MANAGEMENT EPIDEMICS USING HIGH COMPLEXITY MATHEMATICAL MODELING

PART VI: SEIMR/R-S/ML-ST

COVID-19 SURVILLANCE EPIDEMIC STATE AND PARAMETER ESTIMATION USING A DUAL MULTI-STATE KALMAN FILTER (D-MS-KF)

Working Paper Version 1.04.1
Actual version:

Ing. Jesus M. Velásquez-Bermúdez, Ph. D.
Chief Scientist DecisionWare – DO Analytics
jesus.velasquez@decisionware.net

Bogotá, April 2021

EPIDEMIC STATE AND PARAMETER ESTIMATION USING DYNAMIC MACHINE LEARNING BASED ON A DUAL MULTI-STATE KALMAN FILTER (D-MS-KF)

Ing. Jesús Velásquez-Bermúdez, Dr. Eng.
Chief Scientist DecisionWare
jesus.velasquez@decisionware.net

"The Apollo computer used 2k of magnetic core RAM and 36k wire rope [...]. The CPU was built from ICs [...]. Clock speed was under 100 kHz [...]. The fact that the MIT engineers were able to pack such good software (one of the very first applications of the Kalman Filter) into such a tiny computer is truly remarkable."

Interview with Jack Crenshaw, by Matthew Reed, TRS-80.org (2009)

ABSTRACT

From the point of view of managing a pandemic, the mathematical problems to be known:

i) To define the structure of the models of differential equations that govern the behavior of the pandemic
ii) To estimate the parameters that define a specific model within the "infinity" of possible models to describe the dynamic process.
iii) To know the true state of the pandemic, which is defined by the number of people, or by the fraction of the population, which is in each epidemiological state.

The importance of an appropriate estimation of the parameters of mathematical models and the state of the pandemic is determinant of the decision-making process to control the pandemic.

This chapter presents the integration of **Machine Learning (ML)** and **State Estimation** to implement machine learning algorithms to understand the behavior of a pandemic system using advanced algorithms based on the fundamental concepts that support the so-called **Kalman Filter** (**KF**, Kalman and Bucy, 1961) and its combination with Markovian state and Bayesian inference.. The **Multiple-State Kalman Filter** (**MS-KF**, Velásquez, 1978) and the **Dual-Kalman Filter** (**D-KF,** Moradkhani et al., 2005) are integrated to model the state of the pandemic (the distribution of the population in the epidemic states) and the parameters of the differential equation epidemic model (SIR, SEIR, SEIMR/R-S,…).

MS-KF is a methodology that allows to combine: i) the process of determination of possible Markovian states where can be a system (a pandemic process) with ii) response functions identified by Kalman Filter for each of the possible states of the system. The selection of the probability of the Markov system state is determined based on a model of Bayesian inference for multiple possible models. Basic examples are included to illustrate the concepts.

D-KF is a methodology that allows simultaneously estimate the state of the process being analyzed and the parameters of the dynamic equations that govern the process.

The learning process associate to integrated **MS-KF** and **D-KF** (thereinafter **D-MS-KF**) can be termed as Deep Learning process, since it assumes the process of identification of parameters as part of the dynamics of the system, without pre-set a response function as stationary and certain. The dynamics of the **D-MS-KF** allows to re-estimate the system response function considering the information contained in the data more recent, it may include the dynamic selection of the structure of the differential equations that explain the epidemic process.

EPIDEMIC STATE AND PARAMETER ESTIMATION USING DYNAMIC MACHINE LEARNING BASED ON A DUAL MULTI-STATE KALMAN FILTER (D-MS-KF)

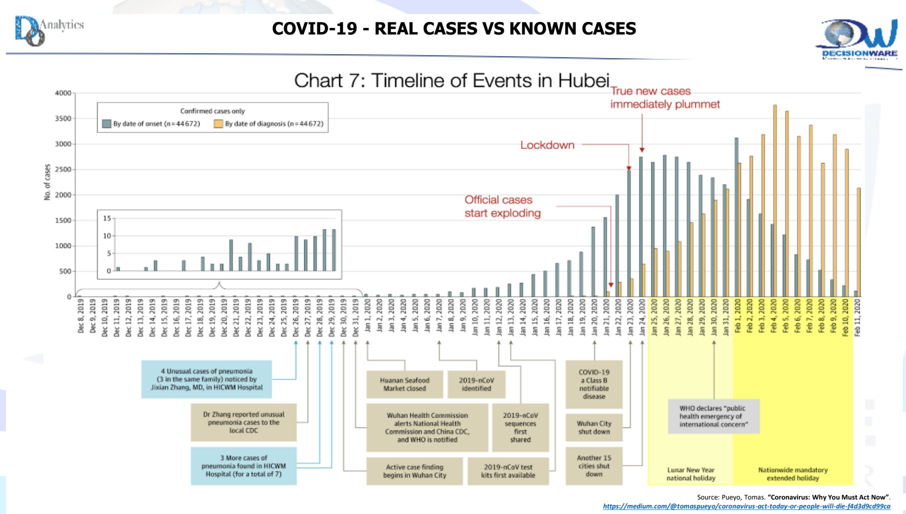
1. THE STATE ESTIMATION PROBLEM

At least two problems must be faced:

i) In new pandemics, such as COVID-19, the structure of mathematical equations can be unknown and is therefore part of the problem to be solved

ii) The system of measuring pandemic states can be very poor, mainly in the early pandemic and in countries with little capacity to widely measure the population and determine their epidemic status.

The following image presents the deface between the "real" information (blue) and information known (yellow) at the beginning of the COVID-19 pandemic in the city of Wuhan in China.

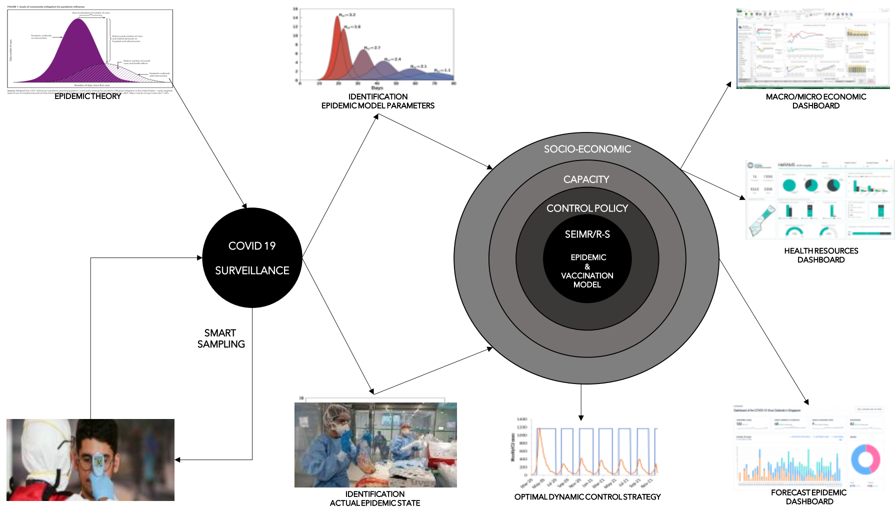


The TRIAGE model supports the pandemic monitoring and control system with the aim of:

i) Closing the gap between the real and the observed epidemic data, and

ii) Provide the information needed to calibrate the mathematical models of epidemic simulation (SIR, SIER , ...).

This document describes the mathematical methodology proposed to face this problem.



2. UNCERTAIN DYNAMIC SYSTEMS

To presents the methodologies for modeling Uncertainty Dynamic Systems, like COVID-19 pandemic, we defined a conceptual framework for discrete-time models of dynamic systems. The selected framework is common for three fundamental methodologies: i) Optimal Control; ii) Dynamic Programming, and iii) State Estimation. The two first methods are oriented to optimization and them are the results of the works of Lev Pontryagin (Pontryagin's Maximum Principle, Pontryagin et al. 1962) and Richard Bellman (Dynamic Programming) during the 1950s, after the contributions to Calculus of Variations by Edward J. McShane (1974). State estimation are methodologies oriented to reconstruct the history (smoothing and filtering) and to forecast the state variables of the system, considering the uncertainty and partial observation of the system in the time-space domain, the most famous application in state estimation if the Kalman Filter methodology (Kalman, 1960).

This common framework is based on:

- Two types of variables:
  - State variables: are the set of variables that are used to describe the mathematical "state" of a dynamical system, that describes, enough about, the system to determine its future behavior in the absence of any external forces that affect the system.
  - Control variables: are the set of variables associated to external forces that can change the "natural" behavior of the dynamic system
- A set of constraints that describes:
  - The dynamic relation between state variables on two consecutive periods (t and t+1)
  - The relations between control variables during each period t
  - The relation between state variables and control variables during each period t

A linear dynamic system can be described mathematically with the following differential equation:

$$\nabla_t \mathbf{X(t) = f(t)\ X(t) + b(t)\ U(t)}$$

without loss of generality, the above equation can be approximated by the following difference equation

$$\mathbf{X(t+1) = A(t)X(t) + B(t)U(t)}$$

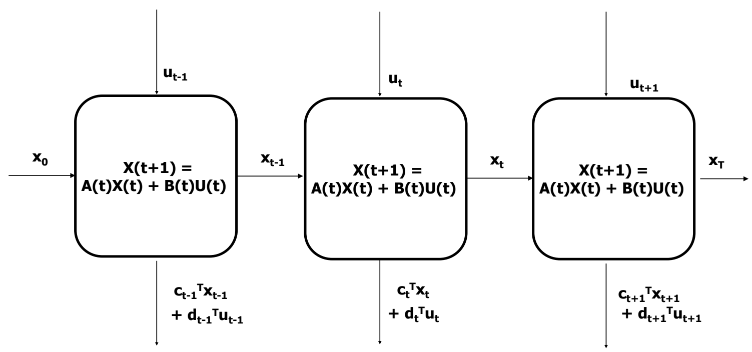
where the matrices **A(t)** and **B(t)** and **L(t)** are directly related to matrices **f(t)**, **b(t)** y **W(t)**.

Dynamic Programming and Control Theory includes the concept of optimization by linking a performance function associated with state variables and control variables, for example **R(X(t),U(t))**, for the linear case it is

$$\mathbf{R(X(t),U(t)) = c(t)^T\, x(t) + d(t)^T u(t)}$$

where **c(t)** represents the cost/revenue vector for **x(t)** and **d(t)** the cost/revenue vector for **u(t).**

DYNAMIC PROGRAMMING



The uncertainty may be incorporated into the model due to multiple reasons like unknown equations (inappropriate modeling), unknown parameters, etc. Then the above equations may be writing as

$$\nabla_t \mathbf{X(t) = f(t)\, X(t) + b(t)\, U(t) + W(t)}\ \varepsilon\mathbf{(t)}$$

or

$$\mathbf{X(t+1) = A(t)X(t) + B(t)U(t) + L(t)}\ \varepsilon\mathbf{(t)}$$

where **L(t)** are directly related to matrix **W(t)** and ε**(t)** represents an error (noise) element associate with each equation. The discrete version will be considered hereafter.

In State Estimation, to complete the system, an observation equation is included, that is

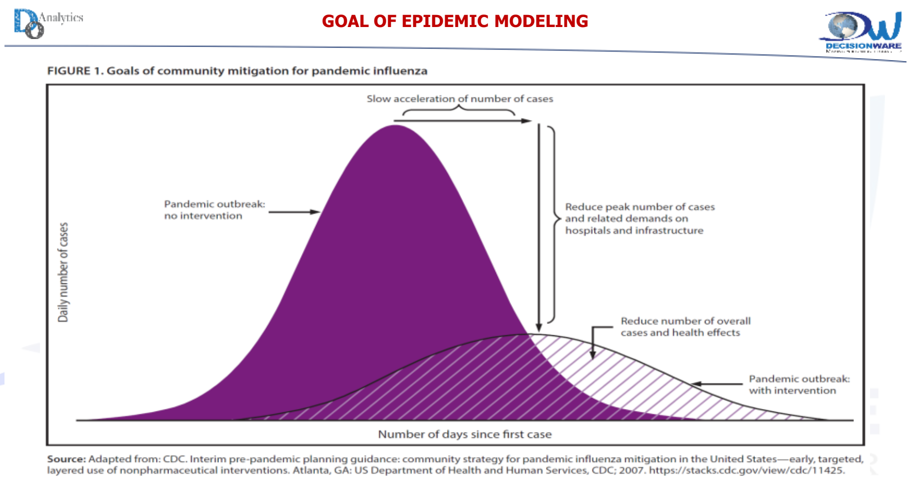
$$\mathbf{Z(t) = H(t)X(t) + \theta(t)}$$

where **Z(t)** represents an observation vector, **H(t) a** measurement matrix, and θ**(t)** noise in the measurement vector.

3. EPIDEMIC MODELING

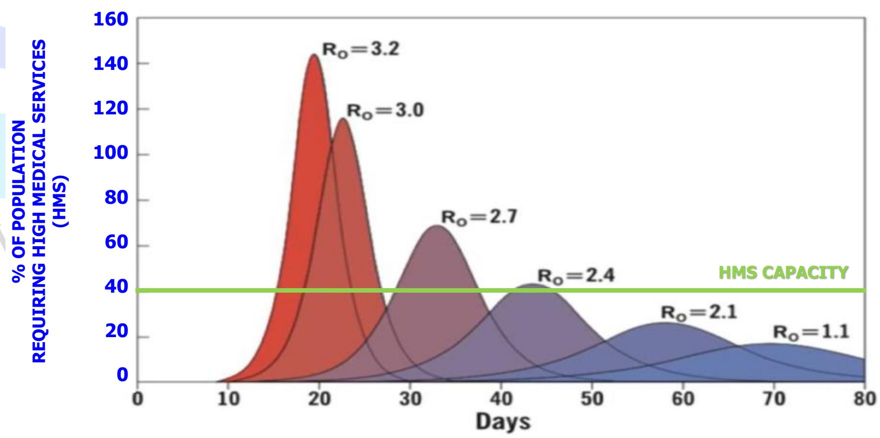
3.1. CONCEPTUAL FRAMEWORK

The goal of epidemic control strategies is to reduce R_0, which estimates the rate at which a disease can spread in a population. This can be achieved by reducing susceptibility or contact rates in the population or the infectiousness of infected populations. The potential effectiveness of medical intervention by varying the infectiousness of infected populations and nonmedical interventions by reducing the contact rates in the population have been examined. In medical intervention, use of vaccines and/or antiviral agents for case of treatment can increase the recovery rate and reduce the death rate. On the other hand, in nonmedical interventions, reducing population contact rates through social distancing and travel restrictions can reduce the impact on the transmission process.



Control of an outbreak relies partly on identification of the disease parameters that lead to a significant reduction of the basic reproduction number $\mathbf{R_0}$ that may be function of several parameters of which γ, the recovery rate for clinically ill and β, the transmission coefficient, are the most sensitive parameters. These two parameters can be controlled by medical intervention and nonmedical interventions.

EPIDEMIC DIFFUSION MODEL AS FUNCTION OF R_0



The modeling of epidemics in a solidly developed area of scientific knowledge, widely studied based on simulation models like

TRADITIONAL EPIDEMIC MODELS		
Model	**DESCRIPTION**	**Reference**
SIR	Susceptibility (S), Infection (I) and Recovery (R)	Kermack & Mc Kendrick (1927) Jing (2018)
SEIR	Susceptibility (S), Exposure (E), Infection (I) and Recovery (R)	Hethcote (2000)

TRADITIONAL EPIDEMIC MODELS		
Model	**DESCRIPTION**	**Reference**
SEIRA	Susceptibility (S), Exposure (E), Infection (I) and Recovery (R)	
SEIMR SEI3RD	Susceptibility (S), Exposure (E), 3+1 Infection States (I3), Recovery (R) and Death (D)	Grimm et al. (2020) Mejía Becerra et. al. (2020)
SEIQR	Susceptibility (S), Exposure (E), Infection (I), Quarantine (Q) and Recovery (R)	Huang (2016)
SIRS	Susceptibility (S), Infection (I), Recovery (R) and Susceptibility (S)	Cai (2015)

An epidemiological model is defined based on differential equations that explain the evolution of the process without human intervention. These differential equations can be established based on the population (number of people) who are in a certain "epidemic" state or based on the fraction of the population that is in that state. The epidemic models are nonlinear systems of ordinary differential equations, traditionally this equations system is solved using simulation models based in a discrete approximation for continuous derivatives, be it over time or space. There are many possible schemes. These models are used to analyze several widely discussed (predefined) scenarios and provide evidence on their effectiveness and are not oriented to get the optimal solution of a mix of control policies.

