

# **Rationing Model : The development of a tool to help managers visualize the impacts of rationing and other decisions**

Anton Sparks, Ninham Shand (Pty) Ltd, PO Box 1347, Cape Town, 8000.

Geoff Pegram, Civil Engineering, University of Natal, Durban, 4041.

Enoch Dlamini, KOBWA, PO Box 5085, Nelspruit, 1200

## CONTENTS

1	Introduction to the Komati system.....	1
2	Rationing in the Komati System.....	4
3	Stochastic generator.....	5
3.1	Overview.....	5
3.2	Tests.....	5
3.3	Constraints .....	6
4	The rationing process .....	8
5	Defining scenarios.....	11
6	graphical output.....	11
7	Experiences during the current drought .....	13
7.1	Severity of the current drought.....	13
7.1.1	Rainfall .....	13
7.1.2	Streamflow .....	13
7.1.3	Storage .....	14
7.2	Comparison of the “Forecasting” and “Historical” streamflows .....	14
7.3	Comparison of the “Planning” and the “Forecasting” streamflow generators .....	15
7.4	Effect of the antecedent flows .....	16
7.5	Progressive curtailment .....	17
8	Conclusions.....	18
9	Acknowledgements .....	18
10	References .....	19

## TABLES

Table 1.1	: Current authorized domestic, industrial and irrigation demands in the Komati system (assuming no rationing).....	3
Table 7.1	: Alternative antecedent flows.....	17

## FIGURES

Figure 1.1	: The Komati and Lomati Rivers Management System .....	2
Figure 1.2	: Diagrammatic layout of the Komati / Lomati system.....	3
Figure 3.1	: Comparing Forecast sequences with the actual historical sequences.....	7
Figure 4.1	: Overview of the rationing process.....	9
Figure 4.2	: Scenario analysis .....	10
Figure 6.1	: Storage trajectories assuming no curtailment.....	12
Figure 6.2	: 30% curtailment of supply to irrigators in Oct 2002 if extreme (worse than a 1 in 100 year drought) conditions are experienced.....	12
Figure 6.3	: Cumulative demand plot to verify system demands .....	12
Figure 7.1	: Comparing cumulative rainfall in the Komati Catchment with the long-term average .....	13
Figure 7.2	: Comparing the current historical inflows with the median inflows to Maguga Dam .	13
Figure 7.3	: Comparing the cumulative current historical inflows with the median inflows to Maguga Dam .....	14
Figure 7.4	: Percentage full of the Driekoppies and Maguga Dams .....	14
Figure 7.5	: Comparing the trajectory of the combined storage in Driekoppies and Maguga Dams since February 2003 under current conditions (red dots) with trajectories assuming the inflows of 1925(black), 1930(red) and 1991(blue) .....	15
Figure 7.6	: Comparison of the 1 in 50 storage trajectory for the combined Maguga and Driekoppies Dams using the “annual” and the “forecasting” stochastic generator.....	16
Figure 7.7	: Comparison of the 1 in 50 storage trajectory for the Driekoppies Dam using the “annual” and the “forecasting” stochastic generator .....	16
Figure 7.8	: Forecasts using different antecedent conditions. (2% forecasts: red(actual), 2% (blue), decreasing (green), median (black) .....	17

## APPENDIX

### A Monthly Flow Generator / Forecaster for Operational Decision-Making - Geoffrey GS Pegram

## E1 ABSTRACT

The Komati Basin Water Authority (KOBWA) identified the need for an interactive tool to demonstrate the consequences of different rationing scenarios to their consumers.

The existing Water Resources Yield Model (WRYM), which simulates inflows, storage and demands on a monthly timestep was enhanced by adding the following:

- **“Forecasting” stochastic generator.** The model is used on a monthly basis using the known initial storage and estimates of possible inflows and future demands on the system to simulate the system’s future storage and supply. The future streamflows are affected by the current state of the system. If antecedent conditions are wet the future streamflows are more likely to also be wet and vice-versa. In the Komati a significant fraction of the streamflows supplying irrigation accrue downstream of the dams. To take this into account, an autoregressive streamflow generation algorithm, whose system flow estimates are conditioned on the streamflows in the catchment during the four preceding months, was developed.
- **Factoring of demands.** Demands can now be factored on a given date to simulate the curtailment or growth of demands.
- **Simplified input.** A pre-processor enables a manager to define rationing scenarios.
- **Graphical output.** A postprocessor displays storage and supply probabilities

This rationing model is being used during the current drought in the Komati Basin to recommend the levels of restrictions. This paper discusses experiences in the use of the model for rationing.

# 1 INTRODUCTION TO THE KOMATI SYSTEM

The Komati River flows through South Africa, Swaziland and Mozambique and the construction of the additional dams in the fertile Komati valley would not have been possible without agreements between these countries. The Komati Treaty resulted in the establishment of the KOMati Basin WATER Authority (KOBWA), to manage both the construction of the dams and later the riverine releases from these dams to South Africa, Swaziland and indirectly to Mozambique.

KOBWA currently manages the Driekoppies Dam, (completed 1998 with a storage capacity of 251 million m<sup>3</sup> and an average inflow of about 200 million m<sup>3</sup>/a), and the Maguga Dam, (completed in 2001 with a storage capacity of 332 million m<sup>3</sup> and an average inflow of 500 million m<sup>3</sup>/a).

Unfortunately, the Maguga Dam is still filling and the two dams cannot yet be operated in an integrated manner. In addition, the supply to existing users downstream of the dams cannot be prejudiced while Maguga Dam is filling. In principle, the Maguga Dam must fill “passively” and cannot impound water until the requirements of the authorized downstream consumers have been satisfied. However, no releases from storage are made to existing consumers downstream of Maguga Dam so they are effectively supplied from run-of-river streamflows – as though Maguga Dam did not exist. This means that existing users downstream of Maguga are curtailed implicitly by the reduction of the streamflow into Maguga Dam during droughts while the users supplied from Driekoppies are curtailed explicitly by reducing the allocation from that dam.

In practice this “passive” rule might have resulted in unnecessary spillage. Initially, when Maguga started filling, Driekoppies Dam was full and could have spilt had wet conditions been experienced. An alternative “active” filling strategy was developed to reduce the risk of spillage from the Driekoppies Dam, in which the Driekoppies Dam provided water to consumers downstream of the Lomati / Komati confluence on behalf of Maguga Dam to encourage Maguga to fill. The understanding was that the additional water accruing in Maguga Dam, as a result of the additional releases from Driekoppies Dam, would be supplied from Maguga Dam to the consumers downstream of the Lomati / Komati confluence if necessary. The current drought has forced Maguga Dam to release these accruals to the consumers downstream of the Lomati / Komati confluence. Once Maguga Dam has released its “liability” to Driekoppies Dam the storages in the system would be as though the “passive” filling rule had been maintained for the full filling period. For this reason, when the system was modelled to simulate the reliability of supply under varying drought severities, a “passive” operating rule was used throughout the filling period. This necessitated adjusting the initial storages in the dams by transferring any accruals in Maguga Dam to Driekoppies Dam.

The location of the river reaches of the Komati system is described schematically in Figure 1.2 and the magnitudes are summarized in Table 1.1. The demand downstream of the confluence of the Lomati and Komati Rivers is about 150 million m<sup>3</sup>/a and can be supplied from either dam. The relative supply from the dams can be varied to reduce the risk of one of the dams emptying or spilling by itself.

The Komati Treaty makes provision for curtailment of the various consumers. Certain consumers such as irrigators have accepted a lower assurance of supply to increase their allocation. The two categories of consumer are defined as:

*High Assurance : means a 2% risk in any one year of only partial availability*

*Low Assurance : means a total unavailability for up to 20% of the time on average in respect of 30% and a 2% risk in any one year of only partial availability in respect of the remaining 70%*

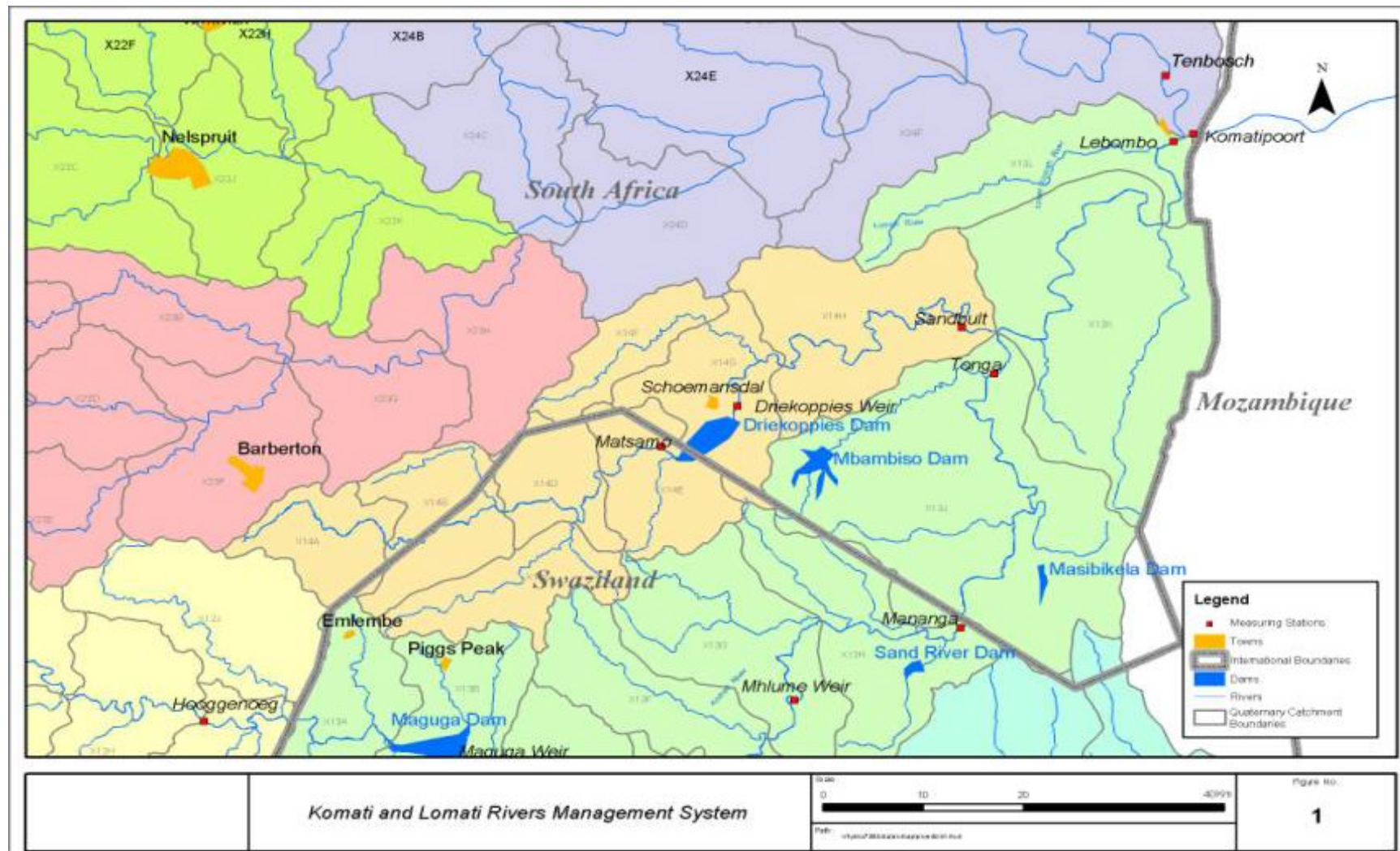


Figure 1.1 : The Komati and Lomati Rivers Management System

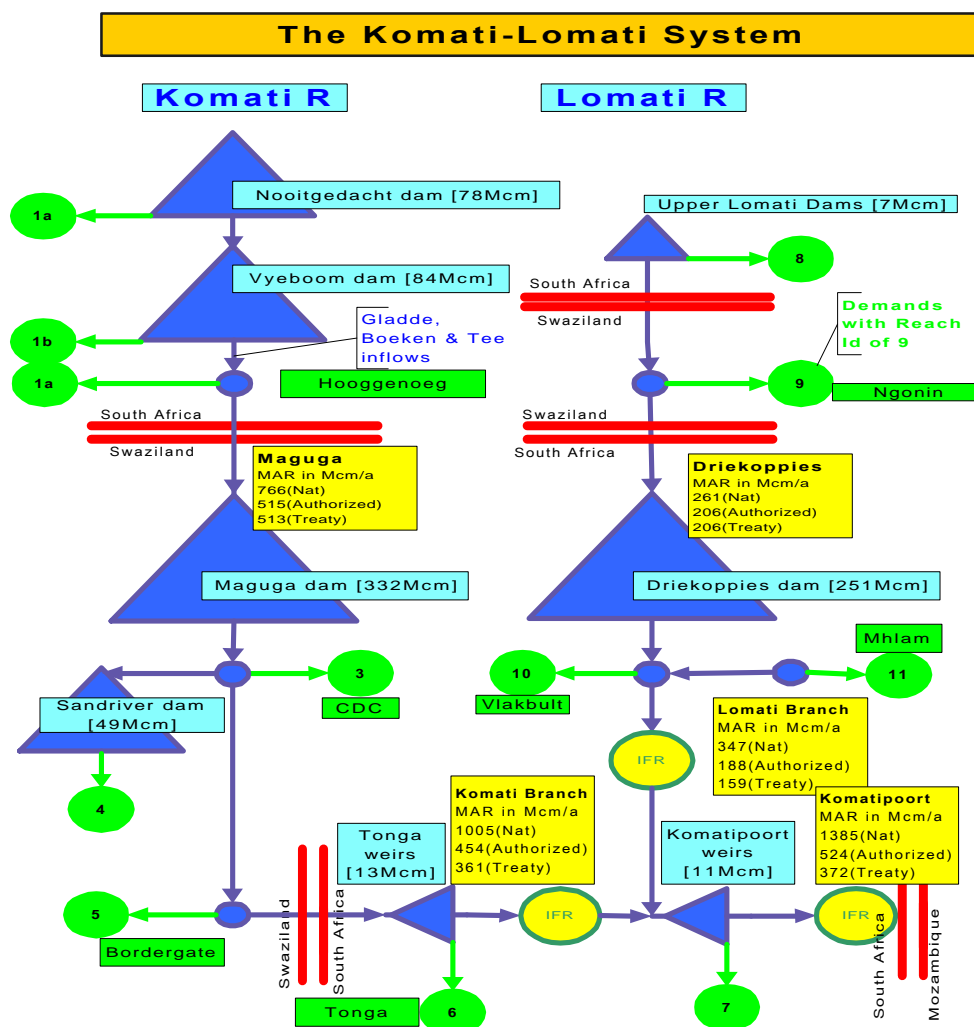


Figure 1.2 : Diagrammatic layout of the Komati / Lomati system

Table 1.1 : Current authorized domestic, industrial and irrigation demands in the Komati system (assuming no rationing)

Location	Reach	WRYM Node	Domestic / Industrial	Irrigation Allocations	Evapo-transpiration losses	Total
Nootgedacht	1a	2	29.2			29.2
Vyebboom	1b	4	99.5	23.8		123.3
Hooggenoeg	2	14	5.4	9.0		14.4
Ds Maguga	new RSA		0.0	22.5		22.5
Tonga	6	50	9.3	56.0	9.9	75.2
ds Lomati/ Komati confluence	7	46	0.2	140.4	9.9	150.5
Upper Lomati	8	32	3.2	4.6		7.8
Vlakkbult	10	42	8.6	71.9	11.1	91.6
Lomati	11	38		18.7		18.7
<b>sub-total</b>			<b>155.4</b>	<b>346.9</b>	<b>30.8</b>	<b>533.1</b>
CDC	3	20	1.9	2.8		4.7
Mhlume / Sandriver	4	48	11.8	168.2		180.0
Ds Maguga	new Swaziland	21	0.0	18.9		18.9
Bordergate	5	28	1.9	0.0		1.9
Ngonini	9	34	1.6	6.1		7.7
<b>sub-total</b>			<b>17.2</b>	<b>196.0</b>	<b>0.0</b>	<b>213.2</b>
<b>Total</b>			<b>172.6</b>	<b>542.9</b>	<b>30.8</b>	<b>746.3</b>

- At present the abstraction by Mhlume Water is controlled by the daily abstraction rules at the weir -not be the annual allocation and the agricultural demand was capped at 200Million m<sup>3</sup>/a rather than 168.2 million m<sup>3</sup>/a.
- The Komati treaty makes provision for additional releases of 30.7 to South Africa (in addition to the new volume of 22.5) and 66.1 to Swaziland (In addition to the new volume of 18.9 above) giving a total increase of 96.8 million over and above the 41.4 increase to date.
- Of the demands in Table 1.1 about 92 million m<sup>3</sup>/a can only be supplied from Driekoppies Dam (reach 10), about 303 million m<sup>3</sup>/a can be supplied from Maguga Dam (reaches 3,4,5,6 and new demands) and 150 million m<sup>3</sup>/a could be supplied from both dams (reach 7) when they are operated in an integrated manner after the filling of Maguga Dam

## 2 RATIONING IN THE KOMATI SYSTEM

The Rationing Model interactively demonstrates the effects of varying the timing and severity of water supply rationing on the reliability and quantity of water supply to enable the selection of an appropriate rationing strategy.

