Connecting Reservoirs and Water Power Stations by Markov Chains

"Anybody who can solve the problems of water will be worthy of two Nobel Prizes, one for peace and one for science"
(President J. F. Kennedy)

"Water is a major challenge for the twenty-first century but presents itself as well as a catalyst for peace. Anyone who invests in water, invests in peace."

Water: At the crossroads of peace and science
by András Szőllősi-Nagy, Zaragosa 2004
Thanks is due to my excellent late teachers, professors at BME then JATE, and I express my separate gratitude to my vocational mentors, István Zsuffa and Pál Rózsa - engineers and mathematician scholars - for their cooperation and stimulation spanning over decades, and to my beloved family members who the burdens of my complex vocational and public life/civil manias borne.

Summary

From 1973, based on studies carried out by Roux and Bernie[1], author took part in research activities regarding sizing reservoirs - that time of Fehérvárcsurgó, Hungary - initiated by István Zsuffa. Shortly after the first results we have extended our investigation to cover multipurpose reservoirs, secondary energy production. Based on the interaction between modeling achievements and real world applications, author targeted the core of the method thorough hydrological reformulation of the hydropower model. In this treatise a thorough review - worked out down to the level of engineering applications - of a queuing theory model - attributable to a basic equation formulated independently by Australian Moran and Russian Savarensky - and files including Fortran sources and data of 19 examples are found hereby to be published electronically. In the first part original methods developed with my collaboration under recent decades is presented. I outline the construction and sample use of more numerical models implying a number of new scientific elements at the time of formation and still cited until today. In the second part - considered as appendix - a handfull of presented sample applications from three continents cover only a fraction of recent and past uses. Overseas publications and research reports reference to the research - paused for nearly two decades despite my intentions - even with the passing of this much time, quoted as practical one emerging from among the theoretical experiments: in "Comparative Evaluation of Generalized River/Reservoir System Models" by Ralph A. Wurbs at Civil Engineering Department Texas A&M University (Reference and Users Manual for the U.S. Army Corps of Engineers) : "In terms of practical usefulness, the most important storage probability theory models are described as probability matrix methods (McMahon and Mein 1986). Zsuffa and Galai (1987) address probability matrix methods from a practical applications perspective and provide computer programs for implementing the methods. Other methods are of theoretical interest." References to our preinternet publication can be found on URL scholar.google.com.

I spent one and a half decade with the development of the methods of storage, based on engineering knowledge. My role was not restricted to programming, it evolved in a big leap.
forward regarding the model's development and the examination of the numerical constructions following from. Emerging questions and doubts affecting the essence of the model arise. Responding to questions stemming from these concerns lead to the reformulation and development of both the model and the technical and computational algorithms stemming from it. New ideas, new elements differing from or complementing the international vocational fashions may be born from literary unawareness, as it is shown by the vocational welcome of our comprehensive handbook (Reservoir Sizing by Transition Probabilities, Water Resources Publication (WRP), Littleton, Colorado, USA ISBN - 0-918334-62-4, Library of Congress Catalog Card Number - 87-51100) published in 1987, see book reviews by McMahon, Kottegoda.

In his 2005 book McMahon refers to our book registered in my copyright proportion of 73%, also his publisher recommends ours as a supplementary accessory to McMahon's. In that very same year of 2005 besides our work having been referred to at the Universities of Athens Technology, Texas A&M, and in the Encyclopedia of Water (Wiley, New York), and in a U.S. Army Corps of Engineers research report, the catastrophe analyses made in 1974 and 1982 plus my water power calculations rewarded me as a co-author with the title of "Best Scientific Publication of the 20th century in Mongolia" by the "Great Khural State", the Mongolian State Assembly / Parliament on the recommendation of Scientific Academy of Mongolia. There should be something time resistant.

Now, that I settled down to sum up an up to date revision wearing the wind of change and wishing to serve the survival of the method and the knowledge accumulated, my role in reformulating Moran's model will be accentuated in a numbered list bellow following this personal retrospection. In the framework of the dissertation - both for the data increase of the time passed and for the informatics environment totally changed in the meantime - I reviewed the descriptions of the models and examples to prepare reconfiguration of the newer variants of the self-developed software package for the sake of the next bigger technological migration to meet requirements of up to date conformity regarding networking practice, standard office automation and near future IT development. During gathering, storing, running downstream and putting on power generation, water does not only evaporate, roll stream deposit, but winter by winter due to the considerable cooling of the environment forms ice on the surface and transports it. Most fundamental feature of ice phenomena is ice cover rate, determination of which could happen by digital (web)camos or mobile phones as well. With this water related computer vision application I enhance my reservoir software package withstanding the test of time as proved by professional literature. Present publishing of the method especially timely, the population explosion causing bigger water shortage than the climate change and the approaching energy dearth, all secure a more important global role for the storage.

Main objectives and methodology

In terms of practical usefulness, the most important storage probability theory models are described as probability matrix methods. We address probability matrix methods from a practical applications perspective and provide computer programs for implementing the methods. Methods of over sophisticated theoretical interest is not dealt with herein. Goal is matching water supply and water and/or energy demand. Waters, our renewable resource - becoming more and more allocated - fluctuate following the random processes of weather. Water users must adjust themselves to these stochastic processes which comprise a random
sequence of highflow and lowflow periods. Based on simple quantitative assumptions and everyday theoretical and numerical basics, a tool for practical engineers and educators aimed to be delivered. This tool is a group of models for performing various analyses for both a single reservoir or some interconnected hydro utilities, even serving multi purpose by attenuating floods, delivering urban, industrial, agricultural supply and/or producing energy. For the selected type of Moran based model, problem identification and model formulation, data preparation and computational details with output presentation and evaluation of results is provided, the latter via examples.