After analyzing the implementation of main (most known) epidemiological models (SIR, SEIR), it was decided to directly model discrete versions of differential equations as they maintain direct connection with biological parameters, which facilitates the connection of these parameters with socio-demographic segments.

Therefore, all epidemiological models considered should be formulated in one of the following terms.

- The time unit of the differential equations is one day.
- The states contain the fraction of the population in each state.
- The time of the optimization model may be divided in periods of multiple days (one week, seven days). In this case, the integration of the differential equations must be made using calculated parameters.

Examples of the epidemic states are showed in the next table. The models will be implemented using this nomenclature. The table includes the symbol used in the original models and the code used in the information system to reference the state.

EPIDEMIC STATES - TABLE: MAE_STA			
MODEL SYMBOL	**EPIDEMIC STATE**	**DESCRIPTION**	**COMMENTS**
S	SU	Susceptible Population	Those individuals who have not been exposed to the pathogen and are susceptible to being infected by it.
E	EX	Exposed Population	Those individuals who are in the latency state; that is, they have been inoculated by the pathogen but are not yet infectious
I	IN	Infected Population	In SIR and SEIR models is infected population. It must be the most critical state for infected people; this is important for models that have more than one epidemic states to describe the infection process.
I_0	I0	Asymptomatic Infectious	Those individuals in the population who have been inoculated by the virus are infectious but have not developed symptoms. Those infected in this state rarely learn that they have been infected.
I_1	I1	Moderate Symptoms Infectious	Those individuals in the population who are infectious and have mild or moderate symptoms. They are those who can be given management of the disease at home.
I_2	I2	Severe Symptoms Infectious	Those individuals in the population who are infectious and have severe but not critical symptoms. Individuals present in this state require hospitalization.
I_3	IN	Critical Symptoms Infectious	It must be the most critical state for infected people; this is important for models that have more than one epidemic states to describe the infection process. In SIR and SEIR models is infected population

EPIDEMIC STATES - TABLE: MAE_STA			
MODEL SYMBOL	**EPIDEMIC STATE**	**DESCRIPTION**	**COMMENTS**
R	RE	Recovered Population	Those individuals recover from infection, having developed antibodies. In most of the models they cannot be re-infected.
	ED	Epidemic Dead	Individuals who fail the infection and die.
	ND	Natural Dead	Individuals who die by other reason different to the epidemic
	NP	New Population	Individuals coming from an exogenous macro-region.

3.2. SIR: EPIDEMIC MODEL

The **SIR** model is the basic model in epidemic modeling (Kermack and Mc Kendrick, 1927). **SIR** process, starting with a susceptible host who becomes infected, the class of infection grow for the infected individuals to be able to transmit the infection to susceptible. When the infected individual is no longer able to transmit infection to susceptible individual, the infected individual is removed from the cycle of diseases transmission in the population. This model is based on the following assumption:

Then, the basic **SIR** model describes the epidemic with three states:

S Susceptible: initially covers all population that potentially can be infected **(SU)**

I Infected: Population that has been infected (IN)

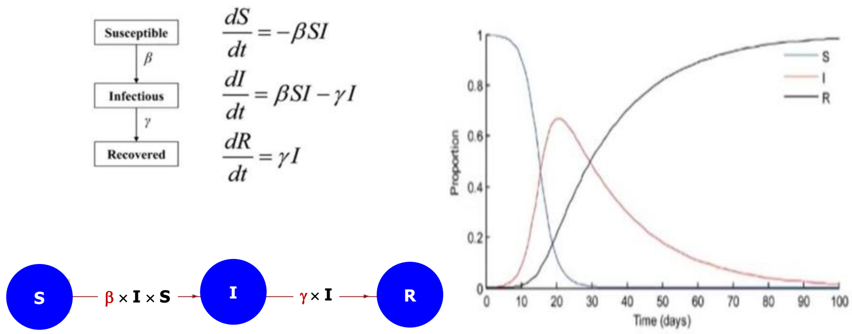
R Recovered: Recovering population (RE)

The diagram shows the behavior of **S(t)**, **I(t)**, and **R(t)** when they are normalized to total of population (**TPOB**) equal to 1. The biological parameters used in SIR and SEIR model are described below.

SIR MODEL - BIOLOGICAL PARAMETERS			
Parameter	**Description**	**Equation**	**Measure Unit**
δ	Contact Intensity – Exogenous Parameter		peo-day
ω	Probability of transmission per contact intensity (infectivity)		
γ	Recovery rate for clinically ill		fpo/day
μ	Epidemic death (mortality) rate		fpo/day
μ^N	Natural mortality rate		fpo/day
κ	The latency period of the virus before developing		day
ψ	Inverse virus latency period	$1/\kappa$	$1/\kappa$
β	Inverse contact intensity $\times$ infectivity	$\delta\delta \times \omega$	$\delta\delta \times \omega$
ρ	Relative removal rate	γ/β	
R_0	Basic reproduction ratio/number		

The diagram resumes the standard SIR epidemic model.

Susceptible-Infectious-Recovered (SIR)
Epidemic Model



SIR model is represented based on three differential equations based on proportions of people in each state (the ratio between the people in a state with the initial population **TPOB**). The measurements between parentheses.

$$\partial \mathbf{S(t)}/\partial \mathbf{t} \text{ (fpo/day)} = - \beta \text{ (1/fpo-day)} \times \mathbf{I(t)} \text{ (fpo)} \times \mathbf{S(t)} \text{ (fpo)}$$

$$\partial \mathbf{I(t)}/\partial \mathbf{t} \text{ (fpo/day)} = \beta \text{ (1/fpo-day)} \times \mathbf{I(t)} \text{ (fpo)} \times \mathbf{S(t)} \text{ (fpo)} - \gamma \text{ (fpo/day)} \times \mathbf{I(t)} \text{ (fpo)}$$

$$\partial \mathbf{R(t)}/\partial \mathbf{t} \text{ (fpo/day)} = \gamma \text{ (fpo/day)} \times \mathbf{I(t)} \text{ (fpo)}$$

where **S(t)**, **I(t)**, **R(t)** represent the population of susceptible, infected, and recovered individuals, respectively. Adding these equations, the following condition must be hold

$$\partial \mathbf{S(t)}/\partial \mathbf{t} + \partial \mathbf{I(t)}/\partial \mathbf{t} + \partial \mathbf{R(t)}/\partial \mathbf{t} = 0$$

Additionally, **SIR** can be extended with other epidemic states for a more complete description of the system/epidemic:

NP New population entering the system, as people from abroad who in many cases are the ones who cause the epidemic (NP).

ED People who die due to the epidemic (these people die regardless of the management of the epidemic (D).

ND People who die from natural death (N)

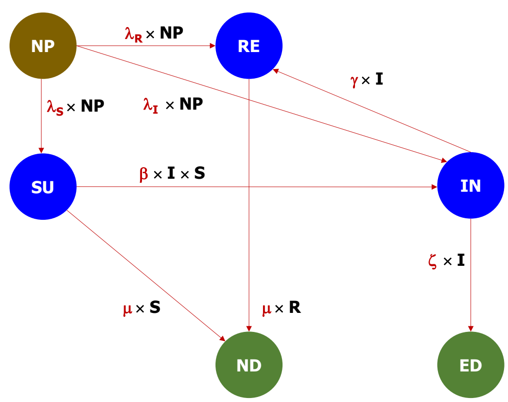
ND and **ED** states should be included if it wants to account for the resources consumed by people who die, who are killed due the epidemic and due by causes other than the epidemic.

For a more general formulation it is included the exogenous variable **NPX(t)** tha represents the proportion of people arriving from an exogenous system, may be births or people arriving from a foreign country/region. The value of **NPX(t)** is a border condition with the foreign system over any value of **t** it is calculated taking as reference the initial population **TPOB**. This adjustment may be important in regions, for example high people exchange rates islands dedicated to tourism. Next table shows the parameters used to this modeling.

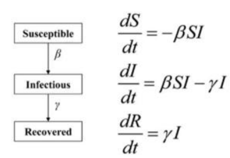
EXOGENOUS SYSTEM PARAMETERS		
Parameter	**Description**	**Measure Unit**
λ^E	Exposed rate coming from the exogenous system	fpo/day
λ^S	Susceptible rate coming from the exogenous system	fpo/day
λ^I	Infectious rate coming from the exogenous system	fpo/day
λ^R	Recovered rate coming from the exogenous system	fpo/day

Next diagram shows the epidemic system modeled.

STATE TRANSITION DIAGRAM - SIR MODEL



STATE	DESCRIPTION
NP	New Population
SU	Susceptive Population
IN	Infective Population
RE	Recovered Population
ND	Natural Dead
ED	Epidemic Dead



Then, the differential SIR equations must be adjusted:

$$\partial S(t)/\partial t = -\beta \times I(t) \times S(t) + \lambda^S \times NPX(t) - \mu^N \times S(t)$$

$$\partial I(t)/\partial t = \beta\,(I(t) \times S(t) - (\gamma + \zeta) \times I(t) + \lambda^I \times NPX(t)$$

$$\partial R(t)/\partial t = \gamma \times I(t) - \mu^N \times R(t) + \lambda^R \times NPX(t)$$

$$\partial D(t)/\partial t = \mu \times I(t)$$

$$\partial N(t)/\partial t = \mu^N \times S(t) + \mu^N \times R(t)$$

where μ^N represents the natural mortality rate and λ^{st} the rates coming from the exogenous system to the state st.

If **NPX(t)=0** the model meets the hypothesis that at all times

$$\mathbf{S(t) + I(t) + R(t) + D(t) + N(t) = 1}$$

If **NPX(t)** is different from zero the previous equation must be adjusted as

$$\mathbf{S(t) + I(t) + R(t) + D(t) + N(t) = 1 + \int_{q \in [0,t]} \partial NPX(q)\, \partial q}$$

To simulate the process the border conditions at the beginning of the simulation horizon are: **S(0), I(0), R(0), D(0)** and **NPX(t),** for all **t.**

Assuming **NPX(t)** equal to zero, the ratio $\rho = \gamma/\beta$ is called the relative removal rate. Thus, dynamics of infectious depends on the following ratio:

$$\mathbf{R_0} = \mathbf{S(0)} \times \gamma/\beta$$

where $\mathbf{R_0}$, called the basic reproduction ratio/number, is defined as the number of secondary infections produced by a single infectious individual during his/her entire infectious period. The role of the basic reproduction number is especially important. However, the following mathematical analysis describes how the basic reproduction number depends on the host population and the infected host.

At time $\mathbf{t = 0}$, $\partial \mathbf{I}/\partial \mathbf{t}$ can be written as

$$\partial \mathbf{I}/\partial \mathbf{t} = (\mathbf{R_0} - \mathbf{1}) \times \gamma \times \mathbf{I(0)}$$

if $\mathbf{R_0} > 1$ then $\partial \mathbf{I}/\partial \mathbf{t} > 0$ and therefore the disease can spread; but if $\mathbf{R_0} < 1$ then the disease dies out.

3.3. SEIMR/R-S GENERAL EPIDEMIC MODEL

SEIMR/R-S epidemic model (Velasquez et al., 2021) is supported on the basic equations of the **SEIMR** model including the effects of considering the development of the epidemic in a territory (macro-region) that includes multiple regions in which the population socio-demographic segments are distributed in a non-homogeneous manner.
http://www.doanalytics.net/Documents/DW-2-ITM-SEIMR-R-S-Epidemic-Model-Theory.pdf

SEIMR/R-S corresponds to a generalized mathematical model of pandemics that enhances traditional, aggregated simulation models when considering inter-regional impacts in a macro region (conurbed); SEIMR/R-S also considers the impact of modeling the population divided into socio-demographic segments based on age and economic stratum (it is possible to include other dimensions, for example: ethnics, sex, …). SEIMR/R-S epidemic model was carried out in a JAVA program. This program may be used by the organizations that considers the SEIMR/R-S will be useful for management the COVID-19 pandemic.

SEIMR/R-S is the core of the SEIMR/R-S/OPT epidemic management optimization model that determines optimal policies (mitigation and confinement) considering the spatial distribution of the population, segmented socio-demographically and multiple type of vaccines. The formulation of SEIMR/R-S/OPT is presented in PART III: SEIMR/R-S/OPT Epidemic Management Optimization Model (Velasquez-Bermudez 2021) describing its implementation in an optimization technology, like GAMS, AMPL, PYTHON, … .
http://www.doanalytics.net/Documents/DW-3-OPCHAIN-Health-Optimization-Epidemic-Model.pdf

SEIMR/R-S can be understood and used by any epidemiologist, and/or physician, working with SIR, SEIR or similar simulation models, and by professionals working on the issue of public policies for epidemic control.

The examples of the theory will be presented using the SIR model, but the implementation must be for the SEIMR/R-S epidemic model.

4. MATHEMATICAL METHODOLOGIES

This section presents a summary of methodologies related to technical aspects of this document.

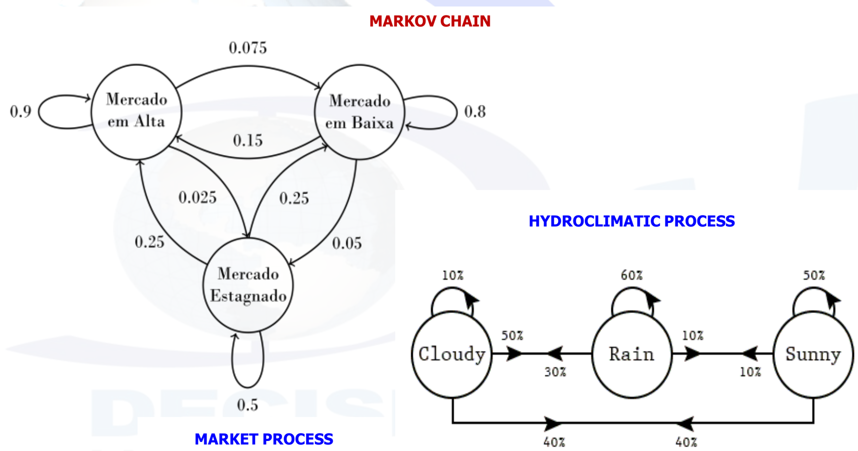
4.1. MARKOV CHAINS

4.1.1. FUNDAMENTALS

A process of Markov, named due to the work carried out by the Russian mathematician Andréi Márkov (1856-1922), is a dynamic random phenomenon for which the Markov property is true: it has no memory, which is a process for which the conditional probability on the present, the future and the past states of the system are independent of the path.

The practical importance of the Markov property is that it can be used to build statistical models of a stochastic process, in such a way to allow transit by a group of states in which the influence of going through a state declines to the lar the time go.

The term Markov chain is used to imply that a Markov process occurs in a discrete state space (infinite or countably). Usually, a Markov chain would be defined on a discrete set of periods (i.e., a Markov chain discrete time); although some authors use the same terminology where "time" and/or states are continuous values. Two simple examples of chains Markovianas are presented in the figure.



The chains in the figure are related to two different systems: a market of assets and a hydro-climate process; each of them can be modelled as a discrete State and discrete time, or as, continuous state and continuous time. The decision of as modeling takes place depends on the data, the specific problem, and the modeler.

Two types of Markov models can be considered:

- **Markovians:** the transition probabilities are supposed to be the constant over time, which can be expressed as

$$P(x_{t+1} = j \mid x_t = i) = \theta_{i,j}$$

  where x_t represents the State of the system during the period **t**, **P(A|B)** the conditional probability of **A** conditioned in **B** and $\theta_{i,j}$ the probability of a transition from state **i** to state **j**.

