Climate System Simulations: An Integrated, Multi-Scale Approach for Research and Decision-Making

Ángel G. Muñoz, Alfredo Nuñez and Ramón J. Cova
Centro de Modelado Científico (CMC). Universidad del Zulia.
Venezuela

1. Introduction

Climate physicists define the Climate System in terms of the different interactions taking place between hydrosphere, atmosphere, cryosphere, biosphere and lithosphere. Being the associated general thermo-hydrodynamic partial differential equation system so complex for analytical solving, it is usually integrated numerically involving a good deal of computational resources.

Moreover, today different spatial scales involve different physical parametrisations, and each forecast horizon (few days, seasonal, annual, decadal and climate change periods) of interest deserves a special treatment, mostly defined by the predictability of the Climate System and the characteristic response time of the interacting components at the corresponding temporal scale. Thus, it is presently customary to distinguish between large-, meso- and micro-scale numerical models, their descriptions ranging from global climate state to basin availability of hydrological resources. An integrated, multi-scale approach considering the output of the various models is of vital to understand at the different levels the behaviour of the Climate System, and thus offer useful tools for decision makers and stake holders.

In this chapter, after introducing a few fundamental concepts and equations in Section 2, and a hierarchy flux of information among the models for different scales in Section 3, the methodologies involving downscaling executions are discussed in detail, regarding both scientific research and policy-making applications. In section 4 we explain the usefulness of conducting several simulations (realisations) to partially reduce the inherent uncertainties. Later, in Section 5 we offer a plausible way to put into operation all the components, presenting several examples based on the experience acquired through a regional initiative known as the *Latin-American Observatory for Climate Events*. Some concluding remarks and suggestions for future research are presented in Section 6.

2. Fundamentals

In this section several fundamental aspects related to the definition of the Climate System, its temporal scales and the governing equations of the atmosphere and the oceans are discussed. While the formal definition in terms of the Climate Subsystems is standard, it is important to remark that different models tend to write the corresponding system of thermo-hydrodynamical equations in different ways, due to the employment of particular numerical methods, simplifications, coordinate systems or physical assumptions.

Nonetheless, for the sake of simplicity, we follow here an approach suggested by (Marshal et al., 2004), which enables a mathematical representation of both atmosphere and ocean dynamics via isomorphisms. This methodology has the advantage of providing an elegant way to better understand the underlying similarities between these subsystems by coupling physical quantities at the boundaries in a clear manner.

2.1 The Climate System and its scales

The Climate System (S) can be defined in terms of five subsystems, namely the atmosphere (A), hydrosphere (H), cryosphere (C), biosphere (B) and lithosphere (L) , and their complex, nonlinear interactions. These components can be described as non isolated, open and heterogeneous thermo-hydrodynamical systems, which are physically characterized by their corresponding mechanical (e.g. forces) and thermodynamical states (e.g. pressure, temperature, salinity). Actually A, H, C and B act as a cascading system involving mass, momentum and energy fluxes through their boundaries (Peixoto & Oort, 1992).
There is a general consensus in defining S as

$$S \equiv A \cup H \cup C \cup B \cup L \tag{1}$$

Due to the great complexity in the different interactions among the climate subsystems and taking into consideration the timescales of the involved phenomena, it is useful to define a hierarchy of internal systems in terms of the characteristic response time to stable state perturbations. First we consider the systems with shortest typical times, and the other components are considered to be part of the external systems. For example, the characteristic times associated to perturbations in the planetary boundary layer are in the order of minutes to hours, while for tectonic dynamic the timescale is in the order of tens of millons of years (Peixoto & Oort, 1992).
This hierarchy gives the possibility of defining the Climate System for a particular timescale in a simplified way without losing any important physical phenomena. If one is interested in weather forecast with timescales from a few hours to 1-2 weeks, then the only internal component is the atmosphere:

$$S \equiv A \tag{2}$$

considering in the boundary conditions the effects of the oceans, sea ice, land surface processess and other phenomena like being part of the external system. In a similar way, for studying the Climate System from months to decades, it is possible to write

$$S \equiv A \cup O \cup B \tag{3}$$

where the oceans O belong to the hydrosphere subsystem.
Basically, equations 2 and 3 are normally used today for hindcast simulations, weather and seasonal forecasts and climate change scenarios.

2.2 Coordinate systems

It is customary to use latitude and longitude as rectangular coordinates on a tangent plane located at a certain point on Earth's surface or above it. However, the real world is better described through non-orthogonal, *terrain-following* coordinate systems, especially if we need high resolution domains and to consider the local orography. As one should expect from this discussion, the selection of the vertical coordinate is one of the most important issues in modern numerical models. In this section a general approach is presented, reviewing briefly some typical vertical coordinates.

The altitude z (or the radial distance) is not the most convenient vertical coordinate for most applications (cf (Haltiner & Williams, 1980)). There is a general tendency to employ the pressure coordinates p or $\ln p/p_0$, being p_0 a reference level pressure (e.g. at sea level). Other common vertical coordinates are $\sigma = p/p_0$, a close modification called *eta* coordinate, the potential temperature θ and the isentropic vertical coordinate. For details the reader can consult (Haltiner & Williams, 1980) and references therein.