Stochastic storage theory and related models have been addressed extensively in the research literature but applied very little by constructors and operators of reservoir systems. Group of analysis methods is based largely on the theory presented by Moran (1959) and expanded by Gould (1961), Klemes (1981), McMahon and Mein (1986). The objective of this type of stochastic storage theory models is to determine the probability distribution of reservoir storage. Storage probabilities may be computed at steady state or as a time dependent function of the starting conditions. Thus, for a given release policy and initial storage content, the probabilities of the reservoir being at various storage levels at future times during the next several months or several years may be estimated. As the analysis period becomes longer, the storage probabilities at a future time are no longer dependent upon the starting storage contents, a steady state condition is reached which is the base for reliability model for long-term planning of a multiple-purpose reservoir.

The stochastic storage theory models assess system performance based on describing inflows by a probability distribution or stochastic process. The methods typically applied to single reservoirs should also be developed into multiple reservoir analysis procedures. While inflows are assumed independent, - through fitting/reading from historical streamflow record - probability distribution of reservoir levels are determined via Markov chain of a stochastic process. Based on this stored volume distributions, the storage versus yield function and corresponding reliability estimators are calculated. Discrete probabilities are used to approximate the continuous distributions of the inflow process. The assumption of first order Markovian processes for representing the inflow process of a reservoir has generally been considered in the literature as adequate for most purposes. Much of the work published in the literature represents modifications or extensions to the basic Moran and Gould models. Moran (1959) presents various procedures for determining storage probabilities. Numerous other authors have presented solutions or extensions to the basic models formulated by Moran.

Like all group of practical procedures our approach treat both time and volume as discrete variables. A reservoir is subdivided into a number of zones and a system of equations developed which approximate occurrence of the possible states of the reservoir storage. Two main assumptions can be made regarding the inflows and outflows, which occur at discrete time intervals. In a mutually exclusive approach, there is a wet period, with all inflows and no outflows, followed by a dry season, with all releases but no inflows. In the more general simultaneous model, inflows and outflows can occur simultaneously. How they change the basic equations and influence the results, how they relate is shown. This also open the way to instationer models applying the different approach within the yearly cycle of time periods based on the nature of inflows and demands of the time interval of a given sequence. The selection from the two approaches is depends on what the demand is for. Estimating energy production should fundamentally be treated in simultaneous model otherwise significant biased output expectations arise.
A nonsteady state analysis can be useful in developing and implementing reservoir operating plans in which allocations of water to alternative users are made at the beginning of each water year, each irrigation season, or other time period of interest, based upon the likelihood of water being available to meet the allocations during the time period. The likelihood of meeting the allocations would be based upon the reservoir storage levels existing at the time the allocations are made. Under this type of operating plan, during drought conditions, as significant reservoir drawdowns occur, the allotment of water to the various users for the upcoming irrigation season or other specified time period is reduced accordingly. Storage probability theory models provide useful information regarding the probabilities of the reservoir being emptied by the end of the time period given the known present storage level and assuming different alternative withdrawal rates. Steady state probabilities are not dependent upon initial storage levels. In this case, storage probability theory models represent an alternative to regular simulation models, using period-of-record or synthetically generated streamflow sequences, for developing yield versus reliability relationships.

Estimation of water supply and/or power production are discussed below by using queuing theory. The relation between subsequent supply and release systems, and that of simultaneous ones is proven. A minimal approach uses historical yearly or monthly streamflow record and topographical map of the valley. Dam section site selection. Power plants supply a defined energy demand within limits of the turbine capacity. The energy demand driven reservoir model. Expected amount and distribution of shortages. Optimizations. Linear algebraic techniques.

As recent decades proved we have succeeded in giving into the hands of practicing engineers and educators a quite sophisticated versatile tool still having easy to touch relevance to the original physical problem. As in the models based on real technical problems, common sense is needed as the most important necessity, the arising models consist of the coincidence probabilities of independent events describing the phenomenon and quantities examined, resulting from convolution of values of discreet or supposed to be discreet distribution functions. Often plain, but not always obvious transformations of basic relations of the physical model - or algebraic equations stemming from them - turn into starting points of new solutions. I summarize my new scientific results discovered in the course of research in 8 theses.

**Own scientific achievements**

Numerical models - set up from basic relationships formulated while solving technical problems - often lead to predictability problems, or resulted in algorithms later proven to be too slow and/or requesting huge memory. The basic fundamentals of the model is often rewritten for the sake of solving or bypassing of these problems, or even merely for the precise problem definition required by coding task. In the course of reformulation newer and newer model modifications arise, though they are often trivial relations, but applied for the first time however; the exact justification of relationships - accepted earlier as experimental ones - appears now and then. Source of new, surprising results could be ignorance, which is not a miracle, young persons draw their new ideas from this just often. This may as well occur to others - though not beginners - being kept at a distance from vocational circulations, compulsion quickly begat some DIY, handymanness promptly produces very many new tools, methods to solve their tasks. At the beginning this was the situation in my case as well, since
only the more equal ones or merely geographically more favourable situated ones had their access to the international vocational information through decades in our homeland. Nowadays the quantitative explosion of the informational web may cause a "disadvantage" like this, a lot often faster to code - or say old-fashioned to "program" - something, than to sort out the equivalent from among a lot ready to use applications of vague origins put out on www.

The thesises listed hereunder are only a mere slice of results - all begat by practice-, and were born during my programming reservoir sizing. Mostly the basic equations and the wishlist was what I received only, and while I was carrying out patching the numerical part I often had to twist the model. Sometimes of that size to wind, that already not too resembled the original one very much. That might have been the reason for acknowledgement I deserved by renowned Australian professor McMahon's book review.