- **Semi-markovians:** the transition probabilities between states vary as to pass more periods; an example is the modeling of the life expectancy, the risk of death, whose probability increases with age.

$$P(x_{t+1} = j \mid x_t = i) = \theta_{i,j}(t)$$

It can say that reality is semi-Markovian, since almost all processes are dynamic random path, and how the system response a can evolve cyclically or trend over time

A Markov process required:

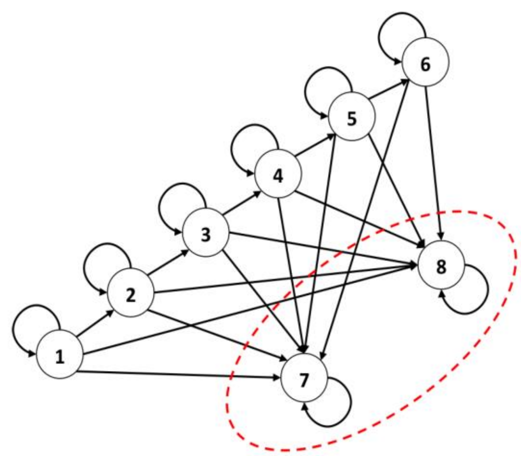
- **States:** The conditions in which the system can be found
- **Transition Probability:** The probability of moving from a current state **i** to state **j** in a transition, or period, which can be stationary or variant over time.

The diagram presents the elements of a Markov system.

STATE TRANSITION DIAGRAM FOR A MARKOV CHAIN MODEL

$$P = \begin{bmatrix} 0.6691 & 0.1588 & 0 & 0 & 0 & 0 & 0.0646 & 0.1075 \\ 0 & 0.5335 & 0.2542 & 0 & 0 & 0 & 0.1495 & 0.0627 \\ 0 & 0 & 0.6992 & 0.1110 & 0 & 0 & 0.1408 & 0.0490 \\ 0 & 0 & 0 & 0.8100 & 0.0800 & 0 & 0.1000 & 0.0100 \\ 0 & 0 & 0 & 0 & 0.8150 & 0.0499 & 0.1282 & 0.0069 \\ 0 & 0 & 0 & 0 & 0 & 0.8380 & 0.1554 & 0.0066 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.0000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0000 \end{bmatrix}.$$

STATE TRANSITION MATRIX



The previous two features should be identified for modeling a process such as Markov: the states, which in many cases are not readily apparent and may come from a study of segmentation that can be done with methodologies of **ML**; subsequent to the definition of the states, required a sample that permits to estimate transitions between states, the process will be more complicated if it considers dynamic transition matrices.

One of the current uses of Markov theory is to represent the behavior of persons in their relations with an organization. But it can also represent the different states in which a system is located, or an element type within a system.

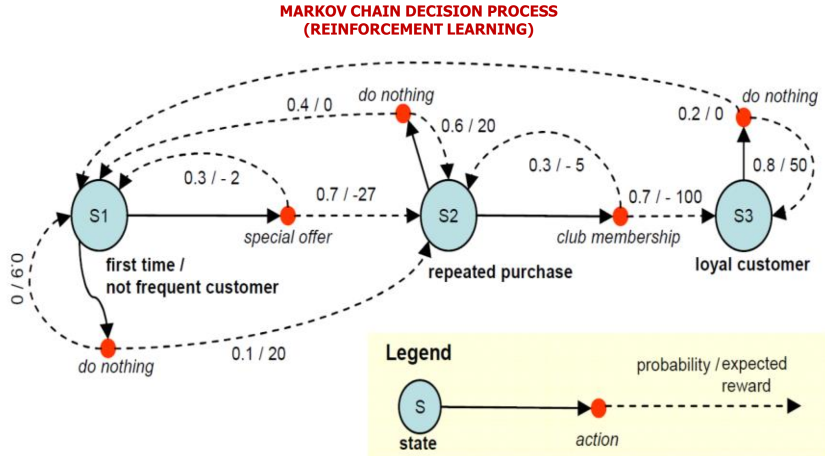
4.1.2. MARKOV CHAIN DECISION PROCESS

Decision-making processes based on a Markov (Markov Chain Decision Process, **MCDP**), today also known as Reinforcement Learning, are the goal of the mathematical modeling of the process of decision-making.

These processes include the actions trying to affect the stochastic process, understood as the path through the system states. A clear application of these methodologies may be associated with the decisions of companies to try to manage the lifecycle of their customers. To achieve this included:

1. The concept of transition probabilities conditional decisions
2. The remuneration/gain/loss to visit a specific state.

For a commercial enterprise, the following diagram describes a MCDP which are:
1. Three states for a client: **S1** (first time), **S2** (repeated purchase) y **S3** (loyal customer).
2. The actions of control of the decision-maker are defined for each state:
    - S1: **special offer** or **nothing**
    - S2: **club membership** or **nothing**
    - S3: **nothing**
3. Each decision has different transition probability and remuneration.



The objective of the modeling is to implement a Markov model that maximizes the profits, in the short/long term, of the decision-maker. Decisions are made in such a way that when the client is coming into a state **i** a decision **D(i)** is made.

4.2. BAYESIAN INFERENCE

The Bayesian inference is a method based on Bayes theorem which is used to update the probability of a hypothesis given the information that is acquired when more evidence or information is available. Bayesian inference has found application in a wide range of activities, including science, engineering, philosophy, medicine, sport, and law.

The Bayes theorem solves the problem known as "probability inverse", that is rating probabilistically possible conditions that govern the event which has been observed. The followers of Bayesian inference say that the significance of the inverse probability lies in that it is that really matters to science, given that it seeks to draw general conclusions (state laws) from the objectively observed, and not vice versa

The Bayes theorem expresses the conditional probability of a random event **B** given in terms of the distribution of conditional probability of event **B** given **S** and the marginal probability distribution of **S**. It is mathematically formulated as:

Considering $\{S_1, S_2, S_3, \ldots S_T\}$ a set of events/states, mutually exclusive and collectively exhaustive, such that the probability of each S_i is non-zero **(0)** and **B** one event either that the conditional probability given an **Si** event are known, $P(B|S_i)$; then, the conditional probability of an S_i event given the event **B**, $P(S_i|B)$, is given by the expression:

$$P(S_i|B) = P(S_i) \times P(B|S_i) / P(B)$$

where
P(S_i) Probability a priori of **S_i**
P(B|S_i) Probability a priori of **B** given **S_i**
P(S_i|B) Probability a posteriori of **S_i** given **B**
P(B) Probability of **B**

The Bayes theorem can be written as (where **P(x)** represents the probability function of **x**):

$$P(Model|Data) = P(Model) \times [\ P(Data|Model) / P(Data)\ ]$$

or

Probability a Posteriori =
Probability a Priori × Likelihood Function (Data) / Probability (Data)

If a process is being observed through multiple (N) models potentially valid for state identification, and/or parameters, the following equation is met for each model n:

$$\Omega_t(Model_n|z(t)) = \Omega_{t-1}(Model_n|z(t-1)) \times \Phi(z(t)|Model_n) / \pi(z(t))$$

where $\pi(z(t))$ represents the probability function of measurement z(t), $\Phi(z(t)|Model_n)$ the likelihood function of z(t) given the model **n** as the true model, and $\Omega_t(Model_n|z(t))$ the probability of that the model **n** is the true model when the observations until time **t** have been processed.

If one of the N models is considered to be true, then the sum of the probabilities should be equal to 1, allowing the value of $\pi(z(t))$ to be calculated as a normalization constant common for all probabilities.

$$\pi(z(t)) = \Sigma_{n=1,N}\ \Omega_t(Model_n|z(t))$$

The Bayesian approach is particularly important in the dynamic analysis of a data streams, and by itself implies a process of deep learning from data.

5. STATE ESTIMATION

In Wikipedia, in control theory, a state estimator is a system that provides an estimate of the internal state of a given real system, from measurements of the input and output of the real system. It is typically computer-implemented and provides the basis of many practical applications.

Knowing the system state is necessary to solve many control theory problems; for example, stabilizing a system using state feedback. In most practical cases, the physical state of the system cannot be determined by direct observation. Instead, indirect effects of the internal state are observed by way of the system outputs. A simple example is that of vehicles in a tunnel: the rates and velocities at which vehicles enter and leave the tunnel can be observed directly, but the exact state inside the tunnel can only be estimated. If a system is observable, it is possible to fully reconstruct the system state from its output measurements using the state observer.
https://en.wikipedia.org/wiki/State_observer

The application for estimating the true state of a pandemic is direct.

5.1. FRAMEWORK

The approach for the State Estimation (**SE**) is supported on a conception of the stochastic processes where the random variables are clearly differentiated of their measures, being understood that the modeler has data of the measurement system that may be different from real values of the random variables. This differentiation is essential to understand the operation of the systems and systems used for observation.

Therefore, all available observations are considered as random variables, having their measure information, but which, by reason of the randomness of the measuring system, are subject of precision errors; therefore, assume them as certain (deterministic) induces errors; moreover, there might be more than one measure for a random variable, obtained by different systems of measurement.

Two types of noise (errors) are considered in state estimation modeling:

- **ε(t)**, errors due to the modeling of the system, reflecting the uncertainty of knowledge of functions (equations and/or parameters) that determine the behavior of the system that is modeling; and
- **θ(t)**, measurement errors that come from the precision with which the random variables associated with the system are measured.

From this point of view, the conventional statistical models only consider one type of error, **μ(t)**, which integrates modeling errors and measurement, whereas the state estimation modeling differs clearly the two types of errors. This may be the principal advantage of state estimation modeling over the traditional statistical modeling.

Another advantage is the possibility of make hypothesis about the dynamics of variation of the parameters of the mathematical models; from this point of view the following aspects must be considered:

- The estimators of the parameters of a system are random variables dependent on the observations and should therefore be subject to adjustment in accordance with the new information is arriving.
- A system parameters vary over time, since many of them change as a result the evolution, that may be technical, physical, economic, social and/or natural.

5.2. MATHEMATICAL FORMULATION

It should be noted, for the ease of the reader, that the notation **X(t)**, capital letter **X**, refers to the random variable and **$x(t_1|t)$** refers to the estimate of **$X(t_1)$** given the information received up to the time **t**.

Consider a system whose state at any moment of time **t** is synthesized based on the value of a set of representative variables grouped in the vector of State **X(t)**. Exogenous actions to the system are made through a set of control variables grouped into vector **U(t)**. A system of differential equations that represent the inter-temporal dynamics of the system can be formulated based on these definitions:

$$\nabla_t X(t) = F_t [X(t), U(t)]$$

where **$F_t[X(t),U(t)]$** is a vectorial function. The previous system of equations has as boundary condition the initial state, **X(0)**. This is:

$$\nabla_t X(t) = F_t[X(t), U(t); X(0)]$$

The function **$F_t[X(t),U(t)]$** represents, accurate or approximately, the dynamics of the system. In the case of an approximation should be considered errors of the modeling of the system by the vector of noise in the modeling **ε(t)**. Keep in mind that if there are no errors in the system modeling, the problem of estimation of state does not make sense since the vector function **F_t[X(t),U(t)]** describes exactly the trajectory of the system. Under this consideration, the dynamic equation becomes a stochastic differential equation and is reformulated as

$$\nabla_t X(t) = F_t[X(t), U(t), \varepsilon(t); X(0)]$$

The state variables **X(t)** are measured through a set of variables that are grouped in the observation vector **Z(t)**. The relationship between the state variables and the observation variables are represented as:

$$Z(t) = H_t[X(t)]$$

where **H_t[X]** is a vectorial function.

If we consider that the observation system is not perfect, either by modeling represented in **H_t[X]** or for the accuracy of the measurements, it is necessary to include errors in the equation represented by the vector **θ(t)** of noise in observation. Then the equation is

$$Z(t) = H_t[X(t), \theta(t)].$$

The purpose of the state estimation modeling, as its name implies, is to make estimates **x(t)** of the state vector **X(t)** from the vector of observations **Z(t)** and from the vectorial function **F_t[X(t), U(t), ε(t); X(0)]** that represents, approximately, the dynamic of the system.

6. KALMAN FILTER

Below, the mathematical basics of state estimation modeling using the theory developed by R.E. Kalman and Bucy R.S., commonly known as the Kalman Filter (KF, Kalman filter), are presented following the Wikipedia information.

In statistics and control theory, Kalman filtering, also known as Linear Quadratic Estimation (LQE), is an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone, by estimating a joint probability distribution over the variables for each timeframe. The filter is named after Rudolf E. Kalman, one of the primary developers of its theory.

KF has numerous applications in technology and engineering. Initially the real-life applications were for guidance, navigation, and control of vehicles, particularly aircraft, spacecraft, and dynamically positioned ships. Furthermore, the KF is a widely applied concept in time series analysis used in fields such as signal processing and econometrics. KF also are one of the main topics in the field of robotic motion planning and control and can be used in trajectory optimization. The KF also works for modeling the central nervous system's control of movement. Due to the time delay between issuing motor commands and receiving sensory feedback, use of the KF supports a realistic model for making estimates of the current state of the motor system and issuing updated commands. Other field of KF applications is to forecast hydroclimatic variables.

6.1. MATHEMATICAL FORMULATION

We consider linear vectorial functions **F_t[.]** and **H_t[.]**:

1. State Equation:

$$\nabla_t X(t) = f(t)\, X(t) + b(t)\, U(t) + W(t)\, \varepsilon(t)$$

where

X(t) system state vector (n,1)
f(t) transition matrix (n,n)
U(t) control vector (k,1)
b(t) control interaction matrix (n,k)
ε(t) vector of noise in modeling (p,1)
W(t) interaction of noise matrix (n,p).

Without loss of generality, the above equation can be approximated by means of a difference equation

$$X(t+1) = A(t)X(t) + B(t)U(t) + L(t)\varepsilon(t)$$

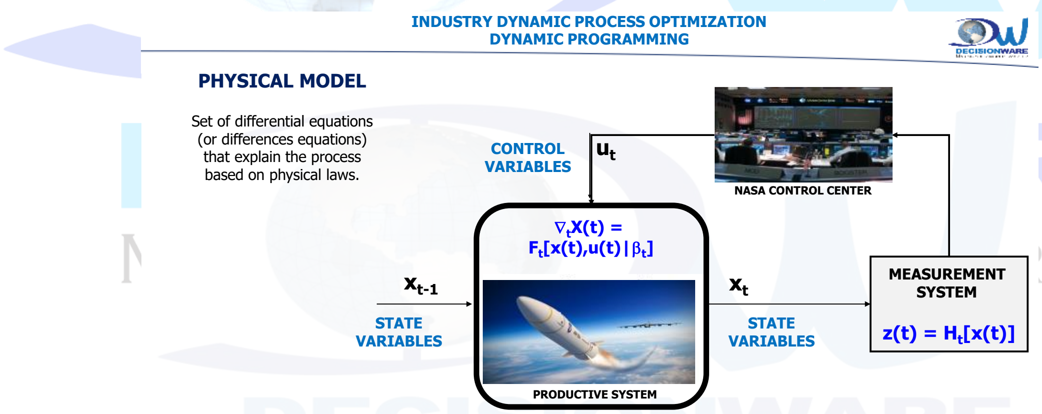
where the matrices **A(t)**, **B(t)** and **L(t)** are directly related to matrices **f(t)**, **b(t)** y **W(t)**. The discrete version will be considered hereafter.

2. Observation Equation:

$$Z(t) = H(t)X(t) + \theta(t)$$

where

Z(t) observation vector (m,1),
H(t) measurement matrix (m,n),
θ(t) noise in the measurement vector (m,1).



6.2. STATISTICAL HYPOTHESIS

The development of the theory of the Kalman filter assumes the following statistical assumptions:

1. Initial state of the system **X(0)**:
   - Expected value:

$$x(0) = E[X(0)]$$

- Variance-covariance matrix:

$$\Sigma(0) = E[\{X(0)-x(0)\}\ \{X(0)-x(0)\}^T]$$

2. Noise of dynamic modeling system ε**(t)**:
   - Expected value:

$$E[\varepsilon(t)] = 0$$

   - Variance-covariance matrix:

$$E[\varepsilon(t)\varepsilon(t)^T] = Q(t)$$

   where **Q(t)** is a diagonal matrix.

3. Noise of observation system θ**(t)**:
   - Expected value:

$$E[\theta(t)] = 0$$

   - Variance-covariance matrix:

$$E[\theta(t)\theta(t)^T] = R(t)$$

4. Control Vector **U(t)**: deterministic.

5. **A(t)**, **B(t)**, **L(t)** and **H(t)**: with deterministic components.