This model is used to make only one decision: by how much and when must the supply to different consumers be rationed/reduced to avoid the dams emptying and the resultant catastrophic failure. The following were considered in the development of this tool:

- **Forecasting Model.** The major irrigation demands coincide (hopefully) with the rainfall season and the best possible estimate of this streamflow will improve the estimate of the need for restrictions. Dry antecedent conditions should reduce the anticipated streamflow as the catchment will absorb a greater proportion of the rainfall. The average inflow downstream of the Driekoppies and Maguga Dams is equivalent to about 68% of the irrigation demand. During the dry periods a greater portion of the demand will be supplied from the major dams and the inflow to these dams will be reduced. Both factors will cause the dams to draw down rapidly during dry periods.  
The existing “annual” or “planning” stochastic generator built into the standard Water Resources Yield Model (WRYM) program does not use the current streamflow conditions to generate its first stochastic streamflow, though subsequent streamflows respect the serial correlation between successive years. For this reason, a stochastic forecasting model considering antecedent streamflow conditions was developed.
- **Flexible Operating Rules.** The operation of the system will change with time. Initially Maguga will be filled passively but at a later stage the Driekoppies and Maguga Dams will be operated conjunctively. The simulation model must be flexible enough to allow the operating rule to change.
- **Changing Demands (growing/curtailing).** The demands will not be static as the treaty allows for an increase of about 140 million m<sup>3</sup>/a in the demands. It should also be easy to specify and curtail groups of demands eg all irrigation on the Komati River downstream of the Maguga Dam and upstream of the Komati / Lomati confluence.
- **Simple operation.** The model should be suitable for routine operation on a regular basis to identify problems early and be suitable for semi-interactive analysis. After specifying the initial conditions, such as the current dam storage, the antecedent flows and the current demands and possibly a trial curtailment level the model should indicate the likely storage trajectories of the dams and enable an assessment of the risks. To simplify the operation it would not be necessary to produce short term curves.

Based on these requirements it was decided to base the Rationing Model on the Water Resources Yield Model (WRYM) and introduce the following enhancements:

- Include a stochastic generator using antecedent flow conditions to make streamflow forecasts
- Allow the factoring of demands
- Provide a pre-processor to simplify input and a post-processor to view output.

### 3 STOCHASTIC GENERATOR

#### 3.1 Overview

The forecasting streamflow stochastic generator was developed specifically for the Komati and is described more fully in Annexure A: Monthly Flow Generator / Forecaster for Operational Decision-Making - Geoffrey GS Pegram. If appropriate parameters are determined for other regions then the forecasting generator can also be used for those regions.

An autoregressive streamflow model was designed, in which the streamflow forecast for the  $i^{\text{th}}$  sequence (each of 16 streamflow records) is based on:

- its four antecedent flows  $z_{i,t-4}$  to  $z_{i,t-1}$ .
- four correlation parameters  $f_{ij}$ ,  $j = 1, 2, 3, 4$  for each antecedent flow
- a noise term  $s_{ai} a_{i,t}$  which introduces variability amongst the different stochastic sequences for each streamflow

The four correlation parameters  $f_{ij}$ ,  $j = 1, 2, 3, 4$  and the standard deviation of the noise term  $s_{ai}$  for each sequence  $i = 1, 2, \dots, 16$ , were extracted. The time series model adopted for each sequence was thus

$$z_{it} = f_{i,1} z_{i,t-1} + f_{i,2} z_{i,t-2} + \dots + f_{i,4} z_{i,t-4} + s_{ai} a_{i,t}$$

where  $a_{i,t}$  are the residuals of the model, or the noise term, assumed to be distributed as  $N(0, s_a^2)$ . These residuals  $\{a_{i,t}\}$  were automatically extracted as part of the model fitting procedure. They were exported and subsequently analysed off-line, to extract their cross correlation matrix.

#### 3.2 Tests

Various tests were performed on the forecasting generator including checking the following properties against the historical sequences:

- Mean
- Standard deviation
- Serial correlation
- Cross-correlation

The yield-storage relationship of the forecast sequences was also compared with the annual stochastic generator. These tests are detailed in Annexure A.

In addition, another test was used to compare year-long stochastic forecasts, seeded using actual historical values, with the subsequent historical flows. The stochastic forecasts were ranked and used to define categories. The frequency with which the historical flows fell within a category was checked to see if it was reasonable.

The 75 year long naturalized historical streamflow records of the Komati were combined into record representing the total inflow from the area. This record was subdivided into sets of 24 month long periods, commencing in February. The forecasting stochastic generator was used to forecast 41 monthly flow sequences each 24 months long from hydro years 1922 to 1994 using the historical antecedent flows. (ie. from February 1922 to January 1924, February 1923 to January 1925 ...and lastly February 1994 to January 1996). The related 24 month long historical and forecast flows were imported into Excel for comparison.

#### Preparing Exceedence Category Plots



- In Excel, the 41 forecast flows for each of the 24 months in each sequence were ranked and the values corresponding with various exceedence values of interest (0%, 5%, 10%, 25%, 50%, 75%, 90%, 95% and 100%) were extracted.
- The exceedence values were used to create boundaries for exceedence probability ranges. The actual historical value was compared with these probability ranges, and the range within which the historical flow fell was identified for each month and year in the record.
- The occurrences were cumulated by exceedence category and month and plotted as per Figure 3.1-Aa (attached).
- A similar process was followed using the cumulative monthly flows for the 24 month period starting in February (see Figure 3.1-Ba).

### Preparing Streamflow Plots

- For each 24 month long period the percentiles of the historical monthly flows and the forecast flows derived above are plotted in Figure 3.1-Ab.
- The cumulative total flows starting from February are plotted in Figure 3.1-Bb.

### Interpretation

In Figure 3.1 Aa) the number of historical inflows less than the median stochastic forecast inflow is indicated by the height of the column below the light green / dark green interface. If this interface corresponds with 50% then half of the historical monthly inflows are less than the corresponding forecast value. In Figure 3.1 Aa) the number of historical values less than the stochastic forecast varies from 45% (Month 11 – December) to 75% (Month 4 – May).

The relative distribution of the cumulative historical and forecast inflows since February are compared Figure 3.1 Ba). Note that in the second year about half the historical flows are less than the stochastic forecast as one would expect. However, the spread of the forecast values is greater than the historical in this latter year. For instance, in the second year no historical streamflows are less than the lowest 10% of the forecast streamflows and the number of historical streamflows greater than the upper 25% of the forecast streamflows is less than expected.

The impact of this in the cumulative flows in Figure 3.1 is less than expected. The correspondence between the historical and forecast flows is very good except for the extreme values. After two years the lowest stochastic forecast streamflows are 30% less than the historical values and the highest stochastic flows are about 40% higher than the historical values. For the lowest streamflows the correspondence between the forecast and the historical inflows was high up till 21 months had elapsed, before the onset of the second period of summer rainfall. The forecasting model should produce reliable answers over this 21 month period.

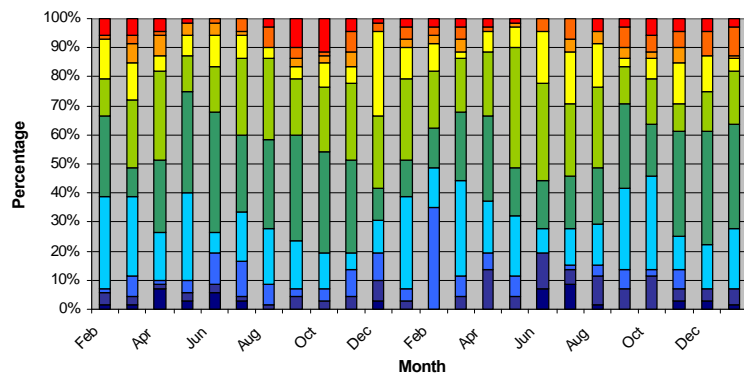
## 3.3 Constraints

To avoid unnecessary costs the model had the following constraints:

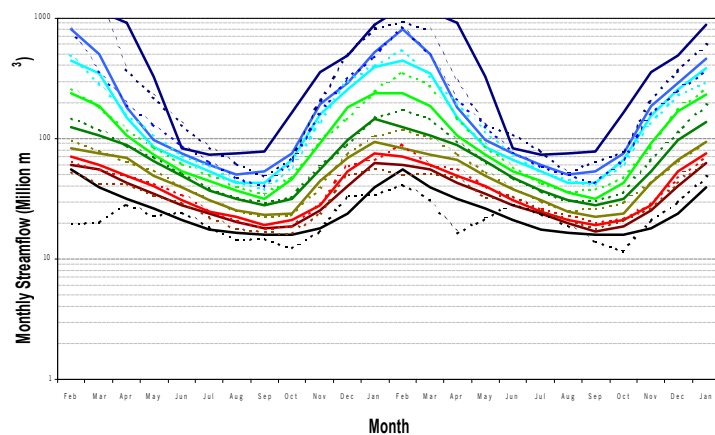
- requires application to non zero streamflows
- uses average annual cross-correlation. Some cross correlations appear underestimated which could result in the yield of integrated systems being overestimated.
- it does not model annual serial correlation and is valid for a relatively short period (up to say 18 months) and the stochastic generator based on annual streamflows that is already included with the WRYM should be used for longer projections. The technology exists to create a hybrid
- the calculation of the autoregressive parameters is not automated