But not only model changes lead to new results, it was often necessary to drastically reduce the quantity of mass calculation and while solving this numerical problems reflexive novelties with a feed back to model structure arose sometimes. Taking into account the square growth of consumption field and capacity range to be examined and the examinations belonging to all of the steps over this area - on dots each claiming multiplication/division operation around power of $n+2$ - there are still what to win by conceptual considerations. Suggested by common sense, then justified by my numerical examples, intuitions accelerated the algorithms on a large scale. This is not an exact justification yet on the other hand. Shortly after the questions were put to more university algebraists, - I realized that it is not necessary to march with cannon onto sparrows, a mere sling will do - it turned out that a few problem - being not far from triviality - were solved by myself with plain transformations. This it is not indication of that false mirage how much I superseed my algebraist masters, but, that dispite their limited pure mathematical knowledge now and then application engineers, knowing the model better, get occasionally onto results much more quickly, than theoretical mathematicians got used to more complicated, theoretical problems. Wheel and Rubic's cube are plain commodities, but first somebody had to discover them after all! (the person presumably was rather engineer, than a mathematician... their turn comes later... e.g. at CFD flow fundamentals and numericals?:o)

The majority of the ideas I firstly "discovered" belongs to reservoir modeling, covering various algorithmic extensions of application areas of queuing theory models starting from the deficiency distribution across economic optimization until power generation. As the majority of methods resulting in Markov chains - in our case each single application extension eventually winds up in convolution calculations, what is not a miracle since these are generally interpreted as summation of convolution integrals on discreet distributions according to the model equation, materializing in multiplications of various transition matrices. Taking advantage of their numerical and matrix algebraic characteristics, progress in model building and solution to numerical problems can be made. During my various researches, convolution as functional operator turned up in an extended manner like in computer vision.

**Thesis 1: Duality of transition matrices describing subsequent incoming water and random demands**

I described the input-output symmetry of the Moran storage equation by a pair of transition matrices of the same but transposed structure belonging to input and output subsequent
subprocesses respectively. \( \xi_t \) stored water, \( I_t \) inflow volume at time \( t \), \( K \) storage capacity, \( M \) consumption volume, transition matrix \( F \) based on distribution of random consumption \( f_{mt} \), that of inflowing water is \( D \), transition in time period is described by matrix \( A \).

\[
\xi_t = \max(\min(\xi_{t-1} + I_t, K) - M, 0) \quad A_{k,f_m} = F D
\]

Moran published his original model in 1959, and volume were discretized by units selected expediently, reservoir's examined as a Markov chain based on a basic storage equation. According to this storage model's assumptions, independently incoming water quantities of identical distribution are followed by water consumptions of a given fixed volume. Stochastic variable, water quantity left over in the reservoir is the characteristic to be examined, where the construction of transition probability matrices derived from. Inspired by hints of info on parallel modeling activities István Zsuffa raised attention to the case of random water demands of known distribution. Following combinatoric considerations of probabilities of matching supply and demand, while random walk sorted into transitions between given inner states some formulae were formulated for transition probabilities, joined in a matrix, factorized by professor Pál Rózsa for the sake of the numerical examinations. Being present as a kind of go-between, the symmetric regularity of the factors just promptly caught my eye, and the symmetry of the Moran storage equation regarding input and output volumes was immediately hitting my brain as enlightenment that transition matrices of the same but transposed structure should be used for describing both input and output subsequent subprocesses, and with this all the model's calculation, examination, and - what is not a last viewpoint, the education - understanding I greatly simplified. [4]

\[
\xi_t = \max(\min(\xi_{t-1} + I_t, K) - M, 0)
\]

\[
p_i = P(I_t = i) \quad f_m = P(M_t = m)
\]

\[
P(\xi_{t+1} = i) = \sum_{j=0}^{k} P(\xi_{t+1} = i \mid \xi_t = j)P(\xi_t = j) \quad A_{i,j} = P(\xi_{t+1} = i \mid \xi_t = j)
\]

\[
A_{k,f_m} = F D
\]

\[
D_{i,j} = \begin{cases} 
0, & \text{if } i < j \\
p_{i-j}, & \text{if } j \leq i < k
\end{cases} \quad D_{k,j} = p(I \geq k - j) = \sum_{h \geq k-j} p_h
\]
Thesis 2: Factorization of yearly product matrix of non stationary model at the period of the largest water demand in year

In the course of computational practical applications I found out that the size of transition matrices of Markov chains belonging to seasons, or even months - of different non random consumptions - following each other in yearly cycle depends on the number of possible states, their product form yearly transition's matrix which claims minimal time and space when factorized at the time period of largest consumption.

\[
F_{i,j} = \begin{cases} 
0, & \text{if } i > j \\
f_{j-i}, & \text{if } j \geq i > 0 
\end{cases} \quad F_{0,j} = p(M \geq j) = \sum_{h \geq j} f_h
\]

\[
F = \begin{bmatrix}
1 & 1 - f_0 & 1 - f_0 - f_1 & \cdots & 1 - \sum f_{h<k-2} & 1 - \sum f_{h<k-1} \\
0 & f_0 & f_1 & \cdots & f_{k-1} & f_k \\
0 & 0 & f_0 & f_1 & \cdots & f_{k-1} \\
\vdots & \vdots & 0 & f_0 & \cdots & f_{k-2} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & f_0 
\end{bmatrix}
\]

\[
D = \begin{bmatrix}
p_0 & 0 & 0 & \cdots & 0 & 0 \\
p_1 & p_0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & p_0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
p_{k-1} & p_{k-2} & \cdots & p_1 & p_0 & 0 \\
1 - \sum p_{q<k-1} & 1 - \sum p_{q<k-2} & \cdots & 1 - p_0 - p_1 & 1 - p_0 & 1 
\end{bmatrix}
\]

\[\text{[4]}\]

In the course of computational practical applications I found out that the size of transition matrices of Markov chains belonging to seasons, or even months - of different non random consumptions - following each other in yearly cycle depends on the number of possible states, their product form yearly transition's matrix which claims minimal time and space when factorized at the time period of largest consumption.

\[
B A x = x \quad \& \quad A x = y \quad \Rightarrow \quad A B A x = A x \quad \Rightarrow \quad A B y = y
\]

A, B consecutive transition matrices, x and y their consecutive ergodic state distributions.