Under the above assumptions the estimate process developed by Kalman [5] guarantees estimators **x(t)** of the average of the vector **X(t)** that minimizes the trace of the matrix of variance-covariance Σ**(t)**

$$\Sigma(t) = E[\{X(t)-x(t)\}\ \{X(t)-x(t)\}^T]$$

If we also consider that the random components of the system, **X(0)**, ε**(t)** and θ**(t)** follow Gaussian probability distribution functions, we will have the following features for **x(t)**:

- Best linear unbiased estimators (**BLUE**),
- Maximum likelihood estimators,
- Bayesian estimators a posteriori.

For the variance-covariance matrices should be noted: i) **R(t)** can be estimated through measurements; while ii) **Q(t)** is more difficult to estimate, given that it has no access to the state **X(t)**.

6.3. ESTIMATION PROCESS

The estimation procedure proposed by Kalman provides the best Bayesian estimator of the expected value of **X(t)** conditionate on:

1. The information processed up to time **t**

$$U(t_p) = \{\ U(1), U(2), \ldots U(t_p)\ \}$$
$$Z(t_p) = \{\ Z(1), Z(2), \ldots Z(t_p)\ \}$$

   information that it synthetized as $\mathbf{I(t_p)}$

$$I(t_p) = \{Z(t_p), U(t_p)\}$$

2. The initial estimates (**t=0**) of the state vector and the variance-covariance matrix**, x(0)** and Σ**(0)**.

The estimate of **X(t)** given the information to instant t_p is denoted as:

$$x(t/t_p) = E[X(t) \mid Z(t_p), u(t_p), x(0), \Sigma(0)]$$

Similarly, the estimator of the variance-covariance matrix is

$$\Sigma(t/t_p) = E[\{X(t)-x(t/t_p)\} \{X(t)-x(t/t_p)\}^T]$$

the a priori estimates for the system are defined as:

$$x(0/0) = x(0)$$
$$\Sigma(0/0) = \Sigma(0).$$

The Kalman filter defines a sequential process of estimation from the combination of Bayesian priori information up to time **t-1** with the information subsequently obtained at the moment **t.** This process is summarized in six steps:

1. Estimate a priori of the variance-covariance matrix Σ**(t):**

$$\Sigma(t/t-1) = A(t-1)\Sigma(t-1/t-1)\, A^T(t-1) + L(t-1)Q(t-1)L(t-1)$$

2. Estimation ex-post of Σ**(t):**

$$\Sigma(t/t) = \Sigma(t/t-1)[I - H^T(t)\{H(t)\Sigma(t/t-1)H^T(t) + R(t)\}^{-1}H(t)]\Sigma(t/t-1)$$

3. Calculation of the so-called gain matrix **M(t):**

$$M(t) = \Sigma(t/t)H^T(t)R(t)^{-1}$$

4. A priori estimation of the expected value of the state vector **X(t):**

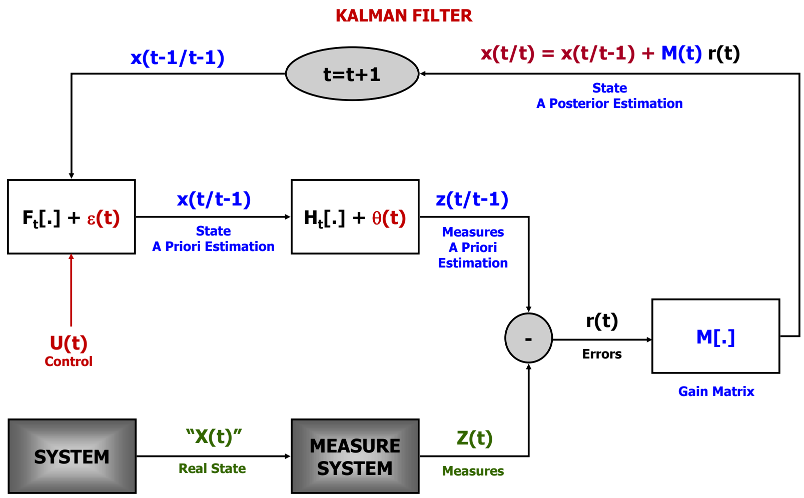
$$x(t/t-1) = A(t-1)x(t-1/t-1) + B(t-1)U(t-1)$$

5. Estimation of the prediction residuals **r(t):**

$$r(t) = Z(t) - H^T(t)x(t-1/t-1)$$

6. A posteriori estimation of the expected value of **X(t):**

$$x(t/t) = x(t/t-1) + M(t)r(t)$$



In the previous process, steps 1, 2, and 3 the calculations for the estimators of the variance-covariance matrix-related **Σ(t/tp)** and for the determination of the gain matrix **M(t)** can be made prior to the obtaining of the observations and are so-called "off-line" calculations. Steps 4, 5 and 6 are dependent on the values of **Z(t)** and correspond to the so-called process of "on-line". This feature is related to the hypothesis of linearity of equations.

By way of synthesis, we will define the set of estimators **Θ(t/t)** as

$$\Theta(t/t) = \{x(t/t), \Sigma(t/t)\} = BAYES[\Theta(t/t-1), Z(t), U(t), Q(t), R(t)]$$

where the operator **BAYES[.]** represents the combination of Bayesian information proposed by Kalman.

6.4. SMOOTHING, FILTERING & FORECASTING

The stochastic processes where the random variables are clearly differentiated of them measures allows the modeler to improve knowledge of the past based on the evidence of the present. It can only be realized by a combination of forward and backward processes, which is known as smoothing. The optimal causal solution is also known as the Wiener filter.

If **X(t)**, capital letter **X**, refers to the random variable and **x(t_1|t)** refers to the estimate of **X(t_1)** given the information received up to the time **t**. The relationship between **t_1** and **t** defines several types of process:

- **Smoothing (t_1<t):** It involves predicting a posteriori the state variable in the past **t_1**, before **t**.
- **Filtering (t_1=t):** It involves the correction (filter) of the prediction of the state variables in **t**, while reaches the information.
- **Forecasting (t_1>t):** It involves the prediction of the State variables for periods **t_1** after the present **t**.

The last two functions are the most used; however, the prediction of the past makes perfect sense when we recognize that in the past we only had estimates of the state variable (we never know the real value), which can be improved by processing the information coming later to the prediction made.

Detailed information about the mathematical formulation of smoothing process is founded in Einicke ().

Smoothing is of fundamental importance in the process of estimating the state of the epidemic, because due to the type of process, it is necessary to reconstruct, in the best possible way, the history of the epidemic since an inappropriate interpretation will give rise to bad decisions since they are made based on measures that do not reflect the reality that is happening.

6.5. STATE ESTIMATION THROUGH OPTIMIZATION

The a priori estimators **x(t/t-1)** can be maximizing the a priori likelihood function of **X(t)** that, in the case of a multivariate normal distribution functions, can be written as

$$\{2\pi|Q(t)|\}^{-n/2}\ \mathrm{Exp}[w(t)^T V(t-1)^{-1} w(t)]$$

where

$$w(t) = X(t) - A(t-1)x(t-1/t-1) - B(t-1)U(t-1)$$

$$V(t) = A(t)\Sigma(t/t)A(t)^T + L(t)Q(t)L(t)$$

The estimators **x(t/t-1)** are obtained maximizing the argument of the exponential function, which can be written as

$$\begin{gathered} X(t)^T V(t)^{-1} X(t) \\ - 2\{A(t-1)x(t-1/t-1) - B(t-1)U(t-1)\}^T V(t)^{-1} X(t) \\ -\{A(t-1)x(t-1/t-1) - B(t-1)U(t-1)\}^T V(t)^{-1} \\ \{A(t-1)x(t-1/t-1) - B(t-1)U(t-1)\} \end{gathered}$$

that is a quadratic function for **X(t)**, where the last term is constant.
The estimators **x(t/t)** can be interpreted as maximum likelihood estimators for **X(t)**. Based on the Bayes theorem, the probability of **X(t)** given the information obtained up to time **t**, **{Z(t),Z(t-1),..}**, can be written as:

$$\begin{gathered} \mathrm{Probability}[X(t)|Z(t),Z(t-1),\ldots] \\ = \mathrm{Probability}[Z(t)|X(t),Z(t-1),\ldots] \\ \mathrm{Probability}[X(t)|Z(t-1),Z(t-2),\ldots]/\mathrm{Probability}[Z(t)] \end{gathered}$$

The likelihood function is

$$\begin{gathered} \mathrm{Probability}[X(t)|Z(t), Z(t-1), Z(t-2), \ldots] \\ = (2\pi)^{-n-m/2}\ |R(t)|^{-m/2}\ |\Sigma(t/t-1)|^{-n/2} \\ \mathrm{Exp}[\ \{Z(t)-H(t)X(t)\}^T R(t)^{-1}\{Z(t)-H(t)X(t)\} \\ - \{X(t)-x(t/t-1)\}^T \Sigma(t/t-1)^{-1}\{X(t)-x(t/t-1)\}\ ] \end{gathered}$$

The estimators **x(t/t)** are obtained by maximizing **Probability[X(t)|Z(t),Z(t-1),Z(t-2),...]** with respect to **X(t)**, what is achieved to maximize the argument of the exponential function which can be written as

$$\begin{gathered} X(t)^T\{H(t)^T R(t)^{-1} H(t) + \Sigma(t/t-1)^{-1}\}X(t) \\ - 2\{Z(t)^T R(t)^{-1} H(t) + x(t/t-1)^T \Sigma(t/t-1)^{-1}\}X(t) \\ + Z(t)^T R(t)^{-1} Z(t) + x(t/t-1)^T \Sigma(t/t-1)^{-1} x(t/t-1) \end{gathered}$$

that is a quadratic function for **X(t)**, in which the latter two terms are consistent.

The vision of the process from the point of view of optimization allows to extend the concepts of **KF** to face two new types of problems:

1. Assume probability distributions different than normal multivariate; in this case the likelihood function must be replaced in the optimization model for those corresponds according to the assumed for probability distribution multivariate function.

2. Another extension of KF is related to the estimation of state variables that must meet a set of conditions that determine the feasible region in which there are valid solutions for the state vector. In this case the estimation process should be adjusted in such a way to have estimators feasible and consistent. The restrictions for the vector of state variables may be expressed as:

$$\mathbf{G[X(t),U(t)] = 0}$$

   where is **G[.]** a vector function. The optimization problem is

Maximizar

$$\mathbf{X(t)^{T}V(t-1)^{-1}X(t)}$$
$$\mathbf{- 2\{A(t-1)x(t-1/t-1)-B(t-1)U(t-1)\}^{T}V(t-1)^{-1}X(t)}$$

sujeto a:

$$\mathbf{G[X(t),U(t)] = 0}$$

The structure of the constraints determines the complexity of the optimization problems. In the case of linear restrictions, the problems are of quadratic programming.

From the statistical point of view, the introduction of restrictions alters the calculation a priori and a posteriori of estimates of **X(t)**; the calculation of estimates of the state vector variance-covariance matrices is also altered.

The problem is simplified if considered **x(t/t)** and **x(t/t-1)** as a result of a process of constrained weighted least squares, where the variance-covariance matrices correspond to matrices of weighting that implicitly represent the credibility that has the a priori information. The estimators obtained in this case are consistent, comply with the restrictions, but are not **BLUE** estimators, because the predictions have bias.

6.6. VARIATIONS AND EXTENSIONS OF KALMAN FILTER

An advanced predictive analytics computing platform that allows the results of an industrial/natural process (described based on partial differential equations) to be projected in advance from real-time input variables (measurements) can be based on the combination of the methodological principles of Machine Learning (ML) and State Estimation (SE), is a means of implementing machine learning processes (cognitive robots) to understand the behavior of an industrial/physical system using advanced algorithms, based on the fundamental concepts that support the Kalman Filter coupled with Bayesian modeling of dynamic systems. This process is called ML-KF.

Specifically, an effective algorithm for early identifying the results of an industrial/natural process based on real-time input variables should consider a combination of the following methodologies:

i) Kalman Filter (KF, Kalman, 1960)
ii) Extended Kalman Filter (EKF, Julier, S. J. and. Uhlmann, 2004)
iii) Dual Kalman Filter (D-KF, Moradkhani et al, 2005)

iv) Bayesian Model Combination (BMC, Rodríguez-Iturbe et al. 1978)
v) Multiple State Kalman Filter (MS-KF, Velásquez, 1978)

The mixture of the above methodologies allows to identify:

i) KF: The value of system state variables when the structure and parameters of differential equations governing system behavior are known for certain. It corresponds to the basic theory formulated by R. E. Kalman, which is limited to systems of linear equations.
ii) EKF: The value of system state variables when the structure and parameters of differential equations are known for certain. Extend Kalman's theory to systems of nonlinear equations.
iii) D-KF: The response functions (parameters of the differential equations) identified for each of the possible system states. This approach is useful when the parameters of differential equations are not precisely known.
iv) BCM: The selection of the model (system of differential equations) that best represents how the system responds. This approach is useful when neither the structure nor the parameters of differential equations are known for certain.
v) MS-KF: Possible Markovian states in which a system/process can be (e.g., stable, stochastic transient, or structural transient) involving different forms of system response. The selection of the probability of the state in which the system is located is determined based on a Bayesian inference model. This approach allows to detect disruptive changes in system behavior in order to adjust response functions according to information obtained from real-time data.

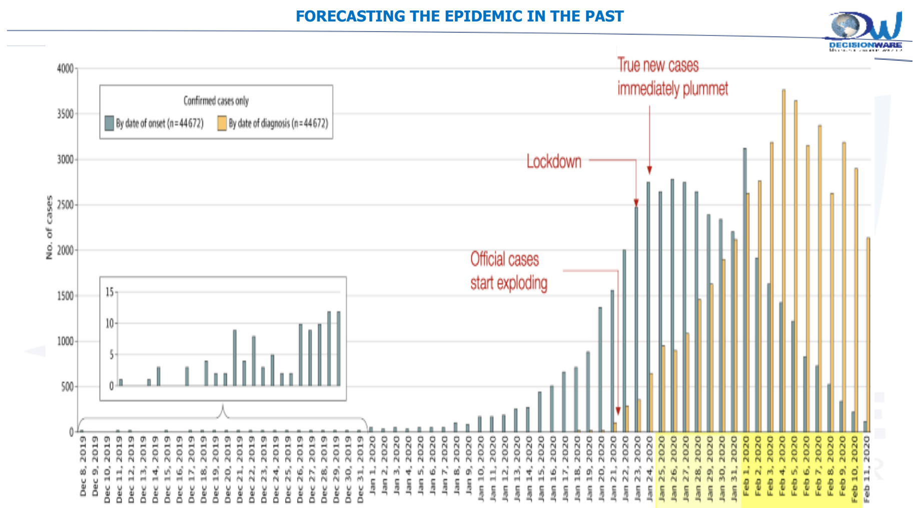
The ML-KF associated learning process is of the Deep Learning type, as it assumes system dynamics as part of the parameter identification process, without pre-establishing a stationary and certain response function. The dynamics of the MS-KF/D-KF allow the system response function to be fully re-estimated considering the information contained in the latest (real-time) data.

6.7. MODELING EPIDEMIC PROCESS

Kalman's methodology can be used for estimating the state of the pandemic, understood as the number of people (or fraction of the population) found in each epidemiological state of the differential equation model selected to describe the pandemic.

Smoothing is fundamental in the process of estimating the state of the epidemic, because due to the type of process, it is necessary to reconstruct, in the best possible way, the history of the epidemic since an inappropriate interpretation will give rise to bad decisions since they are made based on measures that do not reflect the reality that is happening. In the case of the pandemic, “predicting the past” can be much more important than “predicting the future”.

This is easy to visualize when you look at the comparison between known cases and the actual cases reported in Hubei at the beginning of the COVID-19 pandemic.



Applying Kalman's filter to the pandemic can be done in accordance with the following procedure:

1. State variables, X(t), are associated with epidemiological model states. For the case of the SIR model these are: susceptible (S), infected (I), recovered (R), deaths due to infection (D) and deaths by other causes (N).