## Monthly flows

**Aa) Exceedance Category : System's Monthly Streamflows  
from the 1<sup>st</sup> forecast in February**

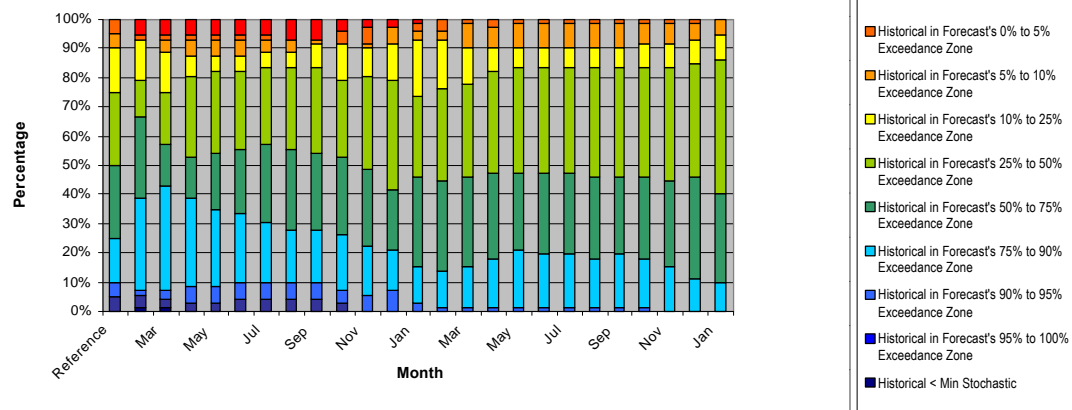


**Ab) System Streamflows with 1<sup>st</sup> forecast estimate in February**

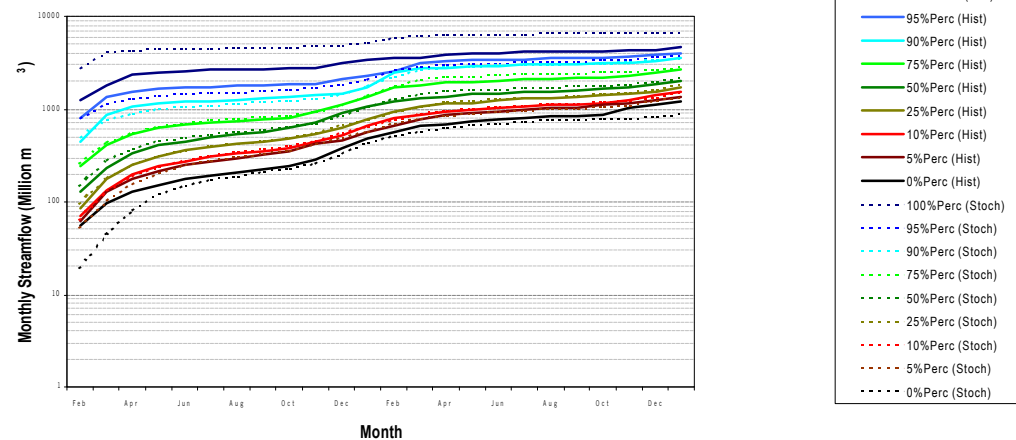


## Cumulative flows

**Ba) Exceedance Category : Cumulative system streamflows  
from the 1<sup>st</sup> forecast in February**



**Bb) Cumulative system streamflows with 1<sup>st</sup> forecast estimate in February**

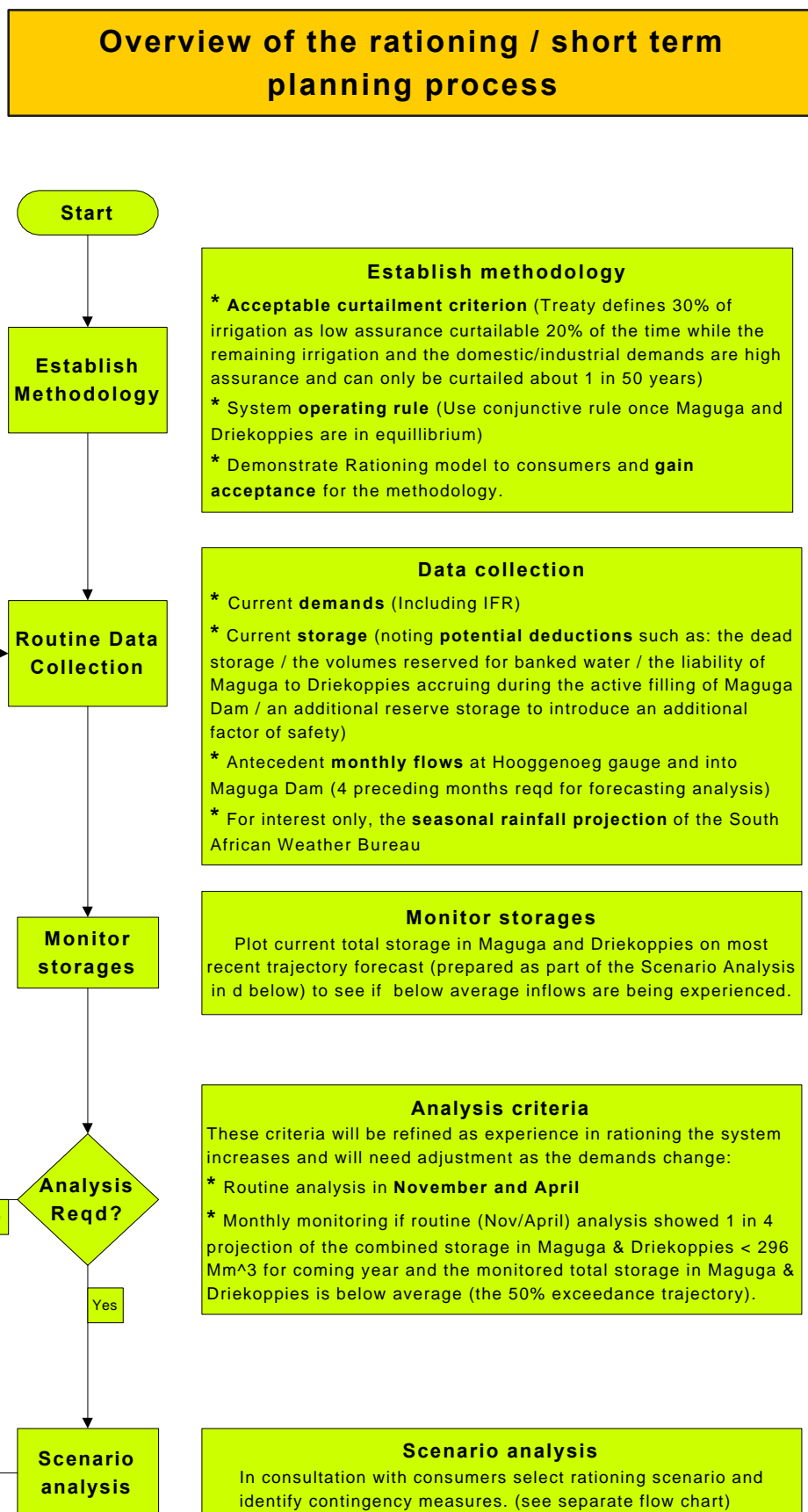


**Figure 3.1 : Comparing Forecast sequences with the actual historical sequences**

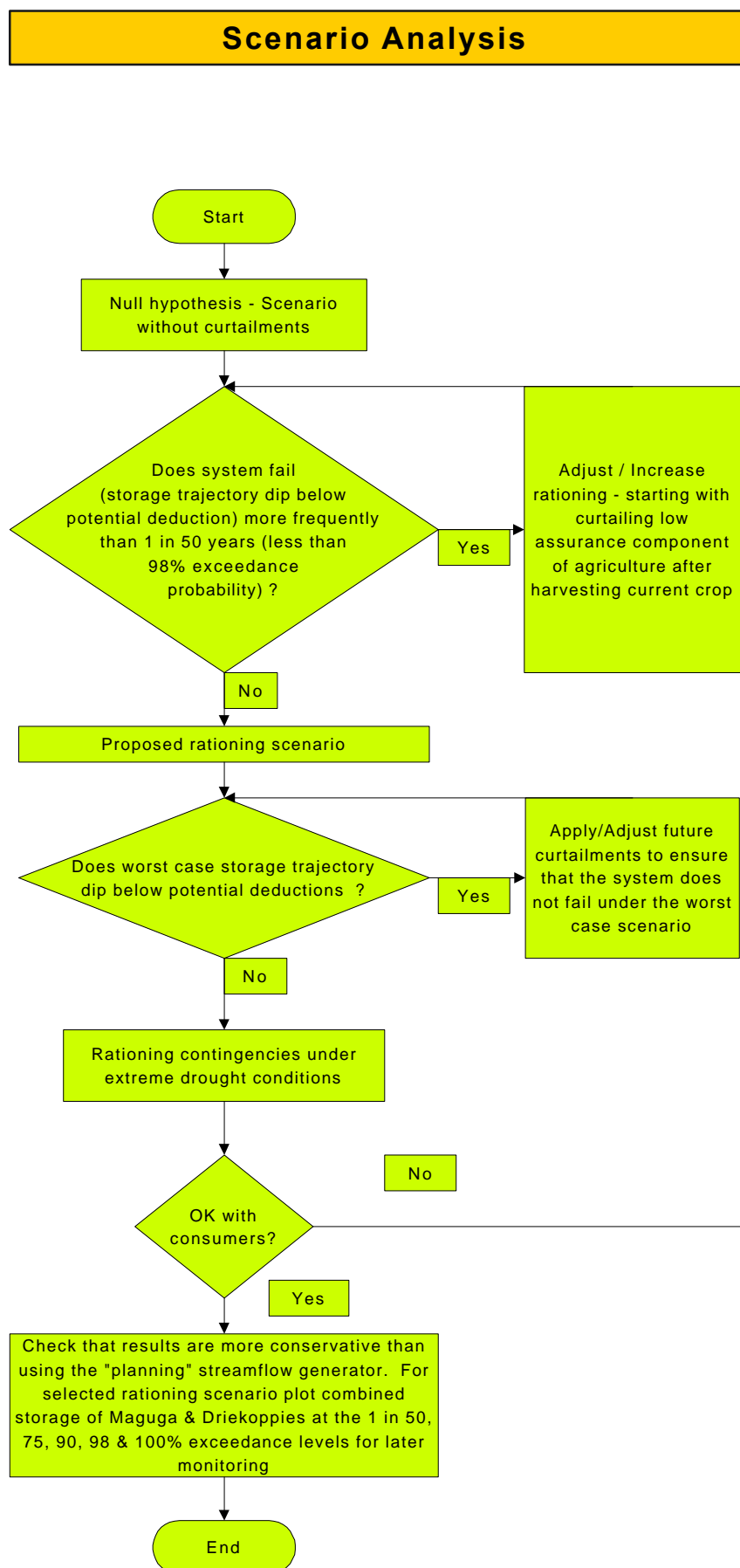
## 4 THE RATIONING PROCESS

Rationing is an integral part of water supply management and is an action that must be planned in advance rather than undertaken in panic. The users should be aware what their long term risk of rationing is when the scheme is first designed. After construction the dam storages and inflows must be monitored regularly and used to evaluate the current risk of rationing. If a system of progressive rationing is adopted then sufficient rationing to survive say a 1 in 50 year event may be imposed initially but sufficient scope must exist for increasing the restrictions at a later stage to survive a more extreme event.

Figure 4.1 and Figure 4.2 illustrate the process in more detail.



**Figure 4.1 : Overview of the rationing process**



**Figure 4.2 : Scenario analysis**

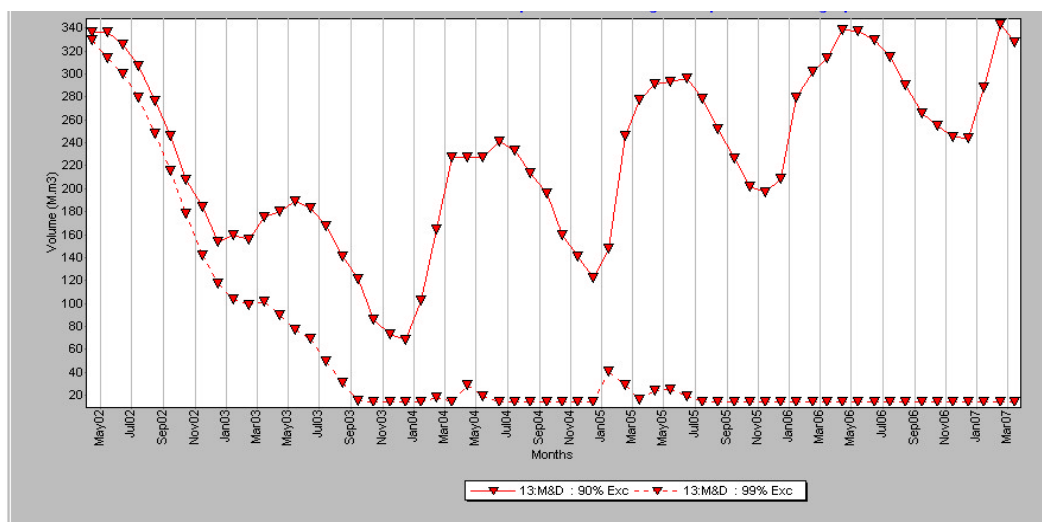
## 5 DEFINING SCENARIOS

Once the Rationing Model has been configured the following information is required to make a routine assessment of the need for restrictions:

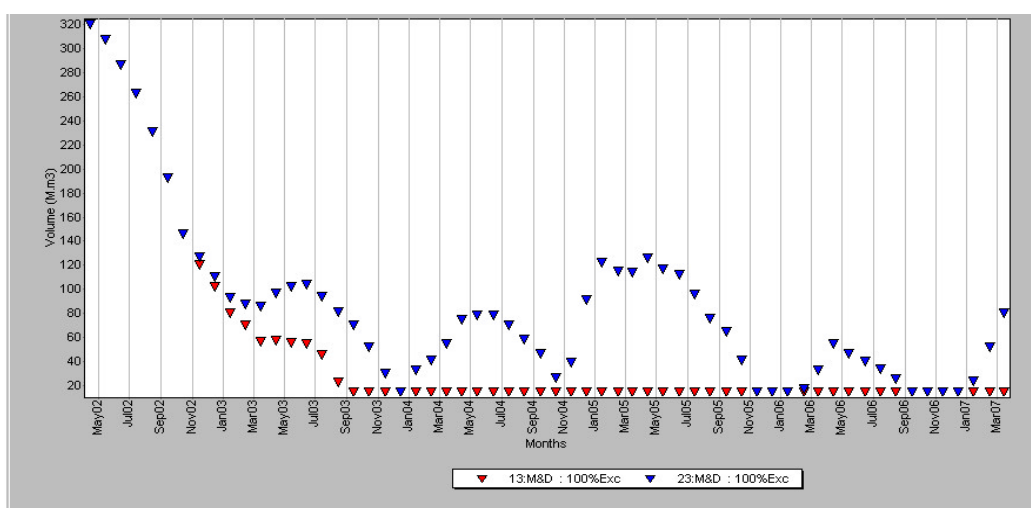
- **storage levels for all the dams in the system.** The storage of the in-channel weirs (approximately 25 million m<sup>3</sup> on the main stem of the Komati River - ignoring the Lomati) was ignored because these must be kept full to enable releases to spill over the weirs and to reach the downstream consumers.
- **antecedent flows.** The flows at key sites for the four preceding months. In the Komati there is little development in the reach between Hoogenoeg and Maguga and the incremental flow obtained by subtracting the streamflows at these two gauges was factored inside the Rationing Model to estimate incremental antecedent natural streamflows for the entire catchment. Additional gauges can be used to improve this estimate.
- **approved demands.** The currently applicable demands are updated.

## 6 GRAPHICAL OUTPUT

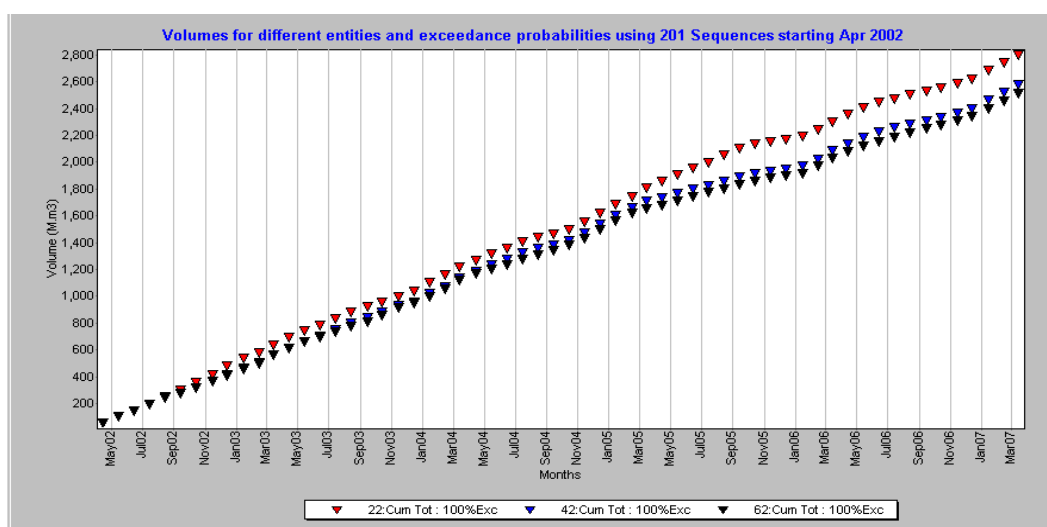
Once a scenario has been defined, the Water Resource Yield Model (WRYM) is run to simulate the streamflows, storage and supply in the system. The Rationing Model reads the Plt file produced by the WRYM to determine the probability of dam storage trajectories and supply. When the dams are fairly full and no curtailments are anticipated a scenario without any curtailments can be analysed (Figure 6.1). Problems may still be experienced if a drought is experienced and a scenario to identify the appropriate timing and extent of rationing restrictions should be analysed (Figure 6.2). It is useful to produce a cumulative plot of the demands in the system to verify that the demands are in fact being supplied and the model is operating correctly (Figure 6.3).



**Figure 6.1 : Storage trajectories assuming no curtailment**



**Figure 6.2 : 30% curtailment of supply to irrigators in Oct 2002 if extreme (worse than a 1 in 100 year drought) conditions are experienced**



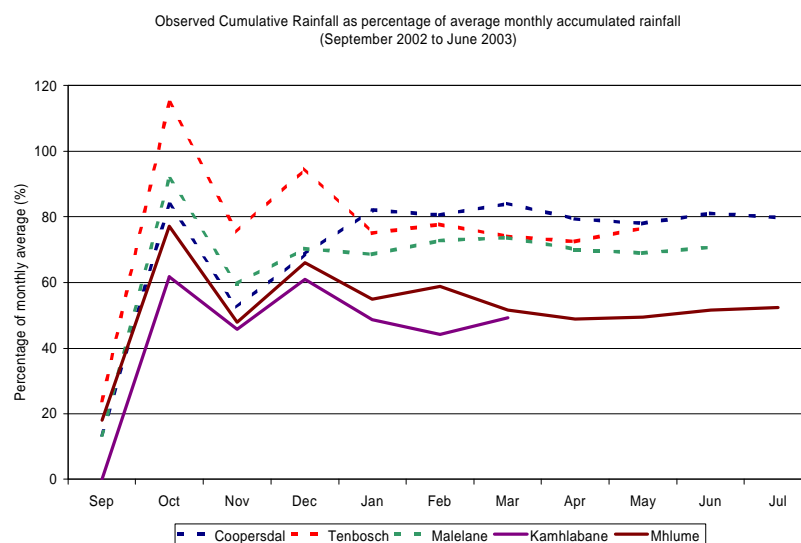
**Figure 6.3 : Cumulative demand plot to verify system demands**

## 7 EXPERIENCES DURING THE CURRENT DROUGHT

### 7.1 Severity of the current drought

#### 7.1.1 Rainfall

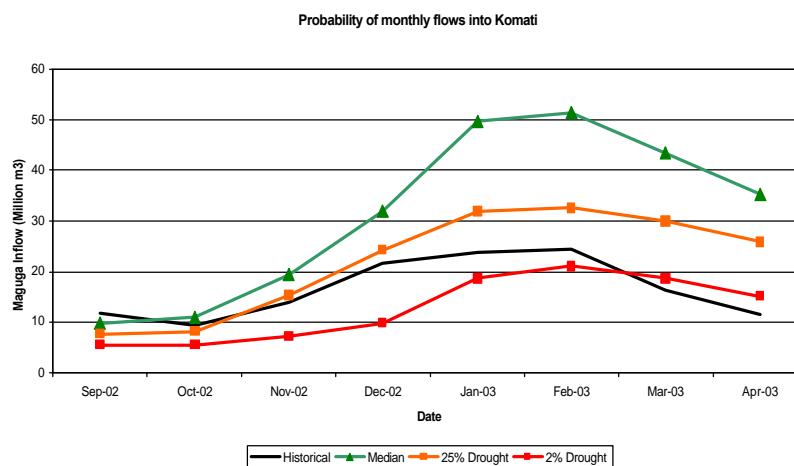
The 2002 / 2003 summer rainfall in the Komati region has been below average. The spatial distribution of the rainfall from thunderstorms is not uniform in the catchment and it is of interest that the rainfall at the stations downstream of the Driekoppies Dam, shown with dashed lines in Figure 7.1, have not decreased by the same proportion as the station upstream of the Driekoppies Dam (Kamhlabane).