As an example of the public private partnership, decision makers of the district of the planned reservoir should agree on the distribution of the water supply needed. Frequent lack of consumption probabilities disable consideration of random water deliveries. Not to force decision makers, economic and agricultural designers to answer emerging questions regarding fluctuating releases, we reckon with fixed water consumption in the majority of the cases. In a more frequenter case however, the proportion rates of the requested water quantity in time - e.g. inside a year - is found desirable to a certain extent. The model represents sequent of events not as consecutive periods of identical feature - that is years -, but repetitive series of seasons, even months, stretching over not necessarily equivalent lengths of time. This time
unit refinement could be increased while the consecutive quantities of arriving waters are independent at the required measure of reliability, otherwise we have to resort to other model or numerical methods other than Markov chains. Since water deliveries over periods inside the year follow unknown distributions, the examination is made for the multiple values of consumptions of fixed proportion over seasons/months. Markov chain state transformation matrices - of every periods following each other cyclically - constituted on the basis of model equation, are of size depending on the number of possible transition states, their matrix product make up the yearly transition matrix, with this multiplication the instationer model is traced back to stationer one of presupposing the consecutive periods with an identical feature. Since size of this product matrix depends on which subperiod we start analysing the year, thus I manifested that at the subperiod (season, month) with largest consumption is where the year is worthy to cut into cyclical subperiods. Not merely place is saved this way, but taking into account the square growth of consumption field and capacity range to be examined and the examinations belonging to all of the steps over this area - on dots each claiming multiplication/division operation around power of 5 - there are still what to win by conceptual considerations. Suggested by common sense, then justified by my numerical examples, the cyclic product can be started at arbitrary subperiod, the results do not change, since [4]:

\[ B Ax = x \land Ax = y \Rightarrow AB Ax = Ax \Rightarrow AB y = y \]

**Thesis 3: Equivalency relationship of sequential and simultaneous water consumption models**

I transformed the basic equations of the two mutually exclusive hypotheses into identical forms, so that the two models just lead to identical results at capacities differing from each other exactly with the value of consumption, and therefore they are numerically equivalent.

Frequent objection to this type of storage models is, that in the original basic equation, inflow precedes consumption thus being separated in time results in oversizing compared to the reality. Twisting the basic equation to simultaneous filling and consumption, contrary finding derives. Inflowing water quantities arriving before possible consumptions are taken into account as staying in the reservoir, although some part of them overflows earlier than any supply releases occur in reality thus latently increasing safety of water deliveries which leads to undersizing. This dual approach catch the boundaries of over- and undersizing and yields a plain solution: adding and subtracting the consumption onto both sides of the equalities of the basic equation derives an identical form of the two equations, capacity is bigger with the consumption value in the subsequent form of the simultaneous model, otherwise we receive the very same results. Thus I manifested the equivalence between the simultan and separated inflow/outflow models. With this finding for the deriving results, I received a field of tolerance between the boundaries representing over- and undersizing where the reality somewhere between the two could be.

\[
\xi_{t+1} = \max(\min(\xi_t + I_{t+1}, K) - M_{t+1}, 0) \quad \xi_{t+1} = \max(\min(\xi_t + I_{t+1} - M_{t+1}, K), 0)
\]

after a minor transformation second becomes equivalent to first
\[ \xi_{t+1} = \max\{\min(\xi_t + I_{t+1} - M_{t+1} + M_{t+1}, K + M_{t+1}) - M_{t+1}, 0) \}\]

All above thoughts turn into contrary in the case of energy production, since if water runs on turbines after inflow finishes, then comparing to real circumstances - of constant feeding, continuous power generation - a bigger hydraulic head is taken into consideration resulting in undersizing. Following common sense in calculating power production - Moran type assumption of subsequent supplies and releases should be rejected and - model adjustment now requests a substantial modification.[4]

**Thesis 4: Modelling simultaneous inflow and power generation**

I presumed all volume changes to be simultaneous and uniform in time, the integral by the volume - to get potential energy function – I transformed by way of substitution into integral by the time and vice versa. In my modified model the proportion of energy desired and the power generated by change of state show together with the change in storage volume how much inflow water needed from which the transition probability follows:

\[
p\left(I_t = \frac{E_t}{\gamma \eta \int_j^i (H(V) - H_0) dV} - 1\right) (j - i) \Delta V
\]

(E\text{t} requested energy in time t, H(V) inverse function of volume-waterlevel V(H), H_0 turbine elevation, \Delta V volume unit of discretization, \gamma, \eta gravitation, water density, rate of efficiency)

In our first energy estimations - probabilities of reservoir states, water surface altitude and quantity of fixed consumption served as the computational basis for power production. Excess turbine capacity makes it possible the use of overflow waters too, thus as a result, beyond fixed consumption which basically determines storage operation, I also took into account water quantities partially or fully delivered to turbines after arriving into full reservoir. Diversion of spillway water form maximum hydraulic head onto turbines showed considerable efficiency increase so leaving overflow out of consideration caused significant oversizing in most cases.

To increase computational reality, expected energy values produced by transitions resulting empty reservoir were also considered since water shortages of different magnitude occur. Although calculations of power generation - based on adjusting traditional Moran model to hydroelectric plant - is repaired somewhat this way, but assuming or ignoring of chronologic separation of random inflow and fixed volume water consumption influences the results fundamentally. Furthermore, after accepting condition on temporal separation, role of fluctuation of hydraulic head regarding water level changes within the time unit arises immediately. My finding, the equivalency relationship between sequential and simultaneous water consumption models does not help either, since regarding power generation, results from either simultan or separated inflow/outflow differ fundamentally only in near full and overflow situations.