2. The control variables, U(t), correspond to the social policies of population confinement that are applied to control the epidemic. They depend on each case. In the Colombian case they depend on temporary measures taken by local governments which vary in details of the form.

3. The basic differential equations are those corresponding to the selected model, in the case of the SIR are:

$$\partial S(t)/\partial t = - \beta \times I(t) \times S(t) - \mu^N \times S(t)$$

$$\partial I(t)/\partial t = \beta \, (I(t) \times S(t) - (\gamma + \zeta) \times I(t)$$

$$\partial R(t)/\partial t = \gamma \times I(t) - \mu^N \times R(t)$$

$$\partial D(t)/\partial t = \mu \times I(t)$$

$$\partial N(t)/\partial t = \mu^N \times S(t) + \mu^N \times R(t)$$

4. Basic differential equations must be adjusted to include the effects of the control in the dynamic process. This process is not simple, as it should be carefully analyzed how control policies relate to the epidemic. For the present case, reference is made to the modelling reported by the Mayor of Bogotá, which is expressed as:

5. The parameters of the differential equation model, including control variables, are:

6. The observation system should be defined based on the information reported in the databases that follow the epidemic (if more than one database exists, it is possible to consider the two sources of information). Since there is no universal data model for reporting epidemic observation, each case should be analyzed separately. Below is the information reported by the mayor of Bogotá.

The KF hypotheses for pandemic observation (surveillance) are discussed below. KF assumes that:

1. The control variables are known and deterministic. These limitation may be managed adjusting the modeling, considering de control variables as new state variables.

2. The dynamic equations are known. These is a big limitation because one of the problems of the epidemic modeling is that the true equation is unknown.

3. The parameters of the dynamic equations are known and deterministic. This is another big limitation for the case of an epidemic process.

4. Co-variance arrays, **Q(t)** and **R(t)** ,related to model errors are deterministic, which is not actually true, as the true value and variation dynamics of the array elements are unknown.

KF variations and improvements, that are alternative to addressing the limitations and issues that have been referenced, will be discussed below.

7. DYNAMIC SYSTEM IDENTIFICATION

The main idea of **KF** is oriented to estimate a set of physical variables that summarize the state in which the system is located. To determine the evolution of the system is available:
i) Control Variables (the decision-maker actions) and
ii) Dynamic Equations (discrete or continuous) for determining the expected path system.

With respect to the dynamic equations, two cases must be considered:
i) They are known with certainty, or
ii) They must be identified during the process of learning about the knowledge (dynamics) of the system.

Two uncertain cases should be considered in: (i) the parameters of the equations, and (ii) the structure of the equations representing the dynamics of the system. This section is concentrating on the latter case, this means using the **KF** to estimate the parameters of an algebraic statistic model.

In the specific case of pandemic modeling, this has to do with the following parameters:

Simultaneous estimation of state variables and statistical parameters seeks to solve the problem of sub-optimization that is generated when the problem is faced in isolation, either by estimating the value of the variables assuming the parameters of the models as deterministic, or when the statistical parameters are estimated assuming the variables, physical or technical-economic, as deterministic.

The study of the selection of differential equations governing system behavior faces the section called Bayesian Selection of Dynamic Equations

7.1. PARAMETER ESTIMATION OF STATISTICAL MODEL

Then, should be noted that, paradoxically, for the estimation of the parameters of a model it is necessary to look at them as state variables and associate them with estimates that have the model parameters at any given time.

A first approximation is to assume that at the initial moment (**t=0**) the response of the system function parameters has been identified and estimated exogenously, it is considering the statistical model as deterministic components, it is strong enough since the parameters come from models in which have been considered as random variables. This initial estimate may correspond to a model probabilistic of the system under study.

Considering that the time **t=0** is only for reference, it is expected that the parameters of a probabilistic response function continue to be random variables after the initial moment, and therefore should continue considering that in accordance with the receipt of new information in new periods **t>0**.

Then, it discusses the use of the approach for the estimation for the identification of models of systems (technical, physical, economic, social,...) that include the estimation of model parameter. For purposes of this analysis the endogenous and exogenous variables will be considered as deterministic, these variables refer to the sample of variables included in a statistical modeling.

This approach allows to assume dynamic structures for the variation of the parameters, which leads to more flexible models that the obtained making use of limited tools that offer classical statistics-based methods. It must be also understood that state estimation process corresponds to a Bayesian Inference process the estimators are recalculated when new information arrived.

From this point of view, the identification of behavior a system is not limited to evaluate parameters that summarize the average behavior of the system as a result of integrating its variability on a last period, and from there to consider the calibrated model as static model; the real problem is the determination of the estimators of the parameters representing 'best' dynamics of the system in accordance with the observed data for the instant of current, time **t**. It implies concentrating the effort in the assessment of the changes that are occurring, then is possible to determine in 'real time' structural system changes that are reflected in changes in the values of the parameters, which in classical models only occurs after these changes have occurred.

The advantages of the approach for the state estimation versus the classical methods are presented below:

1. The biggest advantage of this scheme lies in the possibility of assuming hypotheses about the dynamics of variation of parameters. From this point of view the following aspects must be considered:
   - The estimators of the parameters of a system response function dependent on the observations (random variables) and should therefore be subject to adjustment in accordance with the information obtained.
   - A system parameters vary over time, since many of them are affected by changes that occur in a system as a result its development and/or its natural evolution.
   - From the point of view of monitoring system is suitable to use parameter values that conform to the present circumstances, instead of values that integrate knowledge of their behavior in the past.
   - Many of the changes that occur in the behavior of a system are consequences of unexpected events; before these eventualities schemes based on multiple states, or similar, to determine the moments in which structural changes occur, differentiating them from pure stochastic effects.

2. The integration of Bayesian estimation processes allows to discriminate, or to integrate, different structures of models, selecting one that better fits to observations at any given time.

3. One of the limitations in many systems is the frequency in the data acquisition. In many cases the data are obtained with frequencies lower than those required for monitoring the system (e.g., weekly, monthly, annual). From the observed information, state estimation modeling allows

estimates of all the variables of the system, including which have not been measured at that time.

4. Regarding statistical hypotheses, classical prediction models assume that the variables, endogenous and exogenous, are deterministic, considering that it is not possible to correct the past data in accordance with the processed of new observations.

5. The state estimation allows predictions afterwards (smoothing) such that in the case of backward values of the variables the data used are not the measured initially; the values obtained as a result of the data processing, that correct systematic errors that occur as a result of erroneous or little representative measurements.

6. Classical methods accumulated the errors of measurement and modeling in only one error of the model, which leads to confuse the origin for the noises of the system.

7. From the point of view of control, the state estimation allows the online analysis of the phenomena that are occurring, while conventional methods are aimed at "ex-post" analysis to describe phenomena already past.

There are at least two alternatives to address the problem, the first is the use of the Extended Kalman Filter (**EKF**), which is based on the fact that the above equation involves a non-linear relationship between variables/signals and parameters, which are the random elements involved in the problem.

The second alternative is based on multiple chained linear filters, commonly known as Dual Kalman Filter (**D-KF**). Typically, two filters are used that act in parallel, one on the variables/signals and one on the parameters.

7.2. TIME INVARIANT PARAMETERS

This case is related to the hypothesis that the parameters of the differential equations are permanent over time. Under this hypothesis the estimators of the parameters that would be obtained when t tends to infinity tend to an asymptotic value, the true value of the parameter. As this value is reached, the filter on the parameters no longer provides value.

7.2.1. UNIVARIATE MODEL

Consider a model with only a dependent variable:

$$\mathbf{Z(t) = \beta^T W(t) + \mu(t)}$$

where
Z(t) dependent variable in time **t**; for example, the flow of a river
W(t) vector of independent time variables **t**; for example, flows upstream, rain, wind,
β vector of parameters of the model,
μ**(t)** the error of the model in time **t**.

A model for the estate estimation may be formulate using the following definitions:

1. State equation:

$$\beta(t+1) = \beta(t) = \beta$$

2. Measurement equation:

$$\mathbf{Z(t) = W^T(t)\, \beta(t) + \theta(t)}$$

Estimates of β at time **t** correspond to the estimates of the state variables $\beta(t/t_p)$ when available information has been processed until the t_p. To apply **KF** it is necessary to have estimates of the initial state

$$\beta(0/0) = \beta_0$$
$$\Sigma_\beta(0/0) = \Sigma_0$$

and the variance of the measurement error

$$R(t) = s^2$$

The previous model, equivalent to a classical model, has the following implications:

1. It has been implicitly assumed that $\varepsilon(t)$, the error in modeling of the system dynamics, is equal to zero and its variance-covariance matrix, $Q(t)$, is equal to zero. This hypothesis assumes that the model is perfect, in this case that it is linear, and the parameters are constant and invariant over time. In statistical terms, this implies that when **t** tends to infinity the estimators of variance-covariance matrix tend to zero, converging to a constant vector independent of the value of **t**, it implies that as time elapses the new information does not provide any additional knowledge about the system. This model has not a learning process.

2. **KF** assumes that is called the variance of the error in the measurement s^2; this variance is associated to the precision of the measurement system and it should be known, since the measuring system is known. By not considering the modeling error, the error in measuring integrates the model errors and errors of measurement system, becoming a parameter that must be estimated. There are several alternatives for dealing with this problem:
   - Considering distributions Gaussians to all the stochastic component of the model, it is possible to work based on Bayesian regression models which extends the estimation process to s^2, this solution is the more formal from the statistician point of view,
   - The state vector can be extended to include s^2. This solution is consistent with the fact that the model state variables represent the parameters of the system,
   - Another alternative is the use of a combination of Bayesian models, assuming each one of them different values for s^2, and subsequently determine a posteriori probability of each model being the true [8]. Experimental studies [4] [8] have shown that asymptotically estimators $\beta(t/t)$ and $\Sigma_\beta(t/t)$ are independent of the information a priori.

7.2.2. MULTIVARIATE MODEL

Consider a multivariate model

$$\mathbf{Z(t) = \beta^T W(t) + \mu(t)}$$

where

Z(t) vector of dependent variables in the time **t**,
W(t) vector of independent variables in the time **t t**,
β array of parameters of the model, with vectors rows β_i,
$\mu(t)$ vector of model errors in the time **t**.

The following model may be formulated in terms of state estimation

1. State equation:

$$\beta_i(t+1) = \beta_i(t) = \beta_i$$

2. Measurement equation:

$$Z_i(t) = W^T(t)\beta_i(t) + \theta_i(t)$$

The previous formulation represents a linear system in which the state variables are grouped in an array, with vectors $\beta_i(t)$, and observations in an observations matrix with $Z_i(t)$ components. Filter expressions can be derived from a multivariate Bayesian regression scheme [9].

7.3. TIME-VARIANT PARAMETERS

Classical statistical models have limitations in the modeling of systems involving time-variants parameters. This limitation is resolved traditionally assuming preset forms of variation (for example, models associated with the months of the year, or the seasons,...). State estimation modeling allows to consider systems with time-variants parameters. Temporal variation may occur for reasons purely stochastic, or constant changes in the structure of the system for a period. Consider the univariate model:

$$Z(t) = \beta(t)^T W(t) + \mu(t)$$

assuming the vector of time-variant parameters $\beta(t)$

With respect to the previous modeling, parameter invariance hypothesis is directly related to the elimination of $\varepsilon(t)$, the term corresponding to the noise in the state equation. If it is assumed that $Q(t)$ variance-covariance matrix is zero, variations in the parameters due to stochastic reasons will be allowed. In this case, the $\Sigma_\beta(t/t)$ variance-covariance matrix estimator does not converge asymptotically to zero.

An alternative modeling can be done by allowing trend variations in the parameters. We define the dynamics of parameters based on the following equation:

$$\beta(t+1) = \beta(t) + \delta(t)$$

where $\delta(t)$ represents the change in the value of the parameter at time **t**, this is:

$$\delta(t) = \nabla_t b(t)$$

In this case the state estimation model can be defined as

1. State equation:

$$\beta(t+1) = \beta(t) + \delta(t) + \varepsilon_\beta(t)$$
$$\delta(t+1) = \delta(t) + \varepsilon_\delta(t)$$

2. Measurement equation:

$$Z(t) = H(t)^T\beta(t) + \theta(t)$$

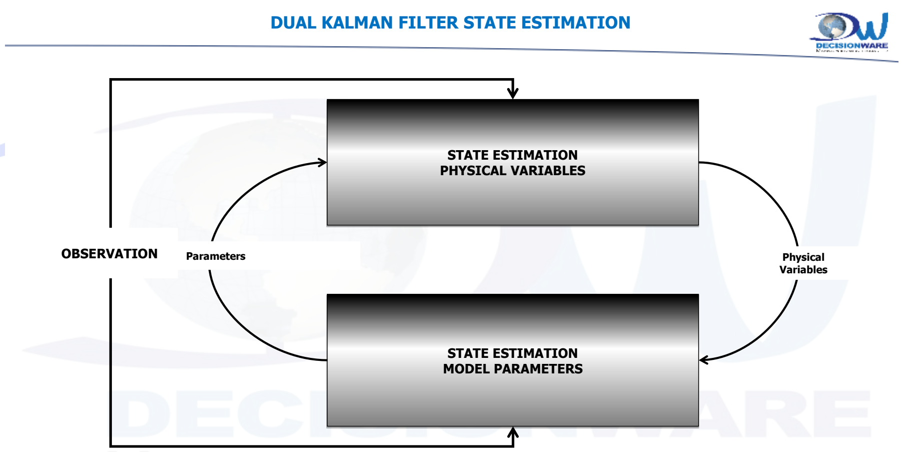
having to make estimates for $\beta(t)$ and $\delta(t)$. If the system is structurally stable must be met that

$$\nabla_t \beta(t) = 0$$

Alternatively, if the system is in the process of setting the trend of change, estimated based to δ**(t)**, which will be different from zero until to return to the equilibrium (a steady state).

7.4. DUAL KALMAN FILTER

Dual Kalman Filter (DKF) uses two filters that act in parallel, one on state variables and one on parameters of the statistical model, or on the physical model. Based on this approach there are multiple theoretical variations, here will be considered the one presented by Labarre et al. (2006). The diagram summarizes the process.



7.4.1. FRAMEWORK

For the presentation of this theory, the **VAR(p)** model that is represented in a vector way as a reference will be used as a reference

$$\mathbf{Y_t} = \Sigma_{i=1,p}\ \Lambda_i\ \mathbf{Y_{t-i}} + \varepsilon_t$$

where **p** represents the number of lags, $\mathbf{Y_t}$ the vector of endogenous variables, Λ_q the matrix of coefficients and ε_t the vector of innovations, which can be contemporary correlated with each other but which is not correlated with its own laggards nor is it correlated with endogenous variables.

7.4.2. MATHEMATICAL FORMULATION

For the formulation of mathematical theory let us consider that the equation of the **VAR(p)** model can be represented by the following expression

$$\mathbf{X(t) = A(t)\ X(t-1)} + \varepsilon_t$$

where the vector **X(t)** represents the state variables and has the following structure

$$\mathbf{X(t)} = \{\ \mathbf{Y_{t-1}},\ \mathbf{Y_{t-2}},\ \ldots,\ \mathbf{Y_{t-p}}\ \}$$

or

$$\mathbf{X(t)} = \{\ y_{1,t-1},\ y_{2,t-1},\ \ldots,\ y_{r,t-1},\ y_{1,t-1},\ y_{2,t-1},\ \ldots,\ y_{r,t-2},\ \ldots,\ y_{1,t-p},\ y_{2,t-p},\ \ldots,\ y_{r,t-p}\ \}$$

where $\mathbf{y_{i,t}}$ represents the value of the endogenous variable **i** in the t-period; **r** being the number of endogenous variables in the system. Matrix **A(t)** is composed of two submatrices

$$\mathbf{A(t)} = \begin{vmatrix} \mathrm{B}\mathbf{(t)} \\ \Phi \end{vmatrix}$$

where the submatrix B**(t)** has the following structure

$$\mathrm{B}\mathbf{(t)^T} = \{ \beta_1\mathbf{(t)^T}, \beta_2\mathbf{(t)^T}, \ldots, \beta_r\mathbf{(t)^T} \}$$

being β_i**(t)** the vector of parameters associated with the endogenous variable **i.** The Φ submatrix corresponds to a deterministic matrix that allows the lagging values of endogenous variables to be shifted backwards, the structure of which can be divided into **p-1** horizontal blocks and **p** vertical blocks, each associated with a square matrix of dimension **r.**

For the case of **p** equal to **4**, Φ is defined as follows

$$\Phi = \begin{vmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{vmatrix}$$

where **I** represents the identity matrix and **0** an array of zeros, each of dimension **r.**

The equation

$$\mathbf{X(t) = A(t)\, X(t\text{-}1)} + \varepsilon_t$$

has the characteristic of corresponding to the filter state equation over variables/signals and having embedded the filter observation equation over the parameters, which is equal to

$$\mathbf{Y(t)} = \mathrm{B}\mathbf{(t)\, Y(t\text{-}1)} + \theta_t$$

An expression that can also be written as

$$\mathbf{Y_t = I} \otimes \mathbf{Y(t\text{-}1)^T\, vec(}\mathrm{B}\mathbf{(t))} + \theta_t = \mathbf{X(t\text{-}1)^T}\, \beta + \theta_t$$

In terms of the vector of state variables, **X(t)**, it can be written as

$$\mathbf{Y(t) = G(t)\, X(t\text{-}1)} + \theta_t$$

where the matrix **G(t)** is defined as

$$\mathbf{G(t)} = \begin{vmatrix} \mathrm{B}\mathbf{(t)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{vmatrix}$$

where **Y(t)** represents the observation vector, **G(t)** the measurement matrix of the state variables and θ_t the noise vector in that measurement.