**Figure 7.1 : Comparing cumulative rainfall in the Komati Catchment with the long-term average**

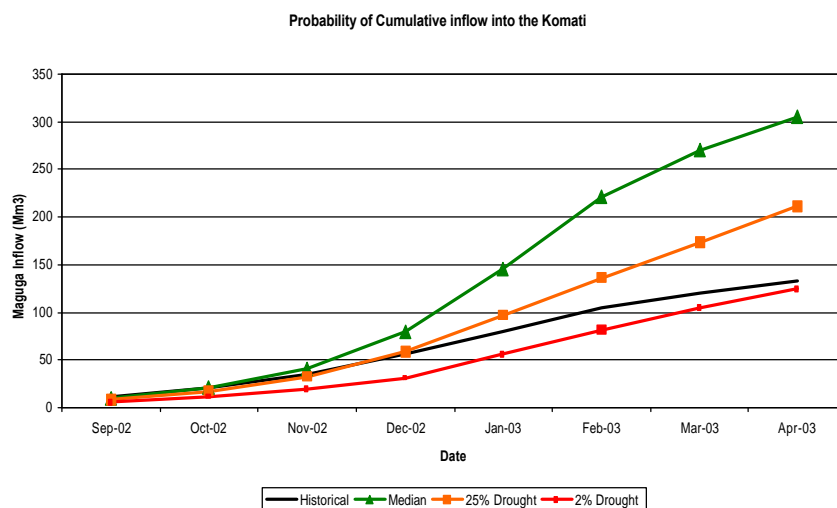
#### 7.1.2 Streamflow

The streamflow entering the Maguga Dam this summer is shown in Figure 7.2 and the flows have decreased from average to below the 1 in 50 year drought by April 2003. The cumulative flow over the season shown in Figure 7.3 also corresponds to the 1 in 50 year streamflow. Similar trends were observed at the Driekoppies Dam.



**Figure 7.2 : Comparing the current historical inflows with the median inflows to Maguga Dam**

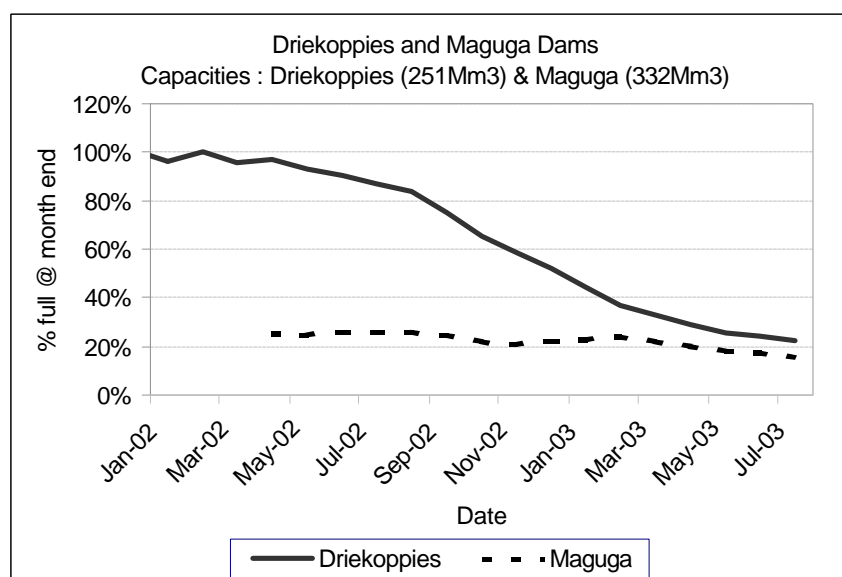




**Figure 7.3 : Comparing the cumulative current historical inflows with the median inflows to Maguga Dam**

### 7.1.3 Storage

The low inflow since September 2002 is reflected in the storage trajectories of the Driekoppies and Maguga Dams. The Driekoppies trajectory kinks downward in September when the increase in irrigation requirements was not matched by an increase in inflows.



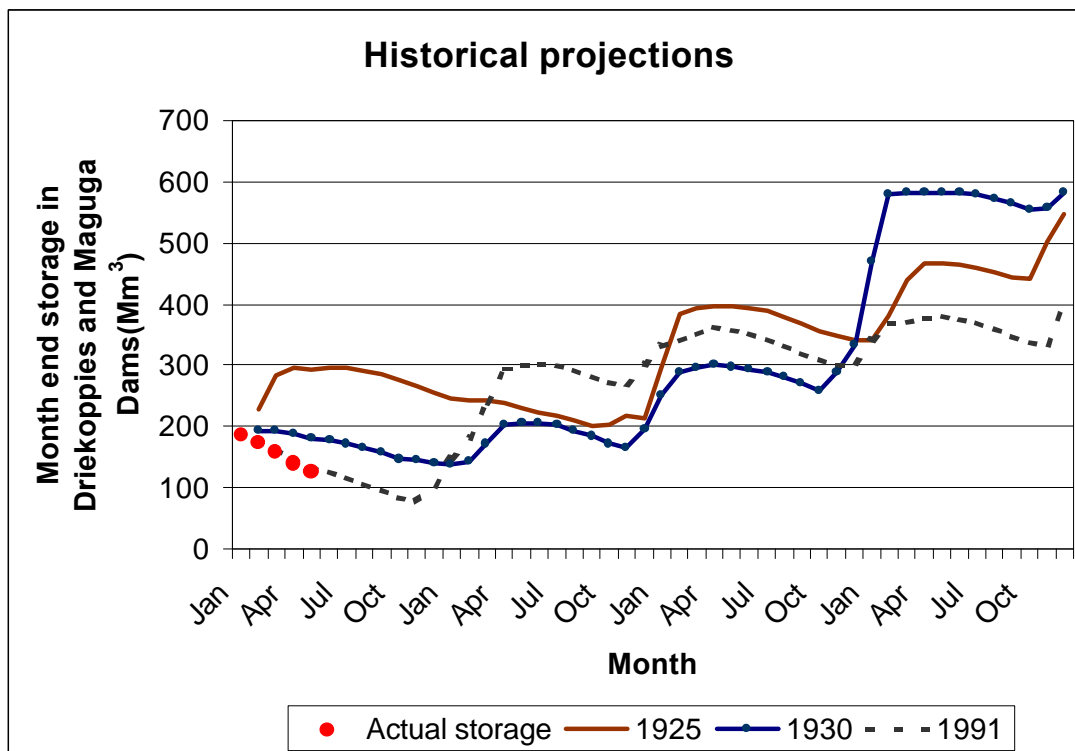
**Figure 7.4 : Percentage full of the Driekoppies and Maguga Dams**

## 7.2 Comparison of the “Forecasting” and “Historical” streamflows

When comparing the current storage with the storages that would be expected under historical or stochastic inflows, the demands on the system for the first four months (February to June) were adjusted so that the releases from Driekoppies Dam equal those released historically for the period February 2003 to June 2003. The following changes were also made to system:

- Transferred liability from Maguga Dam to Driekoppies Dam before simulation (see section 1)
- Filled Maguga Dam “passively” for the rest of the simulation

- Fixed demands supplied from Driekoppies Dam after April 2003 to be only 28% of the allocation.



**Figure 7.5 : Comparing the trajectory of the combined storage in Driekoppies and Maguga Dams since February 2003 under current conditions (red dots) with trajectories assuming the inflows of 1925(black), 1930(red) and 1991(blue)**

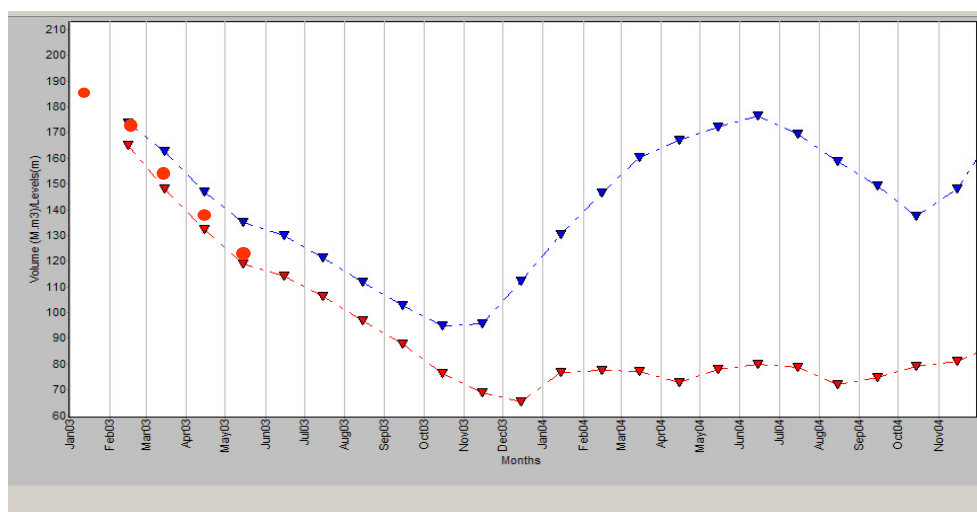
Figure 7.5 shows that the recent dry spell is very similar to that experienced in 1991. In all cases, rationing the releases from Driekoppies Dam to 28% of the allocation did not bring about a dramatic increase in the storage of the system. The uncertainty is in what happens in the second year because it is possible to splice various historical sequences together to obtain a more conservative estimate for the second year. To try to resolve this uncertainty the stochastic forecasting model was developed.

### **7.3 Comparison of the “Planning” and the “Forecasting” streamflow generators**

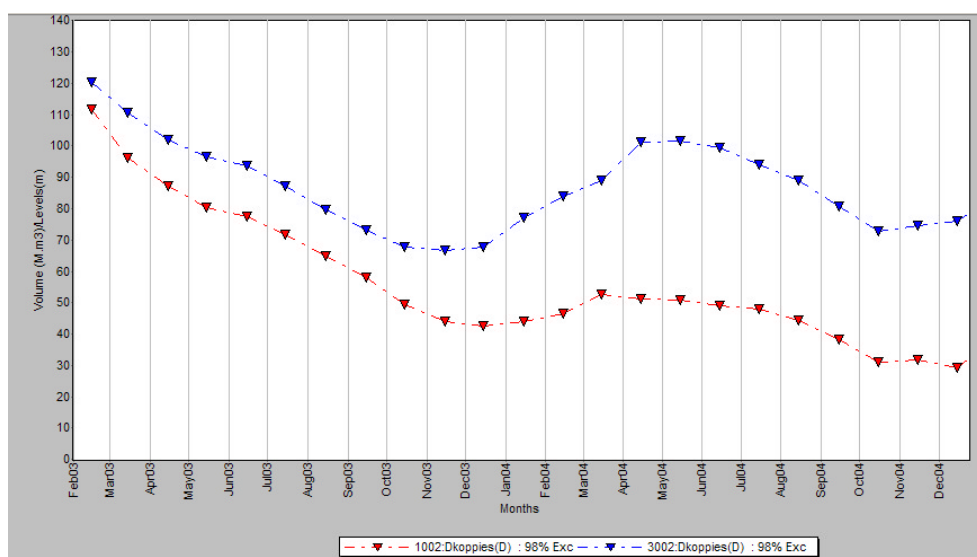
The “forecasting” streamflow generator is influenced by the dry antecedent flows and is more conservative than the “planning” streamflow generator. Figure 3.1 compares the 98% exceedence (1 in 50 year drought) storage trajectories of the combined storage in Driekoppies and Maguga Dams for the “annual” (blue) and the “forecasting” (red) streamflow generators. The 98% exceedence trajectory of the forecasting generator is close to the actual observed storage and seems to agree with the observation in section 7.1 that the drought is currently a 1 in 50 year event.

Figure 7.7 Compares the 1 in 50 year storage trajectories for Driekoppies Dam using the “planning” and the “forecasting” stochastic generators. The “forecasting” generator is about 20 million m<sup>3</sup>/a more conservative than the “planning” generator. The irrigation demands supplied from Driekoppies Dam currently are in the order of 210 million m<sup>3</sup>/a

(see Table 1.1 note 3) so the levels of curtailment required by the two models are within 10% of each other.



**Figure 7.6 : Comparison of the 1 in 50 storage trajectory for the combined Maguga and Driekoppies Dams using the “annual” and the “forecasting” stochastic generator**



**Figure 7.7 : Comparison of the 1 in 50 storage trajectory for the Driekoppies Dam using the “annual” and the “forecasting” stochastic generator**

## 7.4 Effect of the antecedent flows

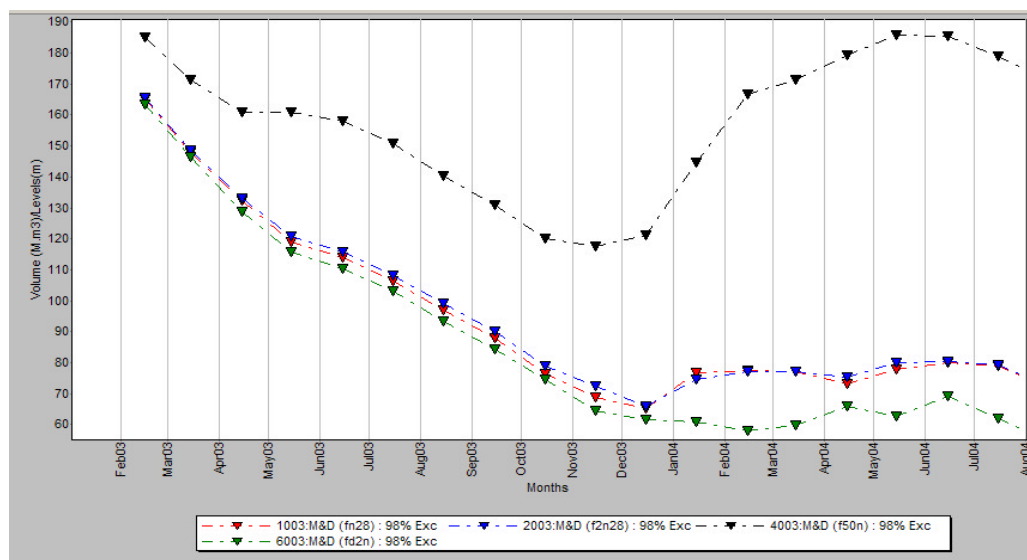
The antecedent flows were varied to determine their effect on the forecast flow (Table 7.1) and the resultant forecasts of the combined storage trajectory for the Driekoppies and Maguga Dams were compared (Figure 7.8). As expected the forecast based on median inflows is significantly higher than that based on drier inflows. The “decreasing” scenario is particularly interesting because although its total inflow is the same as the median inflow, the antecedent streamflows decrease rapidly from a maximum down to the 2% exceedance value. The forecast trajectory for this scenario is actually slightly less than the forecast assuming 2% streamflows for the entire antecedent period. This means that the memory of the higher flows four months earlier does not raise the future storage projection.

**Table 7.1 : Alternative antecedent flows**

Month	Incremental inflow from Hooggenoeg to Maguga			
	Actual	2%	Decreasing	Median
Oct	4	3	31	7
Nov	8	4	16	13
Dec	12	6	13	20
Jan	13	12	12	32
Total	36,7	25,9	71,9	71,8

Notes

- 1) 2% = 1 in 50 year drought
- 2) Decreasing starts with the maximum value in October and decreases to the 2% drought in January



**Figure 7.8 : Forecasts using different antecedent conditions. (2% forecasts: red(actual), 2% (blue), decreasing (green), median (black))**

## 7.5 Progressive curtailment

The Rationing Model was used to apply curtailments progressively. In January 2003 the Rationing Model predicted that there was a 50% chance that the storage in Driekoppies Dam at the end of July 2003 would be almost 180 million m<sup>3</sup>, and there was a 2% risk that the storage would be below 60 million m<sup>3</sup> (see solid and dashed lines in Figure 7.9). At this time the summer rainfall season had not finished so there was still a chance of significant rainfall and this is reflected in the high divergence of these projections. Toward the middle of winter the chance of significant rainfall is much less and the range in projections is correspondingly reduced. For instance, in May the Rationing Model predicted that there was a 50% chance that the storage in Driekoppies Dam would be 60 million m<sup>3</sup> and a 2% risk (98% exceedence probability) that the storage would be below 50 million m<sup>3</sup>. These projections diverge far less than the projection made in January 2003.

At the request of the irrigators the timing of restrictions was delayed till April 2003. Initially (at the end of January), the Rationing Model indicated that supplying 50% (blue number in Figure 7.9) of the demand on the 1<sup>st</sup> April 2003 would reduce the risk of failure to 2% (1 in 50 years). However, because of the low rainfall during the latter part of the summer and the delay in the onset of restrictions, the acceptable supply was reduced to 33% on the 1<sup>st</sup> April 2003. As the drought continues the supply has been further reduced to about 27% from the 1<sup>st</sup> August 2003. This reduction in supply is less than the reductions earlier in the year because the 2% and 50% storage trajectories diverge less in winter.