Regarding reality closer cases of transition matrices - formulated according to inflow with simultaneous fixed consumption; - which hydraulic head should be accounted while computing energy potential of water run on turbines is still an open question. There is no doubt this
approach shows expected energy between more realistic limits, but due to the magnitude of
difference between the energy results of consumptions of simultaneity and sequentiaity,
upper and lower limits of energy output could be a manifold multiple of each other, thus
undermining customer's confidence towards designer's credibility. Serious tasks came up
meanwhile, like estimating power production of a reservoir chain on River Niger collecting
water from a subcontinent-sized river basin. I could not have followed the usual methods, I
must have turned from the regular drive to take the task in a totally new approach; I reached
the model's kernel retaining the consistent part, fundamental ideas and the computational
manner of the method. What was left over then? Discretization of time and volume variables
as well as lower and upper limits: when less water comes reservoir empties, when a lot then
overflows.

Aim is the energy - quantity of which fundamentally depends on the neighborhood
morphology of the dam section -, thus formulation of transition matrices of this Markov chain
process should be done from potential energy function regarding stored water volume
referring to dam section's turbine level, from this special space informatics tool, got by
numerical integration of the map. Based on digitized map, this function has already been
created while selecting geometrically optimal dam section to be examined as power plant
operation site after complemented by hydrographical data, so determining water consumption
from potential energy function seemed to be obvious. But how continously changing
hydraulic heads shall be taken into consideration? From what altitude and until how much
time the water does arrive onto the turbines? I had to solve countless untold additional
emerging questions and doubts by a masterly stroke.

Due to the aspects of power generation all volume changes presumed to be simultaneous and
uniform in time, the integral by the volume - to get potential energy function having been
used in site selection based on geospatial data - could be transformed by way of substitution
into integral by the time and vice versa. This realistic and plausible hypothesis opens the way
to model the processes determining transition probabilities. By accepting this reality nearer
condition neither time nor altitude data are needed to be dealt with, all such information
considering the full morphological data set is already included in potential energy function.

When modifying basic equation for a more accurate estimation of electricity production, it is
not the fixed consumption to take into account but the amount of water needed to produce
required energy which depends on hydraulic head as well, that is: the energy generated
depends on current amount of water stored in the reservoir; power produced is also subject to
inflowing waters arriving at the same time. Presupposing uniformity and simultaneity in time,
now we have to take advantage of all topography interpolations to define random walk
probabilities in a measure falling nearest to the reality.

Potential energy function of $E(h)$ specifies the energy produced by all water quantity run on
turbines during deplating the reservoir from an altitude of $h$. To determine transition
probabilities estimation of power generation between inner states is needed. Power generated
during transition from $h_1$ to $h_2$ is given as $(E(h_1) - E(h_2))$ which should be less than the value of
requested energy to produce otherwise this very transition can not be established because it
generates more electricity than the given time period needs, i.e., its probability is zero. But if
it accurately meets the desired energy to produce, water could not arrive into the depot,
because it would leave through the turbines and more energy would be delivered hereby than
requested.
If the given change of state from \( h_1 \) to \( h_2 \) would not be able to supply the planned energy quantity without additional inflows, a certain calculated amount of arriving water passing through turbines would insure the missing energy then. Since inflow, power generation and change of state happen all at the same time and are uniform throughout the time unit in question (year/season/month), energy produced during the given change of state by a unit water quantity - originating indifferently from either stored water or just flowing in now - is \((E(h_1)-E(h_2))/(V(h_1)-V(h_2))\) which is the quotient of energy produced by a change of state and the difference in stored water volume. Divide the requested energy quantity by unit water generated electricity in order to find out how much water - following the a known distribution - has to arrive into the reservoir which provides the transition probability in question.

As a general rule, no matter whether the calculated water quantity running through turbines during either depletion from level \( h_j \) to \( h_i \) or filling up from level \( h_i \) to level \( h_j \), due to the uniform turbine discharges the same time is spent on the same interim hydraulic heads both in downwards and upwards water level changes. Calculated water quantity to pass onto turbines equals the sum of arriving water and the change of stored water.

The full subsequent state represents a special case in state transitions, not only the water quantity satisfying energy needs is showered onto the turbines, but diversion of spillway water is put on power generation as well up to the turbine capacity. When overflow occures the pressure on turbines is at maximum, the model is extended with the calculation of energy surplus produced by diverted spillway water. Proportion of energy gained from overflow waters indicates that the given power plant section can energeticwise better be utilized.

In case of hydropower the reservoir is not built exclusively to store randomly available excess water for shortage periods but also to ensure sufficient hydraulic head making power generation possible. The subsequent examinations - being based on transition matrices constructed from claimed energy quantity - may bring to light changes deriving from development or change in operation of existing reservoirs, and a current operation may be tested as a sort of forecast with the examination starting from a known state, and is done as a single matrix multiplication in each time unit. This chapter induced the acknowledgement in McMahon's book review cited above, and if the method survives my personal period, I surely hope for my name gets reference for this unusual, new approach requiring common wit merely, as the undermentioned compact formula shows transition probabilities concisely:

\[
p(I_t) = \left( \frac{E_t}{\int_i^j (H(V) - H_0) \, dV} - 1 \right)(j - i)\Delta V
\]

Presumed uniformity of water volume changes in time of simultaneous inflow and power production is not a necessity, fluctuating demand of electricity could be incorporated as a differenciable function factor tag in energy calculation integral thus accumulating some potential energy surplus according to power request distribution within days and weeks is also possible. [5]
Thesis 5: Drastic reduction in calculations regarding random demands on a reservoir capacity range