It would be necessary to include the dynamic equation on the parameters, which for purposes of this presentation is assumed equal to

$$\mathrm{B}\mathbf{(t)} = \mathrm{B}\mathbf{(t\text{-}1)} + \Psi_t$$

where B**(t)** represents the array of estimated parameters for the period **t** and Ψ_t to the error/noise matrix in that equation.

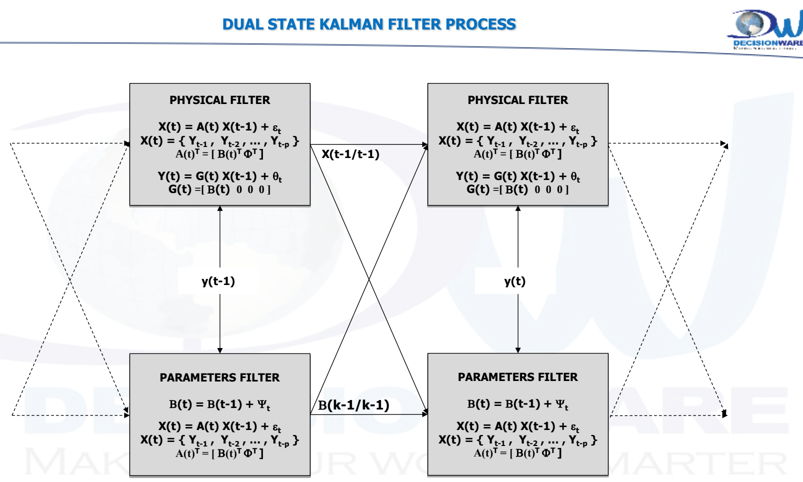
To follow the **KF** format the above expression can be presented as:

$$\mathrm{vec}(\mathrm{B}(t)) = I\ \mathrm{vec}(\mathrm{B}(t-1)) + I\ \mathrm{vec}(\Psi_t)$$

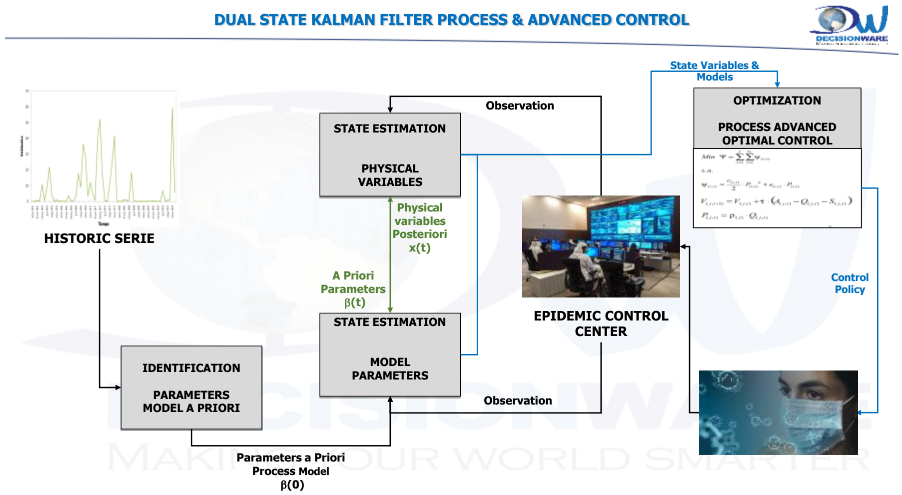
With the above elements you can formulate the dual-filter based on the equations presented in the following table.

DUAL KALMAN FILTER		
EQUATION	**PHYSICAL VARIABLES FILTER**	**STATISTICS PARAMETERS FILTER**
Dynamic State	$X(t) = A(t)\ X(t-1) + \varepsilon_t$ $X(t) = \{ Y_{t-1}, Y_{t-2}, \dots, Y_{t-p} \}$ $A(t) = \begin{vmatrix} \mathrm{B}(t) \\ \Phi \end{vmatrix}$	$\mathrm{B}(t) = \mathrm{B}(t-1) + \Psi_t$ = $\mathrm{vec}(\mathrm{B}(t)) = I\ \mathrm{vec}(\mathrm{B}(t-1)) + I\ \mathrm{vec}(\Psi_t)$
Observation	$Y(t) = G(t)\ X(t-1) + \theta_t$ $G(t) = \begin{vmatrix} \mathrm{B}(t) & 0 & 0 & 0 \end{vmatrix}$	$X(t) = A(t)\ X(t-1) + \varepsilon_t$ $X(t) = \{ Y_{t-1}, Y_{t-2}, \dots, Y_{t-p} \}$ $A(t) = \begin{vmatrix} \mathrm{B}(t) \\ \Phi \end{vmatrix}$

As you can see the filter status equation on the physical variables (flow, precipitation, epidemic states, ...) corresponds to the filter observation equation on statistical parameters and becomes the central axis of Dual Filter variables-parameters, whose central idea is to formulate two filters and exchange the common information between them, in this way a more coherent estimate is achieved than that of using a single filter or two independent filters. The next figure shows the idea.



The following graph presents how the DKF process is integrated into the continuous optimization of a dynamic system under uncertainty.



7.5. EXTENDED KALMAN FILTER

In the Extended Kalman Filter (**EKF**), the state transition and observation models does not need be linear functions of the state but may instead be nonlinear functions. These functions are of differentiable type.

To solve the problem of non-linearity in the status equation, the Extended Kalman Filter (EKF) proposal is based on performing a first-order Taylor approximation at each point at which the state estimation process must be performed.

For example, in parameter estimation the dynamic equation involves a non-linear relationship between the state variables, **X(t-1)**, and the parameters, **A(t)**, which are the random elements involved in the problem.

$$\mathbf{X(t) = A(t) \times X(t-1) + m^X_t}$$

$$\mathbf{A(t) = A(t-1) + m^A_t}$$

Taylor's first-order approach to the vector function **f(x)** is:

$$\mathbf{F(x) \cong F(x^*) + \nabla F(x^*)^T (X^* - X)}$$

Based on the above definition, **X(t)** can be approximated as:

$$\mathbf{X(t) = X(t-1) + \nabla[A(t) \times X(t-1)]\,(X(t-1) - X(t)) + m^X_t}$$

$$\mathbf{A(t) = A(t-1) + m^A_t}$$

8. BAYESIAN SELECTION OF DYNAMIC EQUATIONS

Since there is no "true" model of a physical system, or a technical, an alternative to improve the predictions is to use combination of several models that may potentially represent such a system. In this respect the experience of Alan Blinder (Former Federal Reserve Vice Chairman, 1998) may be illustrative:

1. "My way to approach this problem while I was in the Fed was relatively simple: use a wide variety of models and ever over-rely on one of them".
2. When the Fed experts explored the consequences of various measures, I always insisted on seeing the results of:
   i) Our own quarterly econometric model,
   ii) Some other econometric models
   iii) A variety of Vector Autoregressive (VAR) that I developed for this purpose.

The empirical evidence is consistent with the arguments of Blinder that encourage the use of different methodologies/technologies with the aim of improving the performance of the forecasts. Therefore, we can expect a minor error of forecast using a combination of techniques obtained through a process of searching for the optimal weights of each of the models. The simplest position is to use the average; however, there are other possibilities, as the combination of Bayesian models which has been applied in many cases.

Then, Bayes' theory has been widely used for when it is unknown what is the true model that represents the dynamics of a system. Two cases can be considered:

- Discrimination of a model as the correct model
- Determination of the state in which a system is located that is characterized by the response function (system dynamics) that can change over time.

8.1. BAYESIAN ENSEMBLE OF KALMAN MODELS

8.1.1. MATHEMATICAL FORMULATION

Below, is the extension of the Kalman filter theory in order to compare simultaneously several models of a system. The theoretical development that follows specify a model $\mathbf{M_i}$ based on the set of parameters

$$\mathbf{M_i = \{A_i(t), B_i(t), L_i(t), H_i(t), Q_i(t), R_i(t), x_i(0/0), \Sigma_i(0/0)\}}$$

the subscript **i** indicates the model to which the parameter is associated.

Consider Ψ**i(t/tp)** as a set of estimates of the first moments for the distribution of probability of the state vector at time **t** when the information available up to the time $\mathbf{t_p}$ is processed

$$\Psi_i(t/t_p) = \{ X_i(t/t_p), \Sigma_i(t/t_p) \}$$

π_i**(t)** is the probability a posteriori that the model **i** is true when all the information has been processed up to time t

$$\pi_i(t) = \text{Probability}[\text{model}(t) = i \mid I(t)]$$

Based on the Bayes theorem the value π_i**(t)** is defined as:

$$\pi_i(t) =$$
$$C \times \text{Probability}[Z(t) \mid \text{model}(t-1) = i, I(t-1)] \times \text{Probability}[\text{model}(t-1) = i \mid I(t-1)]$$

where **model(t)** represents the estimator of the true model in time **t**. The previous expression is equal to

$$\pi_i(t) = C \times L_i[Z(t) \mid x_i(t/t-1), \Sigma_i(t/t-1)] \times \pi_i(t-1)$$

where $\mathbf{L_i[.]}$ corresponds to the likelihood function of $\mathbf{Z(t)}$ conditioned in the estimators $\mathbf{x_i(t/t-1)}$ and $\Sigma_i\mathbf{(t/t-1)}$ is defined as

$$L_i[.] = \{2\pi\ |V_i(t/t-1)|\}^{-m/2}$$
$$\text{Exp}\ [\theta_i(t/t-1)^T V_i(t/t-1)^{-1} \theta_i(t/t-1)]$$

where

$$\theta_i(t/t-1) = Z(t) - H(t)\ X_i(t/t-1)$$
$$V_i(t/t-1) = H_i(t)^T \Sigma_i(t/t-1) H_i(t) + R_i(t)$$

and the constant **C** is a normalization constant equal to

$$C = 1.\ /\ \{\Sigma_i\ L_i[.]\ \pi_i(t-1)\}$$

This process is conditioned in a priori probability for each model, $\pi_i\mathbf{(0)}$ and allows:

- Select a model from a set of probable models, taking as true those who have most a posteriori probability; or
- Generate a merged model based on the weighting of models according to their a posteriori probability, in this case the estimate of the state of the system is formulated as

$$x(t/t_p) = \Sigma_i\ \pi_i(t)\ x_i(t/t_p)$$

8.1.2. CASE: PROJECTION OF THE ENSO (EL NIÑO) PHENOMENON

The prediction of the **ENSO** (**E**l **N**iño **S**outhern **O**scillation) phenomenon by the International Research Institute for Climate and Society of Columbia University (**IRI**, http://iri.columbia.edu) uses two types of models for the prediction of **ENSO**:

1. **Dynamic:** based on the physical explanation of the dynamics of the process; and
2. **Statistics:** based on empirical evidence of the process adjusted through statistical models.

TABLE 1. Dynamical and statistical models whose forecasts for Niño-3.4 SST anomaly are included in this study. Note that some models were introduced during the course of the study period, or replaced a predecessor model.

Dynamical models	Model type
NASA GMAO	Fully coupled
NCEP CFS (version 1)	Fully coupled
Japan Meteorological Agency	Fully coupled
Scripps Hybrid Coupled Model (HCM)	Comprehensive ocean, statistical atmosphere
Lamont–Doherty	Intermediate coupled
Australia POAMA	Fully coupled
ECMWF	Fully coupled
UKMO	Fully coupled
Korea Met. Agency SNU	Intermediate coupled
Univ. Maryland ESSIC	Intermediate coupled
IRI ECHAM/MOM	Fully coupled, anomaly coupled
COLA Anomaly	Anomaly coupled
COLA CCSM3	Fully coupled
Météo-France	Fully coupled
Japan Frontier FRCGC	Fully coupled
Statistical Models	**Method and predictors**
NOAA/NCEP/CPC Markov	Markov: Preferred persistence and transitions in SST and sea level height fields
NOAA/ESRL Linear Inverse Model (LIM)	Refined POP: Preferred persistence and transitions within SST field; optimal growth structures
NOAA/NCEP/CPC Constructed Analogue (CA)	Analogue-construction of current global SSTs
NOAA/NCEP/CPC Canonical Correlation Analysis (CCA)	Uses SLP, tropical Pacific SST and subsurface temperature (subsurface not used beginning in 2010)
NOAA/AOML CLIPER	Multiple regression from tropical Pacific SSTs
UBC Neural Network (NN)	Uses sea level pressure and Pacific SST
Florida State Univ. multiple regression	Uses tropical Pacific SST, heat content, winds
UCLA TDC multilevel regression	Uses 60°N–30°S Pacific SST field

IRI publishes results of two methodologies for projections of models combining:

1. **Subjective** based on a consensus among analysts of the **CPC** (**C**limate **P**rediction **C**enter) and **IRI**.

2. **Objective** based on a mathematical model that determines the weighting factors of each of the models available based on the results obtained in the last periods based on a combination Bayesian model.

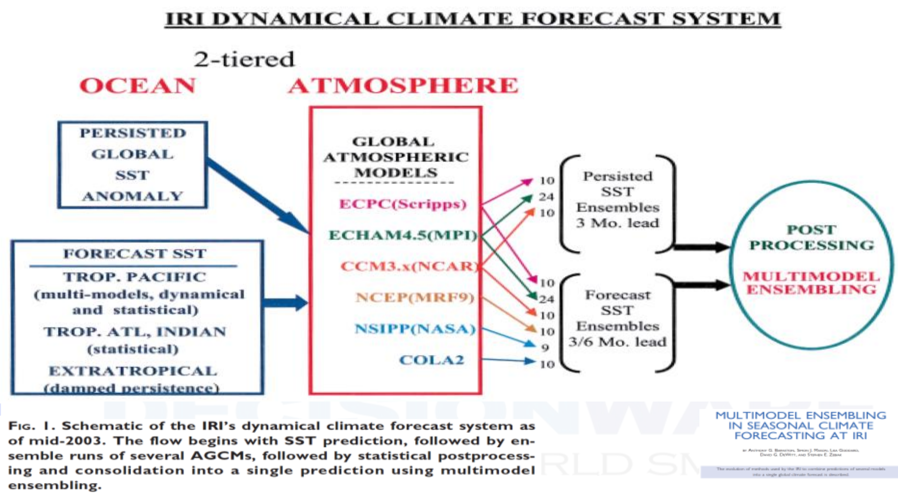
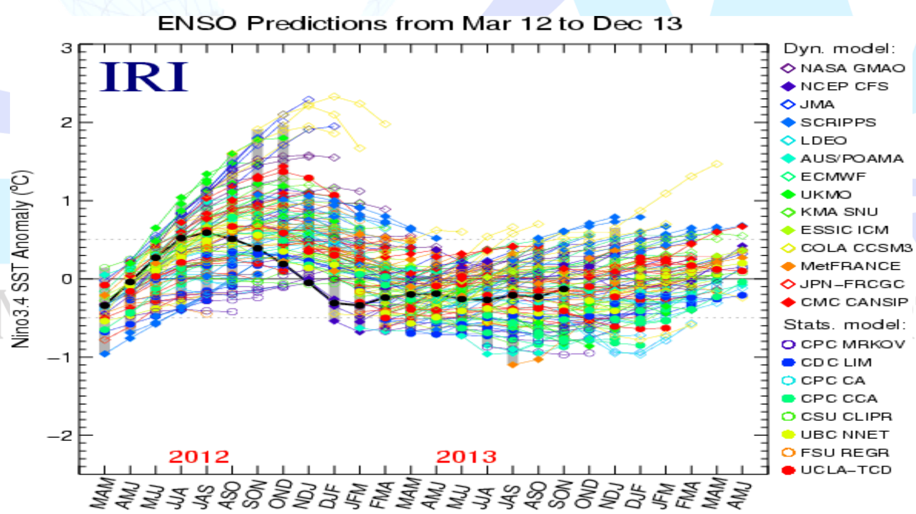


Fig. 1. Schematic of the IRI's dynamical climate forecast system as of mid-2003. The flow begins with SST prediction, followed by ensemble runs of several AGCMs, followed by statistical postprocessing and consolidation into a single prediction using multimodel ensembling.