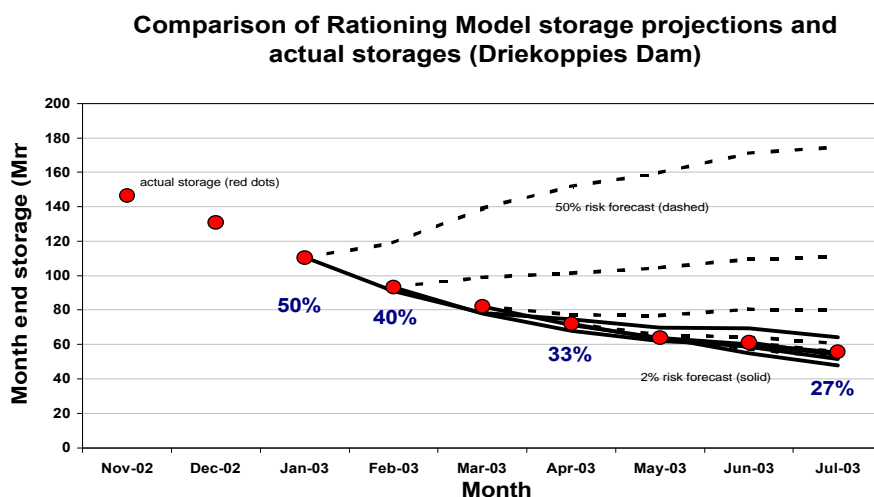


Figure 7.9 : Comparison of the rationing model storage projections with the actual observed storage

## 8 CONCLUSIONS

The Rationing model allows consumers to graphically see the consequences of different curtailment scenarios on the storage trajectories and the reliability of supply from the system.

The “forecasting” stochastic generator provides a good tool for simulating the behaviour of the system for periods just less than two years. It is beneficial if historical drought sequences are also analysed to demonstrate the behaviour of the system under those reference conditions to the consumers. Analyses using the standard “planning” stochastic generator incorporated in the WRYM should also be performed and compared with the Rationing Model especially for periods of over two years.

If the Rationing Model is adopted for other catchments it is recommended that the process of generating the parameters for the catchments be automated. In some catchments it may be necessary to improve the modelling of the cross-correlation and the annual serial correlation and to cope with zero flows.

## 9 ACKNOWLEDGEMENTS

The development of the Rationing Model was only possible because of significant contributions from many individuals and organizations:

- The development of the program was funded by the Komati Basin Water Authority (KOBWA).
- Ben Bonthuys and Charles Sellick developed a customized model to assist with rationing during a drought in the Crocodile-Sabie River. This model incorporated Ben Bonthuys' "bootstrap" stochastic streamflow algorithm whose future flow estimates were conditioned on the existing flows in the system.
- The recent dramatic increase in computer processor speed offered the opportunity to include the Water Resources Yield Model (WRYM) in an interactive tool. The WRYM has been extensively used by the South African

Department of Water Affairs and Forestry (DWAF) to model systems (streamflows, reservoirs, bulk water conveyance, water demands and different operating rules) to facilitate water resources planning. The permission of DWAF to use the WRYM and, in particular, the support of Johan van Rooyen, Malcolm Watson, Bennie Haasbroek and Stephan van Biljon is gratefully acknowledged.

- To enable the WRYM to be used for operational decisions such as rationing, rather than for planning, several enhancements were necessary, including:
  - Incorporating a stochastic algorithm whose forecasts are conditioned on the antecedent flow conditions.
  - Adding the facility to factor (either reduce or grow) a demand after a given number of months from the start of the simulation.

The existing *ANSMK5* code was extended to include an auto-regressive stochastic streamflow forecasting algorithm by Geoff Pegram. This code was incorporated into the WRYM by Pieter van Rooyen of WRP, who also added the facility to factor demands.

- Nadia Nitsche of Ninham Shand, with assistance from André Görgens, Hans Beuster and André Greyling also from Ninham Shand, as well as Ben Bonthuys, of Ben Bonthuys & Associates helped develop the interface of the Rationing program.
- Hans Beuster and Charles Sellick assisted with the testing of the Rationing Model during the current drought.

## 10 REFERENCES

Treaty on the development and utilization of the water resources of the Komati River Basin between the Government of the Kingdom of Swaziland and the Government of the Republic of South Africa. 13 March 1992.

# A MONTHLY FLOW GENERATOR / FORECASTER FOR OPERATIONAL DECISION-MAKING - GEOFFREY GS PEGRAM

A.1	Introduction.....	1
A.2	Model Definition.....	2
A.2.1	Preamble .....	2
A.2.2	Transformation of monthly flows to variates .....	2
A.2.3	Time Series Parameter Estimation.....	3
A.2.4	Estimation of the cross-correlation matrix.....	4
A.2.5	Implementation of the model in software.....	5
A.2.6	Specification of the PARAMFOR.DAT file .....	6
A.2.7	Changes to GNTSTMK5 - becomes GNKOMATI .....	7
A.3	Testing the monthly generator ANSKOM.....	8
A.3.1	GNKOMATI Output.....	8
A.3.2	Short Term Forecasts Compared to Observed.....	13
A.4	Summary and Recommendations.....	16
A.4.1	Suggestions for Future Development.....	16
A.4.2	Recommendations.....	17
A.5	References.....	17

## TABLES

Table A.1	: Cross-correlation matrix R between residuals. Note the blocks of cloned stations highlighted in bold to show near deterministic relationships – the supposed original in upper case. The lowest correlation is between Nooit and Kompoort at 0.380, also in bold	5
Table A.2	: A block from PARAMFOR.DAT .....	6
Table A.3	: Annual and Monthly Means and Standard Deviations of flows generated by ANSKOM and analyzed by GNKOMATI .....	8
Table A.4	: A selection of monthly and annual cross correlations between Nooit on the one hand and Vyge, Magug and Kompoort on the other.....	9

## FIGURES

Figure A.1	: Plot of 900 Magug residuals after normalization and extraction from the AR(4) time series model versus corresponding theoretical residuals with a standard Normal distribution.....	3
Figure A.2	: Estimated and modelled serial correlation functions for the first half of the Maguga record. The models are Autoregressive with lags 2 and 4. Note the better fit with the AR(4) model.....	4
Figure A.3	: Capacity - Yield test for Nooit from MK5.....	10
Figure A.4	: Capacity - Yield test for Nooit from KOMATI .....	10
Figure A.5	: Capacity - Yield test for Vyge from MK5.....	11
Figure A.6	: Capacity - Yield test for Vyge from KOMATI .....	11
Figure A.7	: Capacity - Yield test for Magug from MK5 .....	12
Figure A.8	: Capacity - Yield test for Magug from KOMATI.....	12
Figure A.9	: Monthly forecasts for Nooit starting in month 5 of 1995.....	14
Figure A.10	: Monthly forecasts for Vyge starting in month 5 of 1995.....	14
Figure A.11	: . Monthly forecasts for Magug starting in month 5 of 1995 .....	14
Figure A.12	: Monthly forecasts for Nooit starting in month 9 of 1995.....	15
Figure A.13	: Monthly forecasts for Vyge starting in month 9 of 1995.....	15
Figure A.14	: Monthly forecasts for Magug starting in month 9 of 1995 .....	15

## A.1 Introduction

The Komati Basin Authority (KOBWA) needs a numerical model to provide short-term computer simulation of future possible sequences of monthly flows. These are to be based on currently available values of observed streamflows, at any time in the future, in the Komati system. It was decided to modify the existing annual disaggregation stochastic streamflow generator (ANSMK5), without affecting its usual long-term planning function.

The new hybrid model can thus, by the selection of a parameter IFLAG, produce the conventional annual flows by stochastics or Bootstrapping or, with the preparation of an extra PARAMFOR.DAT file, offer to the calling routine sequences of monthly flows conditioned on the most recently observed ones.

At each call to ANSKOM, an array of monthly flow values for each station selected is offered to the calling routine. This array is for a given year and for a given step in a sequence of 41 (say). The first step in the sequence returns a median (zero noise or deterministic) estimate of the future flows; the remaining 40 steps return plausible future forecasts, each of which includes the right amount of uncertainty and which are of course different from each other. These 40 stochastic forecasts will tend to be more tightly bunched around the median estimate in the near future and will spread out to the natural variability observed in the historical sequences as the forecast horizon increases and the simulated flows “forget” the starting values. It is unlikely that there is detectable skill in forecasting the Komati flows with any routine further than 18 months into the future.

In spite of the relatively short dependence horizon, there is advantage to be gained by using the observed persistence in the streamflow sequences (which depends upon the natural storage in the catchment) to estimate likely values of flow up to (say) 6 months into the future. However, because there is so little correlation in the amount and timing of rainfall from one hydrological year to the next, it is quite unlikely that precise forecasts can be expected from any stochastic model during the wet season and beyond the 12 months ahead. These remarks refer specifically to the streamflow forecasts.

This is not to say that longer sequences of flows (which have short “memory”) cannot be used in reservoir storage based forecasts; the memory of the storage system is enhanced by the large amount of storage in the reservoirs which can be put to good predictive use. The improvement due to the modelling of the persistence in the flows, although marginal, provides a sensible means of getting as much information out of the hydrology as is feasible and sensible.

The brief for the development of the stochastic monthly flow generator/forecaster was firmed up in a letter from Ninham Shand dated 20<sup>th</sup> November, 2001, in which the task was specified as:

- **determine parameters** for the multiple time series
- **use FORTRAN 77** to code the software from an algorithm that would return a set of 12 monthly natural streamflow values for one complete hydrological year
- **test** the reliability of the generated sequences
- **report** on the suitability of the model in the context of the Komati system, with enough detail to enable peer review of the technical background.

This introduction has given an outline of the modelling procedure. The report will continue to give a technical description in Section 2, will describe the implementation



and give a guide to the use of the software in Section 3, will describe the testing procedure and discuss the results in Section 4 and will conclude with a summary and recommendations for further work as requested by Benny Haasbroek of DWAF in his e-mail of 29 November 2001 to Ninham Shand.

## **A.2 Model Definition**

### **A.2.1 Preamble**

There are 16 streamflow sequences in the Komati, but only 6 of these are measured flows, the remainder are “clones”, pro-rated (presumably) by area. The result is that although verification of the parameters of all the sequences is required, validation can be confined to the six key stations. This is done in the sequel. Nevertheless, it was decided to treat all the sequences as if they were genuine and accept them at their face value.

There were 4 straightforward steps in the analysis of the data:

- transformation and de-seasonalization of the natural flows to Gaussian variates
- time series analysis of the individual sequences of variates to extract appropriate sets of time series model parameters
- extraction of the white noise residuals from the individual series
- estimation of the cross-correlations between pairs of sequences of residuals.

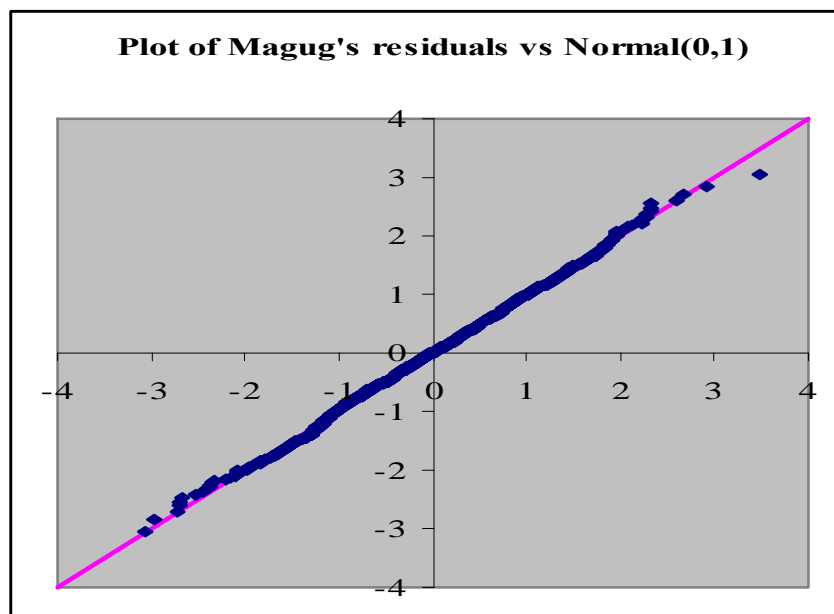
These will be described in detail in turn.

### **A.2.2 Transformation of monthly flows to variates**

The simplest transformation of monthly flows from streams that are perennial (there is not one record of a zero monthly flow in 75 years of data in the Komati system) and which is appropriate in the South African context, is the logarithm. The procedure to produce the variates was to take logs of the flows and then estimate the means and standard deviations of the logs of the flows. These logged flows were then standardized to produce sets of approximately Gaussian variates, each sequence being of length 900 months. It should be noted in passing that no data-screening was done; the data were accepted at face value, even though the “genuine” records were most likely considerably patched and extended.

The results of this transformation was 16 sets of variates of length 900:  $\{z_{it}\}$ ,  $i = 1, 2, \dots, 16$ ,  $t = 1, 2, \dots, 900$ .

An example of the fit to a Gaussian distribution is given in Figure 1.1 where Maguga residuals are ranked and plotted coaxially with the Normal values based on order statistics. The transformation is very acceptable. Worth noting is that, besides the time series analysis of the standardized variates, all the rest of the above analysis (monthly statistics, extraction of variates, cross correlation of residuals) was done on a spread-sheet.

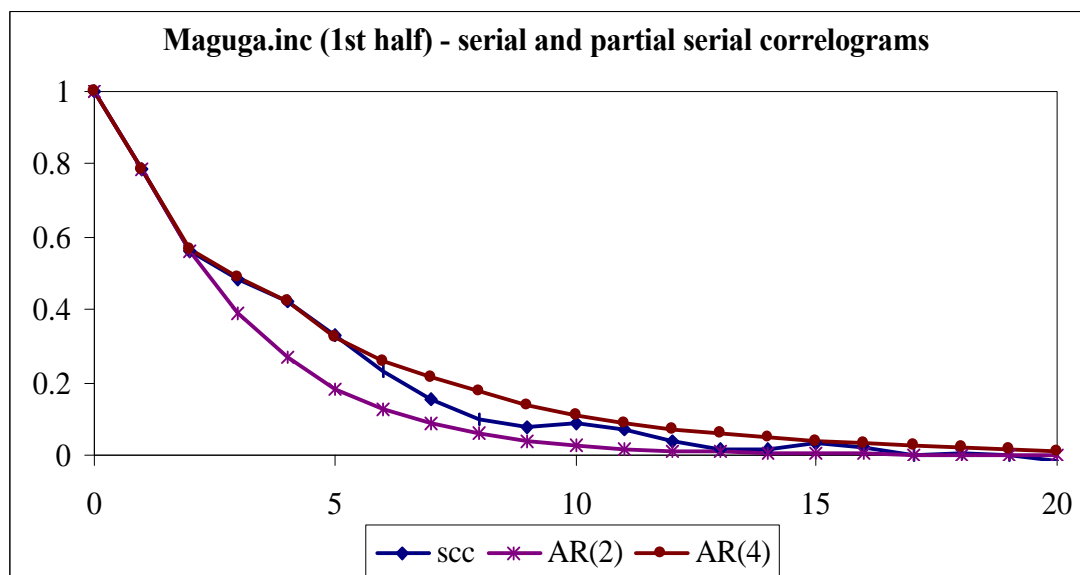


**Figure A.1 : Plot of 900 Magug residuals after normalization and extraction from the AR(4) time series model versus corresponding theoretical residuals with a standard Normal distribution**

### **A.2.3 Time Series Parameter Estimation**

A program to analyse a vector sampled from a (supposed) stationary time series had previously been devised for the Vaal System Analysis. This was modified to allow the modelling of ARMA(p,q) series where  $\max(p+q) = 4$ . The model selection is based on the corrected AIC and the likelihood function includes the determinant term and its approximation. The procedure is outlined in detail in Basson, Allen, Pegram & van Rooyen: *"Probabilistic Management of Water Resources and Hydropower Systems"*, Water Resources Publications, (1994). It was found that in all cases, the best purely Autoregressive model was a lag 4. There were marginally better (and more parsimonious) models suggested such as ARMA (1,2), but it was decided for ease of manipulation to use the AR(4) which depends purely on previous values of observed flows.

