the chemistry of planktonic foraminifera due to secondary calcification and selective dissolution.* Paleoceanography, 10(3): 445–457, 1995.

[6] B.A. Malmgren, M. Kucera, J. Nyberg and C. Waelbroeck. *Comparison of statistical and artificial neural network techniques for estimating past sea surface temperatures from planktonic foraminifer census data.* Paleoceanography, 16, 1–11, 2001.

[7] J.P. McCreary, K.E. Kohler, R.R. Hood and D.B. Olson. *A four-component ecosystem model of biological activity in the Arabian Sea.* Progress in Oceanography, 37: 193–240, 1996.

[8] A. Mix. *The oxygen-isotope record of glaciation.* In: W.F. Ruddiman, H.E. Wright (Eds.), North America and adjacent oceans during the last deglaciation (The Geology of North America, Vol. K-3). The Geological Society of America, 1987.

[9] S. Mulitza, T. Wolff, J. Pätzold, H. Hale and G. Wefer. *Temperature sensitivity of planktic foraminifera and its influence on the oxygen isotope record.* Marine Micropaleontology, 33:223–240, 1998.

[10] F.J.C. Peeters *The distribution and stable isotopic composition of living planktic foraminifera in relation to seasonal changes in the Arabian Sea.* Ph.D.-thesis, Vrije Universiteit, Amsterdam, The Netherlands. pp.184, 2000.

[11] F.J.C. Peeters, G.-J.A. Brummer and G.M. Ganssen. *The effect of upwelling on the distribution and stable isotope composition of Globerina bulloides and Globigerinoides ruber (planktic foraminifera) in modern surface waters of the NW Arabian Sea.* Global and Planetary Change, in press.

[12] E.R. Pianka. *Competition and niche theory.* In: R.M. May (Ed.), Theoretical Ecology: Principles and Applications. Oxford: Blackwells Scientific 1981, pp. 167–196.

[13] G.A. Schmidt and S. Mulitza. *Global calibration of ecological models for planktic foraminifera from coretop carbonate oxygen-18.* Marine Micropaleontology, 44, 2002.

[14] H.C. Urey *The thermodynamic properties of isotopic substances.* Journal of the Chemical Society, 562–581, 1947.