The benefits of this approach are:

1. Discrimination of a model such as the "real" model (the most probable ?)
2. Composition of a general model based on the weighting of different models (using its probability to be the true model). This dynamic selection of the structure of the model implies a learning process.



8.2. MULTI-STATE KALMAN FILTER (MS-KF)

The Kalman filter theory can be extended to consider the modeling of systems whose dynamic representation depends on the state, or regime, in which the system is.

Multiple-**S**tate **K**alman **F**ilter (**MS-KF**, Velasquez, 1978) corresponds to the extension of the theory of the Kalman Filter to model systems whose dynamic representation depends on the state, or regime, in which the system is. Example of a system status can be:

- A person: good temper, temper, quiet.
- In a watershed: dry, wet, in transition
- In a client: happy, unhappy,...
- Epidemic State Estimation: stable, changing

There are plenty of examples that can be cited. In all of them the response of the system (person, basin, client, epidemic state, ...) can respond to stimuli in the environment in accordance with the state in which the system is located, which in many cases is not fully identified, because is a random variable.

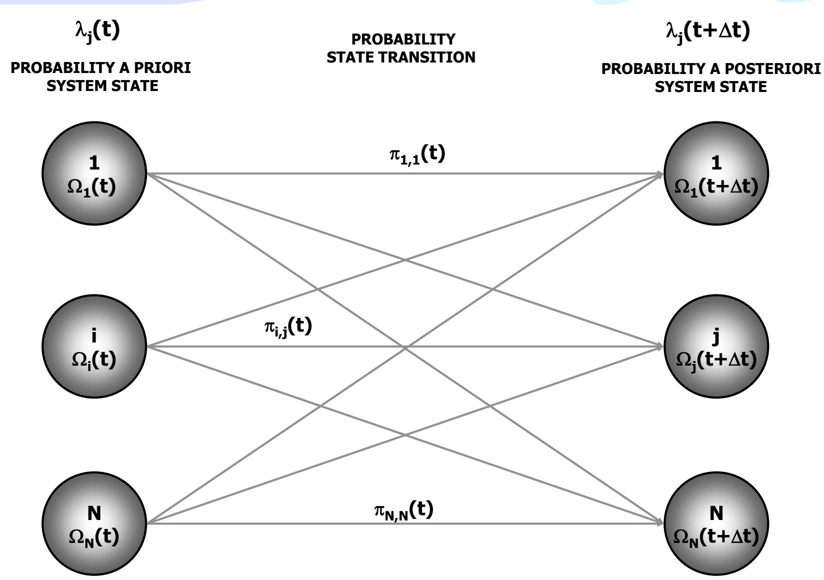
The propose way is to chain the Markovian and the Bayesian modeling. It can be considered the states **S$_i$** of a Markov process as the random variables of Bayesian inference process that produces estimators of the probability of the transition probability between the state **i** and state **j** of the system Markov process.

The following section, which corresponds to an example of modeling with multiple States (**KF-MS**), discusses in more detail the problem of changes in the parameters of the system response function.

8.2.1. MATHEMATICAL FORMULATION

Therefore, the concept of states may be associated with a semi-Markov process (time-dependent transition matrices) where the probability of transition from one state to another, at any interval of length Δ**t**, which is not stationary and can be changed to the extent that develops the process over time.

In this case the regime in which the system is located should be considered as a variable of additional state which must be estimated "in line"; the state may be associate to a state of a Markov chain.



States are identified by the subscripts **i** and **j**. The set of parameters associated with each state **i**, which can be identified using the Kalman filter.

Kalman filter must be changed to calculate the probability a posteriori that the system is in a state **j** considering all the transitions that may have occurred from any state **i** to state **j** during the period **t**. Each State **j** is characterized by the following set of parameters

$$\Omega_j(t) = \{ \pi_j(t), A_j(t), B_j(t), L_j(t), H_j(t), Q_j(t), R_j(t) \}$$

where $\pi_j(t)$ represents the probability a priori that the system is in state **j** at time **t**, which complies with the following dynamic equation

$$\pi_j(t) = \lambda_j(t-1)$$

where $\lambda_j(t)$ is the probability a posteriori that the system is in state **j** in time **t.**

Consider the probability that between **t-1** and **t** is made a transition from state **i** to state **j** defined as

$$\pi_{i,j}(t) = \text{Probability}[\text{state}(t) = j, \text{state}(t-1) = i \mid I(t)]$$

Based on the Bayes theorem the value of $\pi_{i,j}(t)$ is defined as

$$\pi_{i,j}(t) = C \times \text{Probability}[Z(t) \mid \text{state}(t)=j, \text{state}(t-1)=i, I(t-1)]$$
$$\times \text{Probability}[\text{state}(t)=j \mid \text{state}(t-1)=i, I(t-1)] \times \text{Probability}[\text{state}(t-1)=i \mid I(t-1)]$$

that is equal to

$$\pi_{i,j}(t) = C \times L[Z(t) \mid \text{state}(t)=j, \text{state}(t-1)=i, I(t-1)] \times \pi_j(t) \times \lambda_i(t)$$

where **L[.]** represents the likelihood function of **Z(t)** conditioned in make the transition from **i** to **j**, and the information processed up to time **t-1**. **C** is the normalization constant.

To express the previous formula in terms of the estimators of the variance-covariance matrix and state variables is necessary to define estimators $\Theta_{i,j}(t/t)$ in accordance with

$$\Theta_{i,j}(t/t) = \{x_{i,j}(t/t), \Sigma_{i,j}(t/t)\}$$
$$= \text{BAYES}[\Theta_i(t/t-1), Z(t), U(t), Q_j(t), R_j(t)]$$

Based on the above estimates **L [.]** is defined as

$$L[.] = \{2\pi \; |V_{i,j}(t/t-1)|\}^{-m/2}$$
$$\text{Exp}\,[\theta_i(t/t-1)^T V_{i,j}(t/t-1)^{-1} \theta_i(t/t-1)]$$

where

$$\theta_i(t/t-1) = Z(t) - H(t)x_i(t/t-1)$$
$$V_{i,j}(t/t-1) = H_i(t)^T \Sigma_i(t/t-1) H_i(t) + R_j(t)$$

The estimates a priori are defined as

$$x_j(t/t-1) = A_j(t-1)x_j(t-1/t-1) + B_j(t-1)U(t-1)$$
$$\Sigma_j(t/t-1) = A_j(t-1)\Sigma_j(t-1/t-1)A_j^T(t-1)$$
$$+ L_j(t-1)Q_j(t-1)L_j(t-1)$$

The information a posteriori may condense through the following process

$$\lambda_j(t) = \Sigma_{i=1,N}\ \pi_{i,j}(t)$$
$$x_j(t/t) = \Sigma_{i=1,N}\ \pi_{i,j}(t)\ x_{i,j}(t/t)$$
$$\Sigma_j(t/t) = \Sigma_{i=1,N}\ \pi_{i,j}(t)\ \Sigma_{i,j}(t/t)$$

determining a posteriori estimator based on integration on all possible transitions to a state **j** starting from **N** possible states.

The prediction of the regime/state of the system can be based on the weighting of all possible systems, or by selecting one who is more likely.

8.2.2. CASE: STREAMFLOW FORECAST

8.2.2.1. NWSRFS (U.S. NATIONAL WEATHER SERVICE RIVER FORECASTING SYSTEM)

Rainfall-runoff models are supported in conceptualizations of the hydrological system that simulate the physical process that occurs in the basin and that converts precipitation into flow at the outlet of the basin. Among the most well-known models can be mentioned the **NWSRFS (U.S. National Weather Service River Forecasting System)**

The high cost of implementing such models, which require a large amount of physical type data, is one of its main features. These models divide the basin into sub-basins of homogeneous characteristics and on them are imposed equations of water propagation.

The process of implementing the models involves a process of calibration of the physical parameters that characterize the sub-basins, which is usually done from a historical sample of precipitation and flow series with high precision, i.e. time level, or lower, for small and daily level basins for large basins.

The research questions faced were:
1. How can we provide improved estimations of basin initial conditions (e.g., soil moisture, snow-pack) at the start of a forecast period; and
2. How do we characterize hydrologic model uncertainties.

The first question is considered under the topic of Data Assimilation, and the second question is addressed through the Multi-model Super-ensemble techniques in hydrology.

1. **Data Assimilation.**
   NWSRFS has implemented the Ensemble Kalman Filter (EnsKF) in the SNOW-17 model. Snow data assimilation using the Extended Kalman Filter (EKF) is an elegant solution to the challenges posed by the EKF.

2. **Multi-Model Super-Ensemble**
   NWSRFS multi-model super-ensemble technique mixes and matches various methods for modeling hydrologic processes to allow the construction of multiple models, all with different structure. We are using the USGS Modular Modeling System (http://wwwbrr.cr.usgs.gov/mms/) as an integrating framework. For any given model automated parameter estimation methodology (e.g., MOCOM-UA) will be used to determine parameter sets. The potential benefits of this approach are:
   - The super-ensemble provides probabilistic information content.
   - The spread of the super-ensemble provides an estimate of forecast uncertainty.

- The approach allows automated configuration of hydrological models over multiple river basins with minimal human effort.
- The multi-model system reduces the commitment of operational agencies to their own model, and thus may allow more rapid transfer of new ideas from the research community to the operational setting.

8.2.2.2. CASE: MS-KF APPLIED IN WATER RESOURCES SYSTEM

The response of a system function is the reflection of the form of functioning of its internal structure. Structural changes produce changes in the response function. One of the main limitations in the modeling of dynamic systems is related to the hypothesis that the response function is known a priori, not being able to easily model structural changes in the system. Examples of structural changes due to events: i) natural (tsunamis, earthquakes,...), ii) cyclic processes (climatic, changes from summer to winter passing through intermediate stations), iii) economic (economic crises,...), …

In order to model structural changes, the following hypothesis is formulated:
i) A system can be in one of multiple states, or regimes.
ii) Characterized each state by the error variance-covariance matrices $\mathbf{Q_\beta(t)}$, $\mathbf{Q_\delta(t)}$ y $\mathbf{R(t)}$.

Consider the case of three States:
- Steady
- Structural Transient
- Stochastic Transient

In **steady** state parameters that define the dynamics of the system do not vary, this is

$$\nabla_t\beta(t) = 0$$

In **structural transient** state, the dynamics of the system is changing translating the impact into changes in parameters

$$\nabla_t\beta(t) \neq 0$$

In the **stochastic transient** state, the system is affected by shocks due to its intrinsic probabilistic nature, that are not always explained by the dynamics of the system, and they are not due to structural changes, reflecting its effect on measurement errors.

Then the system must be in one of three states, the state in which the system is located must be estimated, since it cannot be identified a priori. A real case are the effects of global warming which is changing the shape of the behavior of water systems in unexpected ways; for example, if the system is represented by models which behavior is associated with the months of the year, being the months associated with climate change; the impact of arrears and the overtaking of the seasons (winter, spring, summer, autumn) may not be associated with months of the year; then an stationary model, 12/52 models for the year, cannot properly manage this new process.

Models based on classical methods can not react ("learning") to structural changes since they do not have a clear sign of the changes that are occurring, and no distinction between structural changes of purely stochastic effects. . The consequence of this fact is that models react slowly during periods of structural change and can be interpreted as indication of structural change effects of random origin.

A **MS-KF** model with time-variants parameters may be governed by the following equations

1. State equation:

$$\beta(t+1) = \beta(t) + \delta(t) + \varepsilon_{\beta}(t)$$
$$\delta(t+1) = \delta(t) + \varepsilon_{\delta}(t)$$

2. Measurement equation:

$$Z(t) = H(t)^T\beta(t) + \theta(t)$$

Then, it is possible to characterize the states based on the different terms of error variance-covariance matrices; this based on the response of a system function is the reflection of the form of functioning of its internal structure, the structural changes produce changes in the response function.

The proposed characterization has in mind that the entropy of a dynamic system is not independent of the state in which it is located, and that it cannot be characterized by values of the state variables, since such entropy can be generated exogenously, and reflected in the matrices of variance-covariance change that depend on the state in which the system is located. The following table shows qualitative characteristics of these matrices

State	Type of Noise		
	$Q_{\beta}(t)$	$Q_{\delta}(t)$	$R(t)$
Steady	Normal	Normal	Normal
Structural Transient	Normal	Large	Normal
Stochastic Transient	Normal	Normal	Large

This case can be considered as a particular case **MS-KF**.

In 1978, Velasquez [7] implemented a **MS-KF** for the prediction of flows in the basin of the Caroni River in Venezuela. The objective was to predict the river-flow at different points of the basin of the Caroni River which flows into the reservoir of Guri, which feeds a system of hydro-generating more than 10,000 MW of nominal installed capacity. Given the size of the reservoir, and the low-altitude barrage, 162 meters, slight variations in the height of the water in the reservoir, generates large variations in the curve of hydro-generation.

Then presents the results of a prediction model with anticipation of a day, for two sites in the basin:

i) San Pedro, punto donde se unen el río Caroní con el río Paragua (98.000 $kmts^2$), y

ii) Arekuna, una estación de medición de caudal ubicada sobre el río Caroní aguas arriba de San Pedro (46.000 $kmts^2$).

The statistical models have the following structure:

- **San Pedro Model (SP):**

$$Q^{SP}(t) = \alpha^{SP}(t) + \Sigma_{j=1,2}\,\Sigma_{q=1,TVI(j)}\ \beta_{j,q}(t) \times Q^{j}(t-q) + \varepsilon(t)$$

where $Q^j(t)$ represents the average flow on the site **j** **{Arekuna, La Paragua}** during the period **t**, $\beta_{j,q}(t)$ the parameter estimated period **t** for the fraction of flow reaching San Pedro **q** days after having gone through the site **j** and $\alpha^j(t)$ flow base expected to day **t** in site **j**.

- **Arekuna Model (AR):**

$$Q^{AR}(t) = \alpha^{AR}(t) + \Sigma_{j=1,2}\,\Sigma_{q=1,TVI(j)}\ \theta_{j,q}(t) \times P^{j}(t-q) + \varepsilon(t)$$

where $\theta_{i,q}(t)$ represents the parameter estimated during the period **t** for the fraction of rain, $P^j(t-q)$, reaching Arekuna **q** days after having fallen in the area associated with the site **j, {Arekuna, Wonken, Kanavaten, Uriman}**.

State estimation modeling using the state equation and the observation equations presented previously.

The following table presents the summary of the results when they are evaluated based on the statistics defined by the World Meteorological Organization, **WMO**. **K-I** corresponds to a standard **KF** and **K-V** to a **MS-KF**. As remarkable result can be verified as **MS-KF** still more adjusted flow signal for the extreme flows and the average, which shows improvement in the results due to the process of learning that goes by adjusting the response function and the state of the system of considering the reception of the data.