An example of the sampled correlogram and the theoretical correlograms suggested by AR(2) and AR(4) models fitted to the Maguga flows is presented in Figure A.2. There it is seen how effectively the AR(4) model follows the sampled correlation structure. The sample was split in two and the same model was found to be appropriate in both halves of the record in the case of Maguga. In the final analysis, the whole record was used to refine the estimation procedure.



**Figure A.2 : Estimated and modelled serial correlation functions for the first half of the Maguga record. The models are Autoregressive with lags 2 and 4. Note the better fit with the AR(4) model**

The four correlation parameters  $f_{i,j}$ ,  $j = 1, 2, 3, 4$  and the standard deviation of the noise term  $s_{ai}$  for each sequence  $i = 1, 2, \dots, 16$ , were extracted. The time series model adopted for each sequence was thus

$$z_{it} = f_{i,1} z_{i,t-1} + f_{i,2} z_{i,t-2} + \dots + f_{i,4} z_{i,t-4} + s_{ai} a_{i,t}$$

where  $a_{i,t}$  are the residuals of the model, or the noise term, assumed to be distributed as  $N(0, s_a^2)$ . These residuals  $\{a_{i,t}\}$  were automatically extracted as part of the model fitting procedure. They were exported and subsequently analysed off-line to extract their cross correlation matrix.

#### A.2.4 Estimation of the cross-correlation matrix

The time series model adopted was one based on a proposal that the cross-correlation between streamflow time series could be lumped into the association between the residuals. Thus the cross-correlation matrix

$R = \{\text{corr}(a_{it}, a_{jt})\}$  was computed from the residuals; it appears in Table A.1

Note the almost deterministic dependence between sequences which make up 6 subsets of the set of 16 sequences - 5 of these are obviously clones of a historic sequence formed by linear scaling. Note also that the monthly cross-correlation between the “genuine” sequences is quite high, from 0.4 to 0.8. It turns out that these values are very similar to the cross-correlation between the annual totals. This point will be commented upon later.

**Table A.1 : Cross-correlation matrix R between residuals. Note the blocks of cloned stations highlighted in bold to show near deterministic relationships – the supposed original in upper case. The lowest correlation is between Nooit and Kompoort at 0.380, also in bold**

	NOOIT	VYGE	Gladde	Boeken	Tee	Hooge	MAGUG	Cdc	Border	Tonga	KOMP	UPLDM	Ngonin	DRIEK	Mhiam	Vlakbult
NOOIT	<b>1.000</b>	0.710	0.691	0.711	0.710	0.710	0.452	0.455	0.459	0.456	<b>0.380</b>	0.514	0.514	0.414	0.409	0.414
VYGE	0.710	<b>1.000</b>	<b>0.974</b>	<b>0.998</b>	<b>0.997</b>	<b>0.994</b>	0.652	0.650	0.643	0.639	0.492	0.728	0.730	0.564	0.559	0.563
Gladde	0.691	<b>0.974</b>	<b>1.000</b>	<b>0.978</b>	<b>0.979</b>	<b>0.982</b>	0.650	0.652	0.637	0.631	0.487	0.722	0.718	0.562	0.561	0.558
Boeken	0.711	<b>0.998</b>	<b>0.978</b>	<b>1.000</b>	<b>0.999</b>	<b>0.998</b>	0.653	0.654	0.645	0.643	0.490	0.730	0.729	0.562	0.566	0.561
Tee	0.710	<b>0.997</b>	<b>0.979</b>	<b>0.999</b>	<b>1.000</b>	<b>0.999</b>	0.653	0.655	0.645	0.643	0.488	0.730	0.727	0.560	0.554	0.560
Hooge	0.710	<b>0.994</b>	<b>0.982</b>	<b>0.998</b>	<b>0.999</b>	<b>1.000</b>	0.654	0.656	0.647	0.645	0.490	0.729	0.725	0.562	0.555	0.561
MAGUG	0.452	0.652	0.650	0.653	0.653	0.654	<b>1.000</b>	<b>0.994</b>	<b>0.975</b>	<b>0.971</b>	0.791	0.887	0.882	0.838	0.831	0.833
Cdc	0.455	0.650	0.652	0.654	0.655	0.656	<b>0.994</b>	<b>1.000</b>	<b>0.990</b>	<b>0.987</b>	0.800	0.892	0.883	0.840	0.828	0.841
Border	0.459	0.643	0.637	0.645	0.645	0.647	<b>0.975</b>	<b>0.990</b>	<b>1.000</b>	<b>0.998</b>	0.812	0.888	0.881	0.839	0.823	0.851
Tonga	0.456	0.639	0.631	0.643	0.643	0.645	<b>0.971</b>	<b>0.987</b>	<b>0.998</b>	<b>1.000</b>	0.809	0.885	0.876	0.836	0.818	0.850
KOMP	<b>0.380</b>	0.492	0.487	0.490	0.488	0.490	0.791	0.800	0.812	0.809	<b>1.000</b>	0.713	0.713	0.814	0.799	0.833
UPLDM	0.514	0.728	0.722	0.730	0.730	0.729	0.887	0.892	0.888	0.885	0.713	<b>1.000</b>	<b>0.995</b>	0.808	0.798	0.810
Ngonin	0.514	0.730	0.722	0.729	0.727	0.725	0.882	0.883	0.881	0.876	0.713	<b>0.995</b>	<b>1.000</b>	0.811	0.803	0.811
DRIEK	0.414	0.564	0.562	0.562	0.560	0.562	0.838	0.840	0.839	0.836	0.814	0.808	0.811	<b>1.000</b>	<b>0.996</b>	<b>0.989</b>
Mhiam	0.409	0.559	0.561	0.556	0.554	0.555	0.831	0.828	0.823	0.818	0.799	0.798	0.803	<b>0.996</b>	<b>1.000</b>	<b>0.977</b>
vlakbult	0.414	0.563	0.558	0.561	0.560	0.561	0.833	0.841	0.851	0.850	0.833	0.810	0.811	<b>0.989</b>	<b>0.977</b>	<b>1.000</b>

To use this R-matrix in generating stochastic sequences with the right intercorrelation, it is necessary to correlate the generated residuals  $\{a_{it}\}$ . To achieve this, the cross-correlation matrix of the residuals, R is decomposed into its “square-root” matrix (not unique) using singular value decomposition. Thus R can be written  $UWV^T$  and  $B = R^{1/2}$  can be written as  $UW^{1/2}V^T$  which is symmetrical. Because  $V^T = U^{-1}$ , then  $UW^{1/2}V^T U W^{1/2}V^T = U W V^T = R$ . This (symmetrical) version of the square root was suggested by Todini (personal communication, 1996) because it “smooths” the independent noise vector  $\{e_{it}\}$  when used to produce  $\{a_{it}\}$  in the usual way:

$$\{a_{it}\} = B\{e_{it}\}$$

For further details of this procedure, see (for example) Basson et al (1994).

### A.2.5 Implementation of the model in software

After much thought and examining the alternatives such as:

- including monthly parameters in the annual parameter file or have them in a separate file or perhaps include them in the software
- perform the calculation in a separate subroutine called by the annual flow simulation generator ANSMK5 (derived from ANSIM) or do the calculation inside ANSMK5.

It was decided to do the following.

Perform the calculation inside ANSMK5, but call it by a new name: ANSKOM. This change is to stress that this routine is a special one-off version of the forecaster designed for the Komati system. The reason for doing the computation inside the routine was that if ANSMK5 was to call a subroutine, the code would be changed by one or more lines, so it seemed sensible to do the simplest thing and modify the code a bit more. From the outside, it still looks the same (except for the name, which could be kept the same as the old one if desired) and produces the same format of output, except that the output consists only of monthly streamflows without rainfall, (This can be remedied if one could include average monthly rainfalls at the stations in the PARAMFOR.DAT file - these could be reported as the best future estimate of rainfall; not done now.)

There are two other changes to ANSMK5 to convert it to ANSKOM.

First, in calling the routine, a parameter labelled IFLAG has previously been given values of 0 or 1, depending on whether a full stochastic generation of annual flows is

desired or whether the generation is done by Bootstrap. A new option is to set IFLAG to 3. This choice will launch the monthly forecasting option.

Second, the parameters needed for the monthly streamflow generation are stored in a new file dedicated to monthly forecasting called PARAMFOR.DAT. The existing PARAMK5.DAT which sets up the annual flow generation and orders the data sequences is unaltered, but still necessary. PARAMFOR.DAT is an optional extra, used only in forecasting when IFLAG is set to 3 by the calling program.

Another telling reason for using a separate parameter file for the forecasting model is that the structure of ANSKOM is complicated. It uses a lot of named COMMON blocks to pass data and parameters to subroutines. The file PARAMK5.DAT is read in a subroutine called ANREAD and the file is opened as 25 and then closed. Subsequently 25 is used by a subroutine READ in servicing DISAGG for the AFF, IRR etc files. There is thus no hope of using the 25 to continue to read the PARAMK5.DAT file in ANSKOM. The alternative is to recraft the routine ANREAD to read the monthly data if IFLAG is set to 3. New named COMMON blocks would have to be set up and linked through the appropriate calling routines. It was felt that this might be neat programming in the future but not good for a one-off prototype which has to be made to work on a tight schedule.

#### To summarize:

To use ANSKOM in **conventional** (annual disaggregation or bootstrap) mode, nothing needs to be changed. Call it with IFLAG = 0 or 1 and supply a PARAMK5.DAT file.

To use ANSKOM in monthly **forecasting** mode, call it with IFLAG = 3, but ensure that besides the PARAMK5.DAT file, there is a PARAMFOR.DAT file containing the monthly parameter information.

#### A.2.6 Specification of the PARAMFOR.DAT file

The PARAMFOR.DAT file has the following structure for use in the Komati system. Note that it is read in free format, thus all that is required is that the data appear in order, separated by spaces in each record. There are 16 blocks of information for the 16 stations, named in the same order as in PARAMK5.DAT. This set of blocks is followed by the 16-square B matrix for cross-correlating the monthly residuals. This is not the same as the B matrix in PARAMK5.DAT. Each of the 16 blocks consists of 7 rows. An example follows in Table A.2.

**Table A.2 : A block from PARAMFOR.DAT**

magug												
0.990486	-0.383124	0.290633	-0.108017	0.57617								
2.0589	2.5005	2.9283	3.2530	3.3835	3.2619	2.9544	2.6451	2.4211	2.2368	2.0827	1.9810	
0.4062	0.5370	0.5942	0.6592	0.8093	0.7621	0.5272	0.3516	0.3080	0.2992	0.2866	0.3229	
1994	6.59	9.78	15.72	22.22	18.15	14.41	15.25	13.25	10.70	8.40	6.95	5.94
1995	5.23	10.69	28.54	40.26	-108.52	116.93	38.95	20.65	17.76	13.93	11.38	9.52

- Row 1 is a blank line.
- Row 2 contains the station name (not read)
- Row 3 contains 5 numbers: the 4 AR parameters  $f_i$  followed by the residuals' standard deviation.
- Row 4 contains the 12 means of the logs of the flows
- Row 5 contains the 12 standard deviations of the logs of the flows

- Row 6 contains 13 numbers which are copied from a streamflow file. They must all be positive. The first number is the year (not read), the next 12 are the natural observed flows.
- Row 7 is special. It is identical in format to row 6 except that at least one of the 12 flows must be preceded by a negative sign. In this example, the 5<sup>th</sup> month is flagged negative. This indicates that the flow for that month (and the following months) is “unobserved”. The first negative encountered in *any* of the rows 7 belonging to *any* of the 16 blocks decides the month from which the forecasting must start.

**Note:**

- For testing purposes (with historical flows) flagging an appropriate month with a negative sign in just one of the sequences in row 7 defines the first “unknown” month.
- In operational mode, the monthly flow totals will be available sometime in the following month in real time, so can be inserted by editing the PARAMFOR.DAT file. When the year is full (i.e. September has been recorded and entered) replace the 6<sup>th</sup> row with the 7<sup>th</sup> and replace the 7<sup>th</sup> row with 13 numbers, the year (e.g. 2002) and 12 negative numbers.

**In summary:**

Rows 6 and 7 in each of the 16 blocks in PARAMFOR.DAT can be edited to reflect the currently known flows. All flows in row 6 must be positive; the first negative flow in any of the stations’ row 7 defines the first “unknown” month from which forecasting will be done.

ANSKOM returns the filled set of rows 7 with its forecast flows as the first year for simulation/forecasting.

There is no limit to the number of years that can be generated, but it must be noted that this is a monthly forecasting tool, not a planning tool. It is not designed to preserve annual cross-correlations explicitly. The forecast horizon should be limited to 2 or 3 years - 5 years at most.

**A.2.7 Changes to GNTSTMK5 - becomes GNKOMATI**

In GNKOMATI there are some minor changes of editorial nature and two major ones. All changes are flagged with !DEC01 in column 73+ in the code.

The first major change is the inclusion of the IFLAG = 3 option in the request for user input.

The second major change is a section, only activated in case IFLAG=3, which dumps all the 41 sequences of generated monthly flows to a file called ‘FORECAST.TXT’. These flows appear in blocks belonging to the streamflow stations requested in the input dialogue. The contents of the file can be easily read by a spreadsheet for analysis and plotting. Some examples of this output will be given in the sequel.

### A.3 Testing the monthly generator ANSKOM

The output from ANSKOM can be tested in two ways: firstly by GNKOMATI as long term sequences and secondly by analysing short term sequences output in 'FORECAST.TXT'.

#### A.3.1 GNKOMATI Output

To examine the long term behavior of the monthly forecast generator, 75 years of synthetic data were generated, conditioned on the observed values of monthly flows up to the end of the 1993 hydrological year. The flows rapidly become independent of the starting condition, so the tests in GNKOMATI, are valid. Examination of the output of GNKOMATI shows that there are some verification tests which are the comparisons of annual and monthly statistics and then the validation tests, the most telling of which is the set of storage-yield tests.

**Table A.3 : Annual and Monthly Means and Standard Deviations of flows generated by ANSKOM and analyzed by GNKOMATI**

	NOOIT		VYGE		MAGUGA		KOMPOORT		UPLUM		DRIEPOORT	
	hist	syn	hist	Syn	hist	syn	hist	syn	hist	syn	hist	syn
ann mean	59.9	52.5	182.3	178.1	235.3	231.2	32.5	27.8	21.5	21.4	58.9	56.5
stdev	55.1	41.5	91.4	95.4	139.6	105.9	45.3	26.8	11	9.8	54.1	37.7
oct mean	2.2	1.7	5.8	5.7	8.6	8.6	0.7	0.7	0.8	0.8	1.7	1.7
stdev	5.7	1.7	4.5	3.2	4.1	3.7	0.7	0.7	0.4	0.3	0.9	0.8
nov mean	7.4	5.8	13.5	12.6	14	14.2	2	2	1.4	1.4	2.9	2.9
stdev	14.3	7.3	13.2	9.8	7.7	7.7	2.5	2.5	0.9	0.9	1.9	1.8
dec mean	10.5	9.4	23.5	22.7	22.2	22.8	3.7	3.6	2.3	2.3	5.5	5.5
stdev	15.7	10.5	19.6	17.9	13.5	13.9	5.7	4.5	1.4	1.4	5.5	4.5
jan mean	12.1	10.7	34.3	33.8	33.3	32.8	6.7	6.1	3.2	3.2	9.2	8.8
stdev	17	11.6	28.5	27.4	30.8	24	12.6	7.8	2.4	2.1	10.3	8.6
feb mean	12	9.6	36.2	34.5	43.3	41.6	7.8	6	3.9	3.9	12.1	11.6
stdev	21.1	11.8	38.8	32	45.7	37.4	19.3	10	3.6	3.2	19.2	13.3
Mar mean	6.5	5.6	24.3	23.6	37.8	34.8	6	4.3	3.4	3.2	10.7	9.7
stdev	11.2	6.9	27.4	20	45.8	29.5	18.3	6.1	3.3	2.5	17.3	10.5
Apr mean	3.3	3.2	14.7	14.7	23.5	22.1	2.5	2	2.1	2.1	6.1	5.4
stdev	3.6	2.5	10.4	8	28.2	12.2	6.4	2.1	2	1.1	10.4	4.2
may mean	2.2	2.1	10	9.8	15	14.9	1	1	1.3	1.3	3.3	3.3
stdev	3.2	1.6	6.6	4.3	6.1	5.3	1	0.7	0.5	0.4	2.3	1.8
jun mean	1.2	1.2	6.7	6.8	11.8	11.9	0.7	0.7	1	1	2.4	2.4
stdev	0.7	0.7	2.3	2.5	3.4	3.6	0.4	0.5	0.3	0.3	0.9	1.1
jul mean	0.9	0.9	5	5	9.8	9.9	0.5	0.6	0.8	0.8	1.9	2
stdev	0.5	0.5	1.8	1.8	2.9	3	0.3	0.4	0.2	0.2	0.7	0.8
aug mean	0.7	0.7	4.2	4.2	8.4	8.5	0.4	0.4	0.7	0.7	1.6	1.7
stdev	0.5	0.4	1.6	1.5	2.5	2.5	0.2	0.3	0.2	0.2	0.5	0.6
sep mean	0.9	0.8	3.9	3.9	7.7	7.8	0.5	0.4	0.6	0.6	1.5	1.5
stdev	1.6	0.5	1.7	1.5	2.7	2.6	0.5	0.4	0.2	0.2	0.6	0.6

The annual and monthly means and standard deviations are reasonably well modelled by the monthly model as will be seen in Table A.3. The means are better reflected than the standard deviations which tend to be on the low side.