By computing stationary model’s consecutive capacities for a fixed consumption, I concluded that the matrix elements belonging to the common non full state transitions are identical. Matrices describe random walk’s target transitions by rows; last row of the minor matrix is the sum of the last two rows belonging to the next bigger capacity. Apart from the last row all numerical operations of the matrices are identical during eliminations along the main diagonal, therefore I obtained the results belonging to each of the smaller capacities in the course of the elimination of a transition matrix belonging to the largest capacity. With the help of an elegant transformation I generalized this algorithmic acceleration and model description to random consumption as well:

\[ FDP = P \]  \[ D = F^{-1}P \]  \[ (F^{-1} - D)P = 0 \]

In cases of consumptions and/or capacities being similar to each other, transition matrices formed from Moran type basic assumption perturbs slightly only. Fortunately in parts of the cases this perturbation makes a difference of a single diad merely. One of the answers to arising double questions serves a practical benefit, the other does as theoretical proof of a conceptual predicate. Question of the practice was how to get the result vector more cheaply - with less numerical operation - from another result vector? The emerging conceptual question responds to that obvious fact, that a smaller reservoir supplies the same consumption with higher risk, that it gets into lower states with a bigger probability, or at last empties with a higher risk. Diad related results stem from the fruitfull collaboration of Pál Rózsa who helped to get the distribution of stored water in reservoir of capacity K+1 units from results belonging to capacity of K with less numerical calculations.

From a striking yet trivial transformation of the beginning of above examination demanding mathematical knowledge with yet/already an uncharacteristic depth in engineering practice today, a more universally useful algorithm acceleration resulting in bigger saving crossed my mind which may originate from earlier observations during countless practical examples. The inner elements of the transition matrices belonging to consecutive reservoir capacities were identical. These matrices differ only in size, in last row hereby, that considering their rank was one maller than their order, this last row was foredoomed to be omitted. Since these matrices are filled up with transition probabilities, and we are looking for eigenvectors with sum of 1, by extracting unit matrix from main diagonal which thus turned out to be dominant, not even pivoting is needed. Starting elimination from upper left corner, all respective elements in every single transition matrix of all capacity belonging to the given consumption distribution and dam section are the same numbers throughout the whole elimination. In other words, while calculating any capacity, the number and operations in the matrices are completely the same, only the smaller capacity matrices run out of last columns and rows earlier, so in the course of calculating the maximum capacity all results of the smaller reservoirs is available as a byproduct. Similarly to the case of fixed consumption, with an elegant transformation this algorithmic acceleration is generalized to random consumption as well:

\[ FDP = P \]  \[ DP = F^{-1}P \]  \[ (F^{-1} - D)P = 0 \]
Necessary exponent of polynomial amount of operation is reduced by one, in case of capacity 50 close to 50 times acceleration is attained as opposed to velocity doubling of diad based speeding up.

The question arose was that this numerical acceleration can be used for the more complicated matrices of power generation just similarly to the case of the fixed and random water volume consumption? Examining inner state transitions of a power station on different capacities, concurrent volume changes presumed to be uniform in time, the process is affected by water(level) between starting and final states only, no matter e.g. how much void empty space is unutilized in the reservoir above. This on the other hand has a consequence, that apart from their overflow affected last rows, transition matrices - of reservoirs with different capacity compiled in a manner not depending on available empty surplus capacity - are identical, that is the method of acceleration is useful here as well. Since transition matrices - due to their identical sums in columns - are singular, a row, in our case the "irregular" last one belonging to full and overflow cases, can simply be omitted. [4]

*Thesis 6: Water Shortage Distribution*

I firstly explained that just like the subsequent reservoir state distribution is obtained from the distributions of antecedent reservoir states and arriving waters, in that very same manner and resources the shortage volume distribution of $\mathcal{K}_{t+1}$ can be computed via a discrete convolution integral executed by way of multiplication by a transition probability matrix based on the transformed basic equation:

$$\mathcal{K}_{t+1} = \max(M_{t+1} - \min(\xi_t + I_{t+1}, K), 0)$$

The basic deficiency relation does not differ in the case of random water demands either, only the character of one additional member - value of the planned consumption - turns into a stochastic variable of a known distribution; so an up till now fixed constant turns into a distribution of a single vector containing chances of possible values of consumption, and in the formula providing the probability of the deficiency events an additional sum is found, what means the multiplication with an additional matrix to be constructed adequately. To compile the transition probabilities the similarity of the undermentioned formulae of $\mathcal{K}_{t+1}$ $\xi_{t+1}$ $\mathcal{K}_{t+1} - \mathcal{K}_{t+1}$ provide help:

$$\xi_{t+1} = \max(\min(\xi_t + I_{t+1}, K) - M_{t+1}, 0)$$

$$\mathcal{K}_{t+1} = \max(M_{t+1} - \min(\xi_t + I_{t+1}, K), 0)$$

$$FDP = P - \mathcal{K}_{t+1} = \min(\min(\xi_t + I_{t+1}, K) - M_{t+1}, 0) \Phi DP = S$$

At an early stage deficiencies were considered as a slight exaggeration in favour of the safety, depletion of reservoir as risk of malfunction while complementer value as operational safety.
Similarly to others the expected shortage value were based on that assumptions. Economic calculations started causing problems later (e.g. operational water level regulation of lake Balaton) since there is a real chance of emptying a reservoir in a way that not even a drop is missing from the planned water consumption, and in more serious economic analyses parts of the water delivery claim different safety levels. Shortage distribution was needed for the choice of plantations of different water demand and for expected reservoir operation plans based on the distribution of the safety of water supply. Onsite interactive risk analysis provided ammunition for strategic masterplans, water and electricity of rivers of melted glacier was needed to feed and light the growing population in West Mongolia in the big neighbours' squeeze.