W.M.O. STATISTICAL

Model	DAILY STREAMFLOW				MONTHLY MAXIMUM FLOW RATE				MONTHLY MINIMUM FLOW RATE				VOLUME	
	mcs	A	R	C	mcs	A	R	C	mcs	A	R	C	MM^3	C
San Pedro														
REAL	4905				7115				1198				12440	
K-I	4851	.032	.010	.050	7110	.040	.000	.056	1206	.024	.009	.033	12613	.019
Arekuna														
REAL	2494				4138				798				6425	
K-V	2497	.131	.001	.174	3917	.044	.033	.122	746	.209	.001	.296	6428	.015
K-I	2537	.154	.017	.193	3545	.139	.131	.182	1251	.507	.567	.589	6502	.143

K-I KALMAN INVARIANT
K-V KALMAN VARIANT MULTIPLES STATES

$C = [\, \Sigma (Y_c - Y_o)^2/n)^{1/2} / Y_m$
$R = [\, \Sigma (Y_c - Y_o) / nY_m$
$A = [\, \Sigma |Y_c - Y_o| / nY_m$

Y_c Calculated Streamflow

Y_o Observed Streamflow

Y_m Mean Streamflow ($Y_m = \Sigma Y_c / n$)

9. EPIDEMIC STATE AND PARAMETER ESTIMATION

This section presents the proposed formulation for the SEIMR/R-S model. for implementation using Kalman's filter theory and the variations and improvements introduced that have been presented in this document.

From the point of view of managing a pandemic, the mathematical problems to be faced:

i) To define the structure of the models of differential equations that govern the behavior of the pandemic
ii) To estimate the parameters that define a specific model within the "infinity" of possible models to describe the dynamic process.
iii) To know the true state of the pandemic, which is defined by the number of people, or by the fraction of the population, which is in each epidemiological state.

9.1. DIFFERENTIAL EQUATIONS

The differential equations of the **SEIMR/R-S** (regional-segmented epidemic) model are:

$$\partial S_{rg,ss}(t)/\partial t = - S2I_{rg,ss}(t) - \mu^N \times S_{rg,ss}(t) + \lambda^S{}_{rg,ss} \times NPX(t)$$

$$\partial E_{rg,ss}(t)/\partial t = S2I_{rg,ss}(t) - \psi \times E_{rg,ss}(t) + \lambda^E{}_{rg,ss} \times NPX(t)$$

st=I0

$$\partial I_{st,rg,ss}(t)/\partial t = \psi \times E_{rg,ss}(t) - \delta\alpha_{st,ss} \times I_{st-1,rg,ss}(t) + \lambda^I{}_{rg,ss} \times NPX(t)$$

st∈I1F

$$\partial I_{st,rg,ss}(t)/\partial t = \delta\zeta_{st-1,ss} \times I_{st-1,rg,ss}(t) - \delta\alpha_{st,ss} \times I_{st,rg,ss}(t)$$

$$\partial R_{rg,ss}(t)/\partial t = \Sigma_{st\in I1F}\ \delta\beta_{st-1,ss} \times I_{st,rg,ss}(t) - \mu^N \times R_{rg,ss}(t) + \Sigma_{ss\in SSR(rg)}\ \lambda^R{}_{rg,ss} \times NPX(t)$$

$$\partial D_{rg,ss}(t)/\partial t = \Sigma_{st\in I1F}\ \mu\sigma_{ss} \times I_{st,rg,ss}(t)$$

$$\partial NR_{rg}(t)/\partial t = \mu^N \times SR_{rg,ss}(t) + \mu^N \times RR_{rg}(t)$$

where the following rates are defined for the socio-demographic segments

SOCIO-DEMOGRAPHIC BIOLOGICAL PARAMETERS		
Parameter	Equation	Description
$\delta\alpha_{st,ss}$	$\Sigma_{ag\in AGS(ss)}\ \delta\delta_{ag,st}$	Total exit rate
$\delta\zeta_{st,ss}$	$\Sigma_{ag\in AGS(ss)}\ \delta\sigma_{ag,st}$	Worsening exit rate
$\delta\beta_{st,ss}$	$\Sigma_{ag\in AGS(ss)}\ \delta\gamma_{st,ag}$	Recovering exit rate
$\mu\sigma_{ss}$	$\Sigma_{ag\in AGS(ss)}\ \mu_{ag}$	Mortality rate depending on segment
μ_{ag}	$\Sigma_{st\in I1F}\ \delta\sigma_{ag,st}$	Mortality rate depending on age

The definition equations of the regional-segmented model are:

$$IS_{rg,ss}(t) = \Sigma_{st\in INF}\ \beta_{st,rg,ss}(t) \times I_{st,rg,ss}(t)$$
$$IX_{rg}(t) = \Sigma_{ss\in SSR(rg)}\ IS_{rg,ss}(t)$$
$$II_{rg}(t) = \Sigma_{ss\in SSR(rg)}\ \Sigma_{ro\in ROR(rg)}\ \varphi\phi_{ro,rg,ss} \times IS_{ro,ss}(t)$$
$$IE_{rg}(t) = \Sigma_{ss\in SSR(rg)}\ \Sigma_{rd\in RDE(rg)}\ \varphi\phi_{rg,rd,ss} \times IS_{rg,ss}(t)$$
$$IR_{rg}(t) = IX_{rg}(t) + II_{rg}(t) - IE_{rg}(t)$$

$$SR_{rg}(t) = \Sigma_{ss\in SSR(rg)}\ S_{rg,ss}(t)$$
$$SI_{ro,rg,ss}(t) = \varphi\phi_{ro,rg,ss} \times S_{ro,ss}(t)$$
$$SE_{rg,rd,ss}(t) = \varphi\phi_{rg,rd,ss} \times S_{rd,ss}(t)$$
$$SN_{rg,ss}(t) = S_{rg,ss}(t) - \Sigma_{rd\in RDE(rg)}\ SE_{rg,rd,ss}(t)$$
$$SIN_{rg}(t) = \beta\beta_{rg,ss} \times IR_{rg}(t) \times SN_{rg,ss}(t)$$
$$SIE_{rg,ss}(t) = \Sigma_{rd\in RDE(rg)}\ \beta\beta_{rd,ss} \times IR_{rd}(t) \times SE_{rg,rd,ss}(t)$$
$$S2I_{rg,ss}(t) = SIN_{rg}(t) + SIE_{rg}(t)$$

$$RR_{rg}(t) = \Sigma_{ss\in SSR(rg)}\ R_{rg,ss}(t)$$

$$DR_{rg}(t) = \Sigma_{ss\in SSR(rg)}\ D_{rg,ss}(t)$$

From now on, the above mathematical definitions will be summarized as

$$\{\ \mathbf{S, E, I_{st}, D, N}\ \} \in \Theta$$

The next table shows the equations dividing the increment and the decrement on each state, it must be considered in the implementation of the mathematical models. The table includes the sets that defined the existence of the equations manly for the infected states.

Set	SIR Regional – Segmented Model - Differential Equations				
	State	State Increment	State Decrement	Natural Dead	Exogenous Increment
	REGIONAL - SEGMENT EQUATIONS				
SU	$\partial S_{rg,ss}(t)/\partial t$		$S2I_{rg,ss}(t)$	$\mu^N \times S_{rg,ss}(t)$	$\lambda^S_{rg,ss} \times NPX(t)$
EX	$\partial E_{rg,ss}(t)/\partial t$	$S2I_{rg,ss}(t)$	$\psi \times E_{rg,ss}(t)$		$\lambda^E_{rg,ss} \times NPX(t)$
I0	$\partial I_{st,rg,ss}(t)/\partial t$	$\psi \times E_{rg,ss}(t)$	$\delta\alpha_{st,ss} \times I_{st,rg,ss}(t)$		$\lambda^I_{rg,ss} \times NPX(t)$
I1F	$\partial I_{st,rg,ss}(t)/\partial t$	$\delta\zeta_{st-1,ss} \times I_{st-1,rg,ss}(t)$			
RE	$\partial R_{rg,ss}(t)/\partial t$	$\Sigma_{st\in I1F}\ \delta\beta_{st,ss} \times I_{st,rg,ss}(t)$		$\mu^N \times R_{rg,ss}(t)$	$\lambda^R_{rg,ss} \times NPX(t)$
ED	$\partial D_{rg,ss}(t)/\partial t$	$\Sigma_{st\in I1F}\ \mu\sigma_{ss} \times I_{st,rg,ss}(t)$			
ND	$\partial NR_{rg}(t)/\partial t$	$\mu^N \times (\ SR_{rg}(t) + RR_{rg}(t)\ )$			

SUSCEPTIBLE STATE EQUATIONS
$SR_{rg}(t) = \Sigma_{ss\in SSR(rg)}\ S_{rg,ss}(t)$
$SI_{ro,rg,ss}(t) = \varphi\phi_{ro,rg,ss} \times S_{ro,ss}(t)$
$SE_{rg,rd,ss}(t) = \varphi\phi_{rg,rd,ss} \times S_{rd,ss}(t)$
$SN_{rg,ss}(t) = S_{rg,ss}(t) - \Sigma_{rd\in RDE(rg)}\ SE_{rg,rd,ss}(t)$
$SIN_{rg,ss}(t) = \beta\beta_{rg,ss} \times IR_{rg}(t) \times SN_{rg,ss}(t)$

	SIR Regional – Segmented Model - Differential Equations				
Set	State	State Increment	State Decrement	Natural Dead	Exogenous Increment
	$SIE_{rg,ss}(t) = \Sigma_{rd \in RDE(rg)}\ \beta\beta_{rd,ss}\ IR_{rd}(t) \times SE_{rg,rd,ss}(t)$				
	$S2I_{rg,ss}(t) = SIN_{rg}(t) + SIE_{rg}(t)$				
	INFECTED STATE EQUATIONS				
	$IS_{rg,ss}(t) = \Sigma_{st \in INF}\ \beta_{st,rg,ss}(t)\ I_{st,rg,ss}(t)$				
	$IX_{rg}(t) = \Sigma_{ss \in SSR(rg)}\ IS_{rg,ss}(t)$				
	$II_{rg}(t) = \Sigma_{ss \in SSR(rg)}\ \Sigma_{ro \in ROR(rg)}\ \varphi\phi_{ro,rg,ss} \times IS_{ro,ss}(t)$				
	$IE_{rg}(t) = \Sigma_{ss \in SSR(rg)}\ \Sigma_{rd \in RDE(rg)}\ \varphi\phi_{rg,rd,ss} \times IS_{rg,ss}(t)$				
	OTHER EQUATIONS				
	$RR_{rg}(t) = \Sigma_{ss \in SSR(rg)}\ R_{rg,ss}(t)$				
	$DR_{rg}(t) = \Sigma_{ss \in SSR(rg)}\ D_{rg,ss}(t)$				

9.2. PARAMETER OF DIFFERENTIAL EQUATIONS

The parameters that are part of the dynamic differential equations are

BIOLOGICAL PARAMETERS – SEI3RD & SEIMR/R-S MODELS			
SEI3RD Parameter	**SEIMR/R-S Parameter**	**Description**	**Measure Unit**
μ^N	μ^N	Natural mortality rate	fpo/day
κ	κ	The latency period of the virus before developing	day
μ	μ_{ag}	Epidemic mortality rate	fpo/day
ω	$\omega_{rg,ss}$	Probability of that a person may be contagion	prob
δ_{st}	$\delta_{ag,st}$	Probability of I_0, I_1, I_2, I_3, … recovering without worsening the clinical condition.	prob
π_{st}	$\pi_{ag,st}$	Time a patient in I_0, I_1, I_2, I_3, … recovers	day
η_{st}	$\eta_{ag,st}$	Time a patient in I_0, I_1, I_2, I_3, … to next infected state	day
ζ_{st}	ζ	Total contact free rate in I_1, I_2, I_3, …	1/day
ζ^Q_{st}	ζ^Q	Total contact confined rate in I_1, I_2, I_3, …	1/day
c_{st}	$c_{ag,st}$	Probability of contagion in free state I_1, I_2, I_3, …	prob
c^Q_{st}	$c^Q_{ag,st}$	Probability of contagion in confined state I_1, I_2, I_3, …	Prob
β_{st}	β	Transmissibility rate of an individual in state st	
β^Q_{st}		Transmissibility rate of an individual in state st on quarantine	
$\beta\delta_{t,st}$		Dynamic rate of transmissibility calculated as $\beta\delta_{t,st} = (1 - \theta_{t,st}) \times \beta^Q_{st} + \theta_{t,st} \times \beta_{st}$	
$\theta_{t,st}$		Epidemic control parameter (proportion of the st-state that circulates freely)	

9.3. MEASUREMENT SYSTEM

The measurement system is related to the information provided by government offices regarding the state of the pandemic. Measurements must be linked to the epidemiological states of the **SEIMR/R-S** model. The measurement system involves defining the measurement matrix, **H(t)**, that transforms the measurement into epidemic states, it should be noted that this matrix depends on each case since it cannot be assumed that there is a standardized model of measuring the epidemic process, there may be states that are not measured (e.g., exposed, and asymptomatic), or states that are measured in aggregate form (e.g., severe and critical).

In addition to measuring epidemic states, consideration should be given to measuring epidemic control measures (confinements, mitigations, ...) which affect the evolution of the epidemic. The basic theory of KF assumes that control corresponds to an action whose effect can be measured accurately (deterministic assumed), however, in the reality of the pandemic the effect of control measures is uncertain (for example, the number of people who actually confine as a result of a confinement order is unknown). This involves adjusting KF to consider uncertainty in control policies.

REFERENCES

[1] Jazwinski, A. H. **"Stochastic Process and Filtering Theory"**. Academic Press (1970)

[2] Poveda, G. y Mesa, O. **"Metodologías de Predicción de la Hidrología Colombiana considerando el Fenómeno El Niño/Oscilación del Sur (ENSO)"**. Revista Atmosfera No. 20.

[3] Smith, R. Informe de Investigación entregado a **ISA** (1995).

[4] Rodríguez-Iturbe, I., Valdez, J. y Velásquez, J., **"Keynote Lecture on the Applications of Kalman Filter in Rainfall-Runoff Studies"**. American Geophysical Union Chapman Conference on Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources, Pittsburg, U.S.A. (1978).

[5] Kalman, R. E., (1960). **"A New Approach to Linear Filtering and Predictions Problems"**. Trans. ASME, Journal Basic Engr.82(2) 35-345

[5A] Kalman, R.E. and Bucy, R.S. **"New Results in Linear Filtering and Prediction Theory".** Trans ASME, Ser D. J. Basic Eng. 83 (1961)

[6] Velásquez, J., **"Modelo de Kalman con Múltiples Estados para Predicción de Caudales"**. IX Congreso Latinoamericano de Hidráulica, Mérida, Venezuela (1980)

[7] Velásquez, J., **"A Multi-State Kalman Approach to Short-term Forecasting of Streamflows"**. Seminario Internacional de Hidrología y Recursos Hidráulicos, Universidad Simón Bolívar, Caracas, Venezuela (1978).

[8] Valdez, J., Velásquez, J. y Rodríguez-Iturbe, I., **"Bayesian Discrimination on Hydrology Forecasting Models based on the Kalman Filter"**. American Geophysical Union Chapman Conference on Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources, Pittsburg, U.S.A. (1978).

[9] Box, G. and Tiao, G. **"Bayesian Inference in Statisticas Analysis"**. Addison-Wesley Publishing Co. (1973).

[10] Natale, L and Todini, E. **"A Constrained Parameter Estimation Technique for Linear Models in Hydrology"**. Artículo en Mathematical Models for Surface Water Hydrology. Jhon Wiley & Sons (1977).

[11] Doodge, J. **"Linear Theory of Hydrologic Systems".** United States Department of Agriculture. Technical Bulletin 1468 (1973)

Julier, S. J. and. Uhlmann, J. K. (2004). "Unscented Filtering and Nonlinear Estimation'. PROCEEDINGS OF THE IEEE, VOL. 92, NO. 3, MARCH 2004

Moradkhani H. et. Al. (2005). "Dual State–Parameter Estimation of Hydrological Models using Ensemble Kalman Filter", Advances in Water Resources 28 (2005) 135–147.

Einicke G. A. "Smoothing, Filtering and Prediction: Estimating the Past, Present, and Future". Published by InTech Janeza Trdine 9, 51000 Rijeka, Croatia