A subset of the output from GNKOMATI as compared with the output from GNTSTMK5 (which works on the annual disaggregation model) is presented in Table A.4. The annual cross-correlations are (with some exceptions) not as well modelled by the monthly model as by the annual model as shown in Table A.4. This could be rectified by modifying the cross correlation estimation procedure, to a more complicated one than that adopted here as the first attempt. The measured monthly cross-correlations of the residuals taken from Table A.1 and used in the generation of the flows appear in the penultimate row of Table A.4.

**Table A.4 : A selection of monthly and annual cross correlations between Nooit on the one hand and Vyge, Magug and Kompoort on the other**

	nooit/vyge			nooit/magug			nooit/kompoort		
	hist	komati	mk5	hist	komati	mk5	hist	komati	mk5
Oct	0.910	0.675	0.743	0.660	0.446	0.518	0.573	0.327	0.377
Nov	0.871	0.667	0.735	0.541	0.417	0.445	0.464	0.330	0.347
Dec	0.818	0.664	0.767	0.522	0.436	0.470	0.356	0.353	0.365
Jan	0.714	0.669	0.716	0.331	0.434	0.316	0.126	0.313	0.138
Feb	0.580	0.681	0.593	0.341	0.438	0.371	0.229	0.326	0.286
Mar	0.613	0.676	0.563	0.414	0.454	0.367	0.331	0.338	0.300
Apr	0.778	0.687	0.639	0.443	0.445	0.356	0.377	0.312	0.353
May	0.866	0.678	0.687	0.467	0.426	0.337	0.313	0.326	0.197
Jun	0.904	0.674	0.675	0.590	0.415	0.411	0.386	0.326	0.278
Jul	0.917	0.701	0.650	0.706	0.426	0.481	0.524	0.320	0.325
Aug	0.925	0.700	0.620	0.732	0.441	0.475	0.474	0.343	0.293
Sep	0.923	0.683	0.696	0.683	0.442	0.541	0.488	0.358	0.380
$R_{ij}$		0.710			0.452			0.380	
Ann	0.774	0.666	0.751	0.603	0.439	0.576	0.465	0.324	0.446
		Note							
		-	'hist' are measured from the record						
		-	'komati' are estimated from the output from ANSKOM						
		-	'mk5' are estimated from the output from ANSMK5						

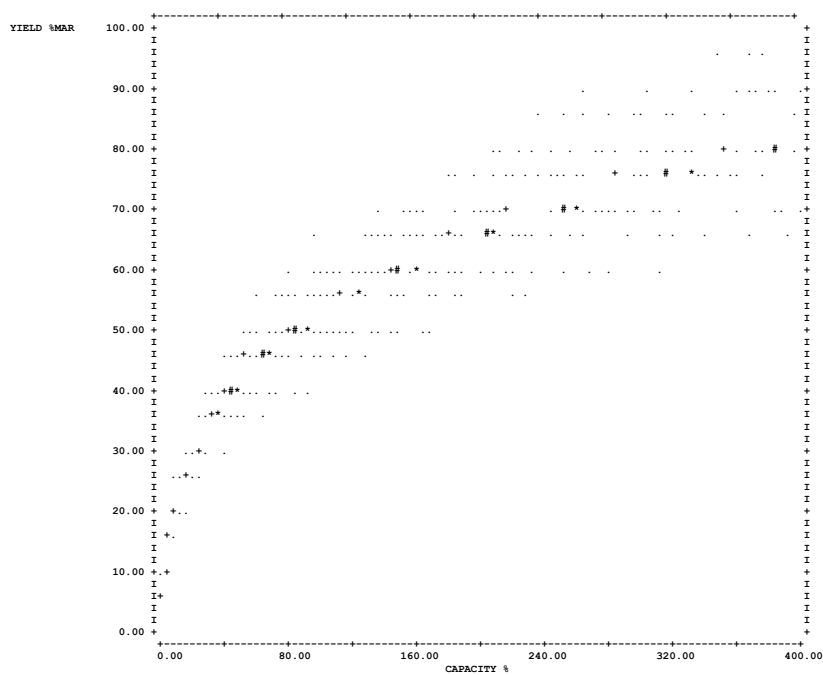
The monthly cross-correlations of the monthly model reflect the historical ones reasonably during the wet months but under-estimate them during the recession limbs of the dry months, which is to be expected, and which is not a matter of concern because the flows are low. By and large the monthly correlations from ANSKOM are relatively uniform over the year as expected and the annual value seems to be close to the month values. This set is one of the worst comparisons between the two modelling procedures. The ANS output from GNKOMATI contains the full comparison.

The capacity-yield tests are satisfactory, in that the outputs from ANSMK5 and ANSKOM are not apparently very different from each other. These test results follow in Figure A.3 vs Figure A.4, Figure A.5 vs Figure A.6 and Figure A.7 vs Figure A.8 where it will be seen that the monthly model does well to recapture the historical curve.

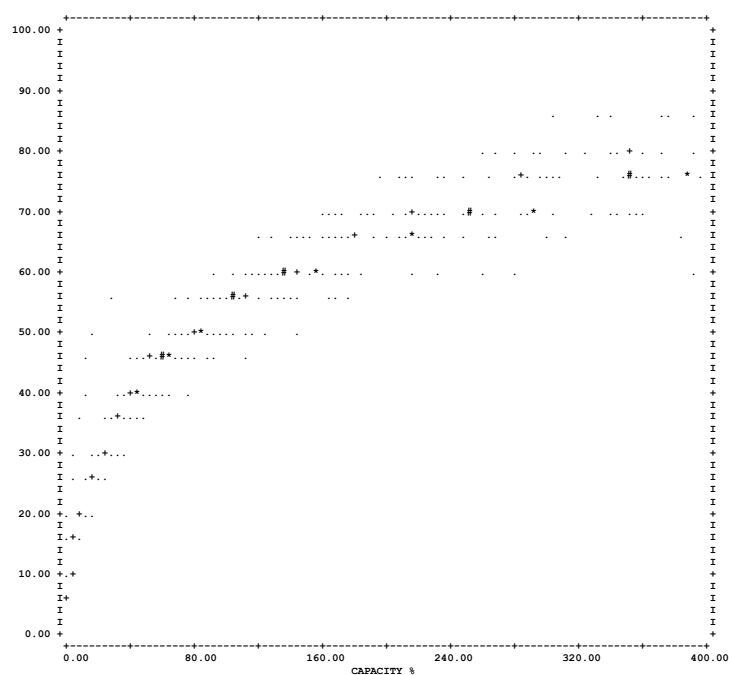
The only apparent difference of importance is the spread of points in the Nooit curves (Figure A.3 vs Figure A.4) - the average is right but the spread for the KOMATI model is less than for the MK5 model, possibly due to the choice of marginal distribution.

The consequence of these tests is that it is safe to use the monthly model in short terms forecasting because it intrinsically embodies the long term variability. The exception is that it does not recover the annual cross-correlation as well as one would like for a long term planning model. This would have to be rectified by designing a hybrid annual/monthly model, which can be done - the theory has already been developed.

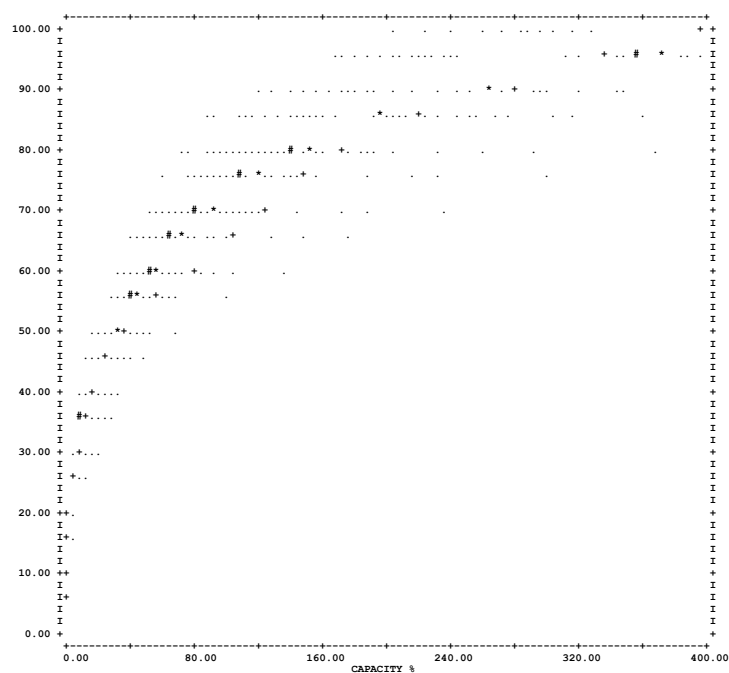




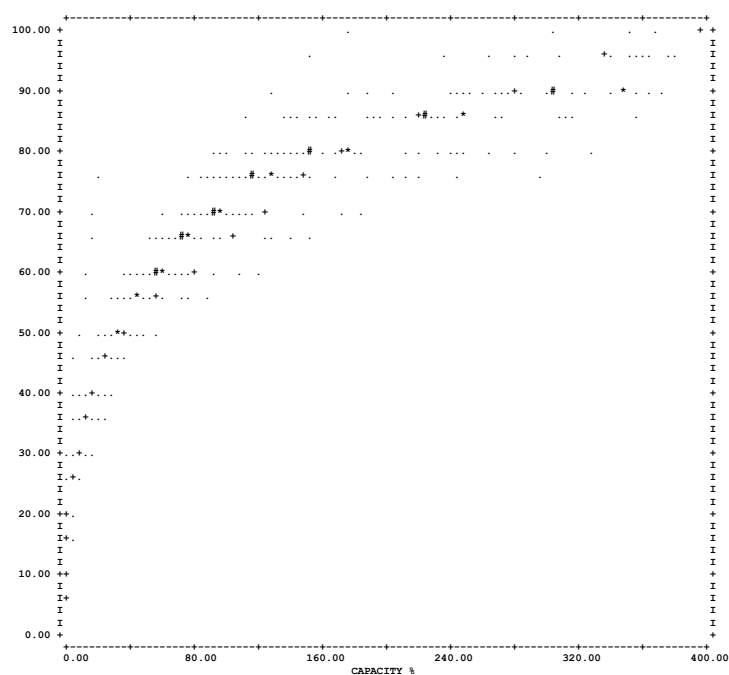
**Figure A.3 : Capacity - Yield test for Nooit from MK5**



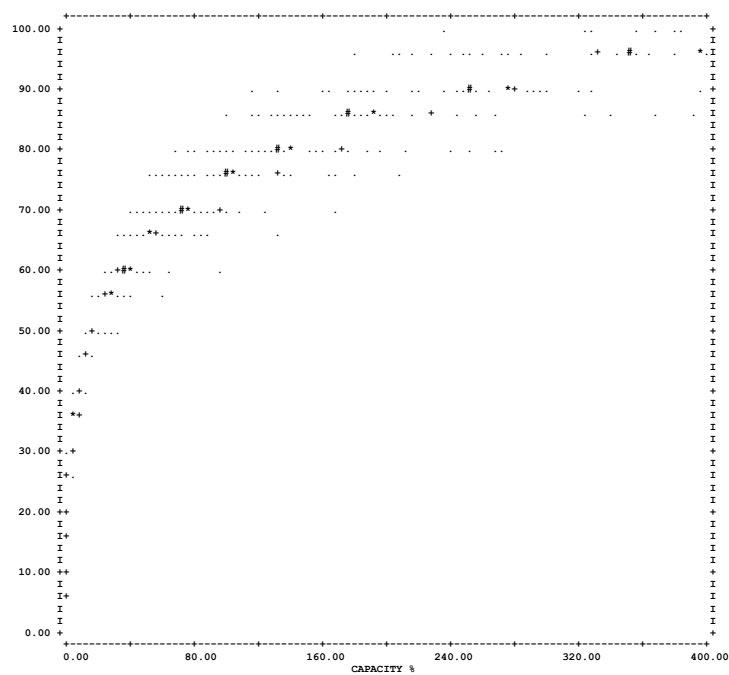
**Figure A.4 : Capacity - Yield test for Nooit from KOMATI**



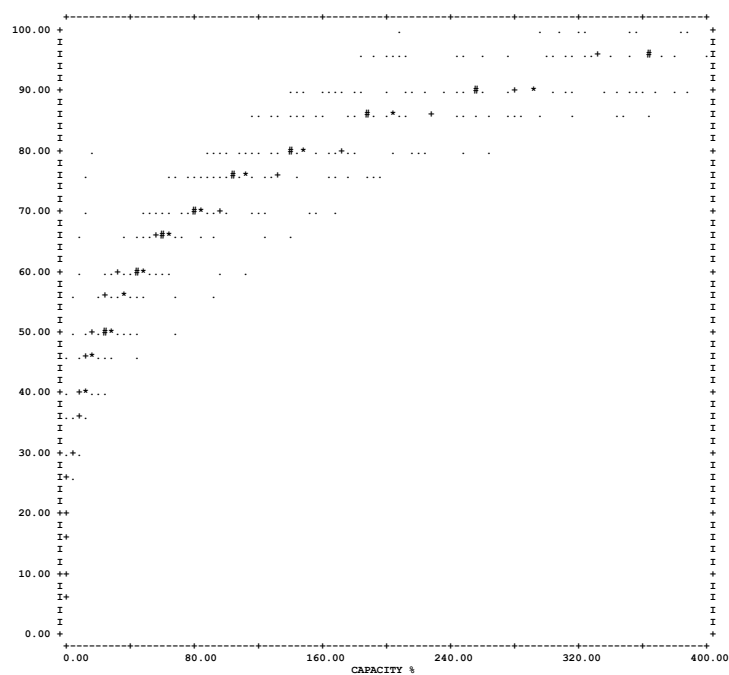
**Figure A.5 : Capacity - Yield test for Vyge from MK5**



**Figure A.6 : Capacity - Yield test for Vyge from KOMATI**



**Figure A.7 : Capacity - Yield test for Magug from MK5**



**Figure A.8 : Capacity - Yield test for Magug from KOMATI**

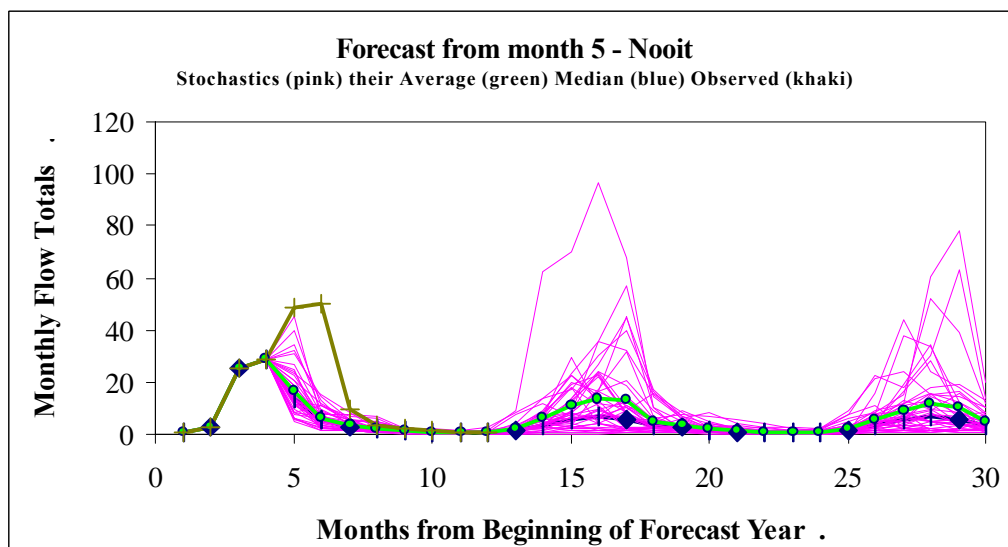
### A.3.2 Short Term Forecasts Compared to Observed

The output from ANSKOM was dumped in FORECAST.TXT and then imported to a spreadsheet. Two sets are presented here. The first is for forecasts from the 5<sup>th</sup> month of the 1995 hydrological year; the second set is from the 9<sup>th</sup> month.