To change the estimate of deficiencies, it was necessary to improve the calculations of model refinement based on assumption closer to the reality. Shortages are just the same type of events like inflow, outflow. The feature to be calculated is the rate of failure of water consumption exactly. I described shortage event just in the same manner as Moran made it with the model's basic equation. Beside the values of consumption and capacity, reservoir state probabilities and distribution of arriving waters served as a basis to the calculation of shortage distribution. The same independence conditions apply to deficiency relations as the hypothesis assumed in the model's basic equation, so shortage distribution can be reckoned via convolution which due to the discretization is a matrix multiplication.

Estimating the distribution of the shortages I had to be aware of that particular fact, that in normal economic environment the damages caused by deficiencies are polyvalent, and only sections by sections can be considered for linear one. A striking example of section linearity is the case of multipurpose reservoir, where communal water shortage causes the largest damage, since to drink is a necessity, and hygiene also requests a prescribed water quantity per capita until a certain level. Second harm grade ensues when industrial water fails to supply factories demanding bigger safety of continous water services than agricultural water utilisation which is characterized with the lower next damage value. Anyway planning of plantations might already take it into consideration to irrigate less water if there is not enough, or to plant not so much from water demanding plants in the precognition of deficiency distribution. In many cases - out of formulae leading generally to matrix multiplications, those calculating expected value of shortage are not enough, - distribution of deficiencies or that of restricted consumption \( M \), is needed.

\[
P_r = p(M_r = \mu \Delta V) = \sum_{i=0}^{\mu} p(\xi_{t-1} = i \Delta V) \ast p(I_t = (\mu - i) \Delta V) = \sum_{i=0}^{\mu} P_i P_{\mu-i}
\]

The basic deficiency relation does not differ in the case of of random water demands either, only the character of one additional member - value of the planned consumption - turns into a stochastic variable of a known distribution; so an up till now fixed constant turns into a distribution of a single vector containing chances of possible values of consumption, and in the formula providing the probability of the deficiency events an additional sum is found, what means the multiplication with an additional matrix to be constructed adequately. To compile the transition probabilities the similarity of the undermentioned formulae of \( \xi_{t+1} \rightarrow \xi_{t+1} \) provide help:

\[
\xi_{t+1} = \max(\min(\xi_t + I_{t+1}, K) - M_{t+1}, 0)
\]
\[ x_{t+1} = \max(M_{t+1} - \min(\xi_t + I_{t+1}, K), 0) \]

\[ FDP = P \quad -x_{t+1} = \min(\min(\xi_t + I_{t+1}, K) - M_{t+1}, 0) \quad \Phi DP = S \]

Matrix multiplications claim huge operational load however, reducing their number would possibly be beneficial by performing internal subproducts only once and then storing these subproduct matrices. No matter that in case of deficiency the water run out because of water consumption or power production, the reservoir is empty, irrespective of how large a dam built or how much void empty space is currently unused in the reservoir above. The matrices used for the calculations are identical or very similar in their common parts. When computing deficiency the full identity of the matrix elements allows to foreshadow, that although these matrices belonging to different capacities differ structurally in their last row only. Last row - describing overflow, full state - may not have a big effect on the deficiency, since value of consumption should be smaller than capacity, otherwise the reservoir would be empty in all years, and there isn't a Markov chain and there is nothing to plan.

I introduced the subproduct matrix to generate the shortage distribution belonging to the consumption distribution because of this, the elements of which depend on the distribution of arriving waters and consumption only, it is necessary to create only once and we receive the deficiency distribution of the random consumption as subproduct matrix is multiplied by the vector of state distributions belonging to the different capacities after inflow.

\[ p(S_{\text{shortage}} = i) = s_i = \sum_{j=i}^{\infty} f_j q_{j-i} = \sum_{j=i}^{\infty} p(M = j) p(\xi_{\text{afterInflow}} = j - i) \]

\[ FDP = P \quad \& \quad DP = Q \quad \& \quad \Phi DP = S \quad \Rightarrow \quad \Phi Q = S \]

\[ AP = P \quad \& \quad BP = S \quad \Phi_{i,j} = f_{i+j} \]

To find out easily the expected value of the shortages of random consumption, subtract the the expected value of the consumption distribution from the difference between the expected values of water quantity in the reservoir after inflow and after consumption. [4]

**Thesis 7: Optimising reservoir on a given reliability and/or investment return**

*During my practical calculations, isoclines of probability field showed monotony, therefore I bounded grid dots of the capacity consumption subfield by predefined lower and upper safety. On the examined dots, cumulative values of expected consumption and loss functions weighted with deficiency distribution were compared with the investment and operational expenses. I applied this type of optimum search to the case of power generation as well.*
Capacity consumption grid of $k^2$ step number magnitude claims the execution of an algorithm of $k^3$ multiplicative operation in each dot. This in total represents $k^5$ order of magnitude operational step. Examination should be carried out over the morphologically appropriate range of reservoir capacities. Not necessary however to do the heavy calculation for all consumption from zero up to the values reaching as far as the order of magnitude of the capacity. Due to the large number of possible states resulting in huge operational load particularly in cases of low consumptions, reservoir empties quasi never and is quasi always full - too big the safety is. In opposit cases of few states belonging to big consumptions on the other hand quickly may we reckon, but the reservoir empties quasi always, and because of this not too much to reserve, the reservoir retain so few water that storage space is only exploited within time unit. Designers would generally like to examine around a prescribed safety, thus to determine optimal reservoir size and operation. This on the other hand presents the necessity to do the heavy calculation along only a line belonging to given water delivery safety on capacity-consumption field and as a result the examination of $k^5$ magnitude turns into one step lower in power number that is $k^4$.

Even more interesting is the optimization - between the claimed operational safety limits - carried out on the basis of economic data. As early as the mid-70s I made and finished at the request of and under the guidance of Dr. Zsuffa operational water level regulation examinations of Lake Balaton with the help of losses (eg in tourism) and (eg, fisheries, irrigation) profit functions and the probability of depletion. To refine the examination I formulated the algorithm determining the distribution of shortages serving as the basis of more sensitive economic investigations.
But unfortunately - similarly to the case of the distribution of water consumption - it is difficult to obtain economic data, and however influencing the sizing significantly, return time and operation expenses often does not play a role in many cases. Though apart from the investment expenditure, with the help of the operational costs and the profit hoped for in the function of variable delivery safety depending on communal, industrial, or agricultural type of consumer, the method may possibly turn into a serious long-term decisions supporting tool.

Though not explicitly related to water, morphological and economical optimization of the reservoir had already roots in the earliest period and more delicate has been finely burnished throughout since then. Difficulties were caused by the lack of digital maps and the unavailability or unreliability of economic data representing expenses and prices. We were ready to go ahead, and this kind of development of the model were blocked by the unpreparedness and/or immaturity of the application environment.

Applying the completed algorithms over capacity-consumption grids the convex surface search was justified. Some cases of multipurpose reservoir investigations included flood reduction as well, in addition to the water and energy utilisation, savings in downstream flood control investment were also taken into consideration as a negative supplementary investment expense in the course of optimum search. [6]

[8] Ice Inspection - Estimating Ice Cover Rate

*In my model ice cover estimation is an integral over the water domain of the perspective picture in investigation. The integral over the water of the camera picture is done by dx dy. Since we are summing the icy pixels on the picture itself, the integration is done by an*
integral transformation, it is a change of variables when integrating a function over its domain, here the Jacobian determinant is used. To accommodate for the change of coordinates the Jacobian determinant arises as a multiplicative factor within the integral. Following calibration focused on water plane I arrived on the following closed form of the Jacobian:

\[
\begin{bmatrix}
\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\
\frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}
\end{bmatrix} = \frac{1}{z} \mathbf{AR} - \frac{1}{z^2} \begin{bmatrix}
\bar{x} \\
\bar{y}
\end{bmatrix}
\begin{bmatrix}
r_{31} & r_{32}
\end{bmatrix}
\]

where:

\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = \mathbf{A} \begin{bmatrix}
x \\
y
\end{bmatrix} + t \\
\begin{bmatrix}
\bar{x} \\
\bar{y}
\end{bmatrix} = \mathbf{R} \begin{bmatrix}
x \\
y \\
z = 0
\end{bmatrix} + t \\
\begin{bmatrix}
u \\
v
\end{bmatrix} = \frac{1}{z} \begin{bmatrix}
\bar{x} \\
\bar{y}
\end{bmatrix}
\]

\[
\int \int f_{\text{ice}}(x, y) \, dx \, dy = \int \int f_{\text{ice}}(u, v) \mathcal{J}(u, v) \, du \, dv
\]

where \( \mathbf{R} \) is the rotation matrix and \( t \) is the translation, they are the extrinsic parameters. \( \mathbf{X} \) consists of the real word coordinates while \( \mathbf{Y} \) yields the image pixel coordinates. \( \mathbf{A} \), the camera intrinsic matrix, is given by \((u_o, v_o)\) the coordinates of the principal point, \( \alpha \) and \( \beta \) the scale factors in image for \( u \) and \( v \) axes, and \( \gamma \) the parameter describing the skewness of the two image axes on the light detector CMOS chip.

Ice cover rate on water surface does not depend on distance or unit, it is the function of an angle formed with the plane of the water only. On the other hand, I obtained this angle from the horizon line shown on the picture to be evaluated. There is no need for difficult on site calibration either. Similarly to stereometry or our eye, the real distances can even be calculated based on pictures taken even with mobile phone from two points of known distances. Based on coordinates exempted from radial distortion of pictures taken by mobile phones or cameras - I precalibrated in office environment - to get intrinsic parameters, horizon line equation \( a x_i + b = y_i \) results in closed form:

\[
\begin{bmatrix}
a \\
b
\end{bmatrix} = \left[ \begin{array}{c}
\frac{1}{x_i^2 - \left( \frac{\sum x_i^2}{\sum x_i} \right)} \\
\frac{\sum x_i y_i}{\sum y_i}
\end{array} \right]
\]
From parameters of the horizon defined this way the obliquity of the camera - its differing from vertical - I determined. Based on the pixel distance of the principal point of the camera and the horizon line on the picture, and the focal length of the optics measured in pixel the angle between the camera axis and the water surface I got as well.

Application

Our research was primarily driven by the needs of practical applications which significantly promoted the prompt utilization of the - by nature - basic investigations. The dissertation's appendix consists of real examples from three continents, case studies illustrate the model suitability for the examination of reservoirs of any size, purpose or operational complexity. Furthermore resulting numerical methods are directly applicable on all such storage tasks where approximational approach of the independence of arriving inflow water quantities in subsequent time units is acceptable.

Hundreds of reservoirs - planned and existing - were investigated and designed by public domain software accompanying my publications. There were countless applications even in Hungary without notifying software and modell author. Software pirates also sold rewritten copies, but no modell tuning or development were done later by others except for some sedimentation experiencing (including "discovering" non existent ergodic reservoir state distribution of sedimenting reservoirs:). Extra European applications included a reported master theses in South Korea in late 80s, Nigerian examples in early 1980s, wide ranging Mongolian application of my packages covered from mid 1970s to mid 1980s, while Algerian applications were on top between 1984-1987. By public release of the source package in 1987, wide spread application evolved from then on.
Short enumeration of my theses related literature activity


   [with copyright rate 19/26 (73%) in ARTISJUS contract]