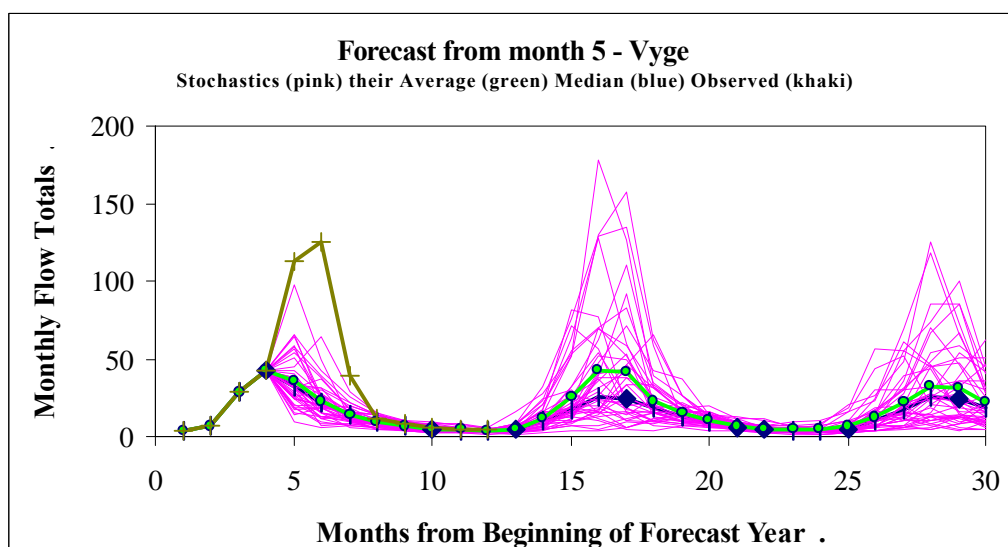
The first set is forecast from the wettest month (February) and the second set from a month when very little rain falls (June). Six figures (Figure A.9 to Figure A.14) are presented, three (for stations Nooit, Vyge and Magug) for each set.

In each figure, the 40 stochastic forecasts appear as pink traces, their average is green, the median (zero noise trace) is in navy blue and the observed data appear in khaki but is only 12 months long, being the last year in the available record - 1995. The median is slightly below the average in the later months because the distributions of the monthly flows are skew. The pink stochastic traces show the variability of the forecasts increasing with the forecast horizon. The wet month forecasts are much more variable than the dry month ones, as would be expected.

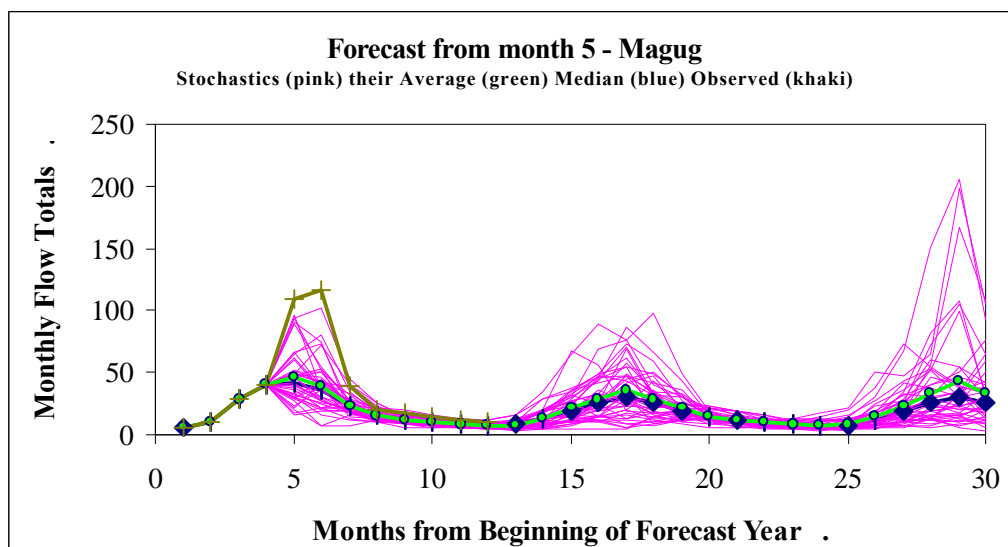
The figures give credibility to the monthly forecasting algorithm's ability to provide as much information about the future as is intrinsic to the data sets.



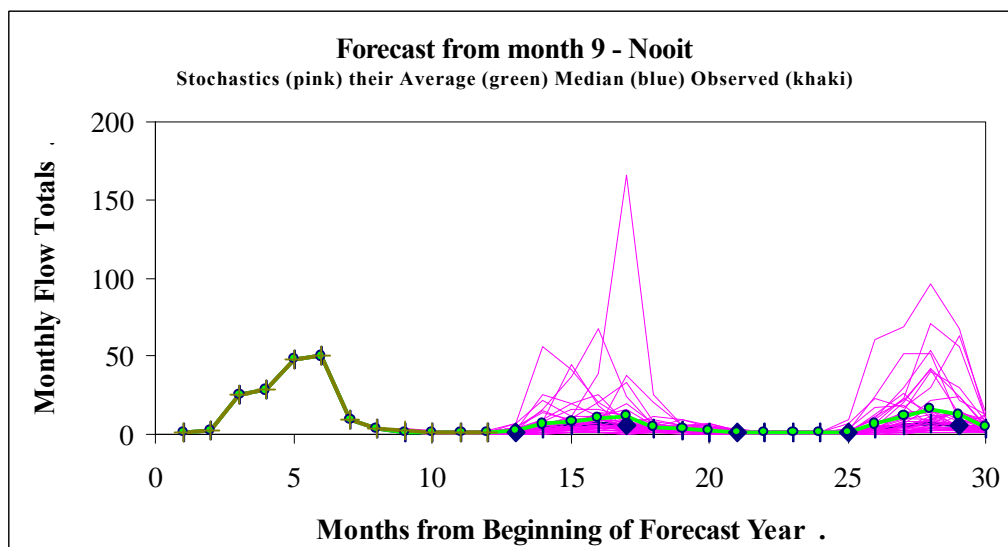
**Figure A.9 : Monthly forecasts for Nooit starting in month 5 of 1995**



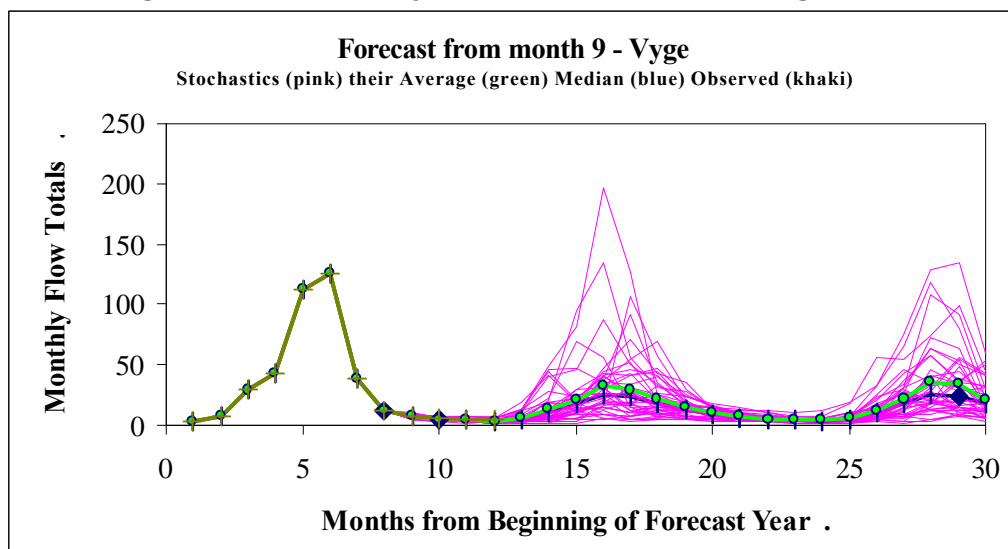
**Figure A.10 : Monthly forecasts for Vyge starting in month 5 of 1995**



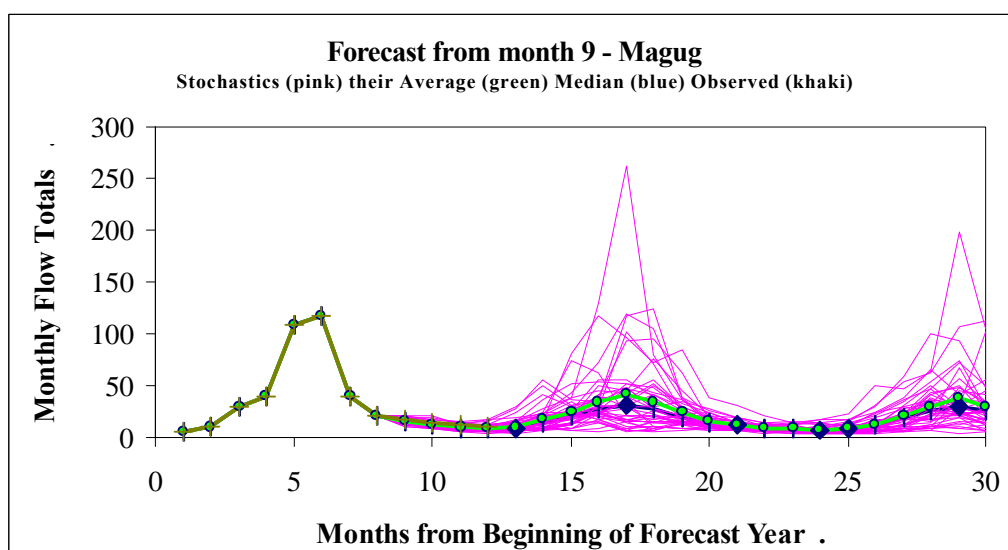
**Figure A.11 : . Monthly forecasts for Magug starting in month 5 of 1995**



**Figure A.12 : Monthly forecasts for Nooit starting in month 9 of 1995**



**Figure A.13 : Monthly forecasts for Vyge starting in month 9 of 1995**



**Figure A.14 : Monthly forecasts for Magug starting in month 9 of 1995**

## A.4 **Summary and Recommendations**

A subroutine ANSKOM has been derived from ANSMK5, the 1996 version of ANSIM for the specific purpose of forecasting monthly flows for the Komati system. Once it has been proven, and users are comfortable with its output, it will be useable in other systems where monthly flow forecasts are required.

There is a warning that must be heeded, however. In none of the flow sequences observed so far in the Komati basin is there a single zero flow. Because of this, the log transform works well. This will not be true if the routine is to be used in parts of South Africa west of the Komati where ephemeral streams become more common. A different treatment will be required to generalize the routine, although its basic architecture should remain much the same.

ANSKOM can be used in place of ANSMK5 in the usual way - it is transparently the same when annual disaggregation flow sequences are desired for long term planning.

To forecast with ANSKOM in the Komati system, the user will need to:

- update the copy of PARAMFOR.DAT to include the most recent data, ensuring that the missing data are indicated by negative numbers. At this stage this file should be in the same directory as GNKOMATI or whatever calling program is used. The code will have to be altered if another option is required
- set IFLAG = 3 to activate the forecasting mode
- ensure that PARAMK5.DAT (the Komati version) is present and is not altered from its current configuration.

The output from ANSKOM is, on the surface, indistinguishable from that of ANSMK5. Each call returns an array of data (through COMMON as before) which comprises a set of the monthly flows for each station, a full year at a time. The first call will return (in all the sequences) whatever intact months there are in the last observed year with the remainder filled with forecasts. Subsequent calls will return sets which are conditioned on the previous ones' last four months. Of the 41 sets of monthly flows, the first one is the median of the future flows, produced by omitting the noise term in the calculation, in the first call of the set to ANSKOM. The remaining 40 are plausible future conditioned stochastic flow sequences.

The median sequence is essentially a deterministic sequence and provides a benchmark. The average of the 40 stochastic sequences will be close to the expected future values and will lie above the median because the monthly values are so skew. This average is, in a statistical sense, the 'best' estimate of the future - the median is a bit more conservative.

### A.4.1 **Suggestions for Future Development**

Future work to improve ANSKOM and allow its generalization to other river systems includes:

- improving the preservation of the cross-correlation
- investigation of the necessity for twelve individual monthly cross-correlation matrices and possibly a seasonally varying serial correlation structure
- consideration of a hybrid annual/monthly model for long-term generation and forecasting combined
- automation of the agreed estimation procedure.

The most pressing of these seems to be the automation of the estimation procedure. This could proceed incrementally and be modified as more sophistication is built into the model.

At this stage it is quite difficult to anticipate the likely problems that will be encountered with South Africa's semi-arid hydrological environment. The problem of ephemeral streamflows in annual sequences was dealt with in ANSMK5 by assuming that the annual flows were uncorrelated. This assumption does not hold with monthly flows, so a greater level of sophistication must be employed in monthly flow modelling (as against annual flow disaggregation).

Monthly flows of major rivers in more arid parts of the country do go to zero. Building in a correct serial correlation structure under these circumstances can proceed in one of two ways: one can set the zeros to a small positive number (this is the old 1960's solution to the problem) and proceed with the methodology used in ANSKOM, or one can explicitly model the occurrence of zero flows as a mixed binary and continuous distribution, as is done with rainfall.

The first option (set the zeros to a small positive number) is the "quick and dirty" one which will probably work quite well in the preponderance of cases. The mixed binary/continuous option would need more care; models exist which could be adapted to suit the application. The decision as to which approach is adopted will materially affect the amount of time required to produce the requisite software.

#### **A.4.2 Recommendations**

Although it is desirable to investigate and model the natural streamflow processes as faithfully as possible, it is recommended that the first option (converting zero flows to small positive quantities) be adopted in the first instance. This would permit the development of a set of software which would yield parameter estimates automatically for river systems (in the Komati format) with no added complexity and no frills.

Existing software could be modified to produce a stand-alone monthly flow analysis package which would produce PARAMFOR.DAT files simultaneously with matched PARAMK5.DAT files. If desired, the two files could be concatenated and ANSKOM (or its successor) modified to accommodate this change with extra COMMON blocks.

The advantage of this approach is that it makes use of tried and tested routines which are familiar to users. The disadvantage is that it locks the procedures into 16 year old code which may need revision and overhaul.

In summary, the recommendation is that the ANSKOM route be taken as a pragmatic first step to establish the efficacy of the software in the South African environment and that thought be given to the overhaul and possible redesign of the generating/forecasting software to reflect the desire for more sophistication and friendlier user interfaces.

## **A.5 References**

Basson, M.S., R.B. Allen, G.G.S. Pegram, A. van Rooyen, (1994). "Probabilistic Management of Water Resource and Hydropower Systems". Water Resources Publications, Colorado, 424 pages.