It is possible to generalise the discussion considering a vertical coordinate ξ, which is a single-valued, monotonic function of the altitude. In general, ξ is a function x, y, z and t:

$$\xi \equiv \xi(x, y, z, t) \tag{4}$$

where x and y are horizontal cartesian coordinates. Thus, a scalar field A can be written as

$$A(x, y, \xi, t) \equiv A(x, y, z(x, y, \xi, t), t) \tag{5}$$

Now, the partial derivative with respect s (a dummy parameter denoting x, y or t), reads

$$(\partial_s A)_\xi = (\partial_s A)_z + \partial_z A \, (\partial_s z)_\xi \tag{6}$$

where the subscript z or ξ denotes the corresponding vertical coordinate. Since the vertical derivatives correspond to

$$\partial_z A = \partial_\xi A \partial_z \xi, \tag{7}$$

so that equation (6) becomes

$$(\partial_s A)_\xi = (\partial_s A)_z + \partial_\xi A \partial_z \xi \, (\partial_s z)_\xi \tag{8}$$

This expression can iteratively be used with $s = x$ and $s = y$ to construct the gradient of A and the bidimensional divergence of the vector field **B**:

$$\nabla_\xi A = \nabla_z A + \partial_\xi A \partial_z \xi \nabla_\xi z \tag{9}$$

$$\nabla_\xi \cdot \mathbf{B} = \nabla_z \cdot \mathbf{B} + \partial_\xi \mathbf{B} \partial_z \xi \cdot \nabla_\xi z \tag{10}$$

For $s = t$,

$$(\partial_t A)_\xi = (\partial_t A)_z + \partial_\xi A \partial_z \xi \, (\partial_t z)_\xi \, . \tag{11}$$

The total derivative in terms of ξ is

$$D_t A = (\partial_t A)_\xi + \mathbf{V} \cdot \nabla_\xi A + \dot{\xi} \partial_\xi A \tag{12}$$

where **V** is the horizontal velocity and $\dot{\xi} = d\xi/dt$ is the vertical velocity.

Naturally, we also have to calculate the elements of Jacobian matrices, as we usually need to go from the rectangular coordinate system used for the computational domain, to the the real-world terrain-following coordinate system, and viceversa. The general curvilinear models are becoming more common today for they lead to better results at regional and microscale levels.

2.3 Atmosphere equations

As mentioned before, here we follow the approach advanced in (Marshal et al., 2004). The equations representing a compressible and hydrostatic atmosphere (for the non-hydrostatic
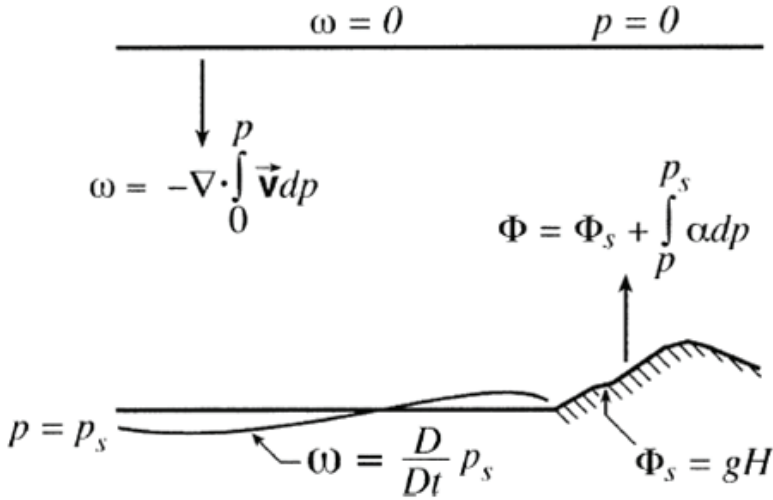
Fig. 1. Mathematical description of the atmosphere boundary conditions. Source: (Marshal et al., 2004)

case see (Adcroft et al., 2011)) in pressure coordinates are

$$D_t \mathbf{V} + f\mathbf{k} \times \mathbf{V} + \nabla_p \Phi = \mathbf{F} \tag{13}$$

$$\partial_p \Phi + \alpha = 0 \tag{14}$$

$$\nabla_p \cdot \mathbf{V} + \partial_p \omega = 0 \tag{15}$$

$$\alpha = \alpha(\theta, p) = \theta \partial_p \Pi \tag{16}$$

$$D_t \theta = \frac{Q_\theta}{\Pi} \tag{17}$$

$$D_t q = Q_q \tag{18}$$

where $\omega = D_t p$ denotes the vertical velocity in pressure coordinates, f the Coriolis parameter, $\mathbf{k}$ a versor along the vertical coordinate, $\Phi = gz$ the geopotential, α the specific volume, T the temperature, $\theta = c_p T/\Pi$ the potential temperature, $\Pi = c_p(p/p_0)^\kappa$ the Exner function and q the especific humidity. c_p corresponds to the especific heat at constant pressure, $\kappa = R/c_p$, with R the Rydberg's constant and $D_t = \partial_t + \mathbf{V} \cdot \nabla_p + \omega(\partial_p)$ is equivalent to (12) in pressure coordinates.

The terms $\mathbf{F}$, Q_θ and Q_q represent sources and sinks of momentum, heat and humidity, respectively, and are parametrised in the numerical models.

The boundary conditions are written as

$$\omega = \begin{cases} 0, & p = 0 \\ D_t p_s, & p = p_s(x, y, t) \end{cases} \tag{19}$$

The boundary condition for integrating the hydrostatic equation (14) is

$$\Phi = \Phi_s = gH, \tag{20}$$

valid for $p = p_s$, where p_s is the sea surface level pressure and H stands for a scalar field describing the terrain altitude in the lower boundary.
The surface pressure evolves according to

$$\partial_t p_s = \nabla \cdot (p_s < \mathbf{V} >) = 0 \tag{21}$$

where

$$< \mathbf{V} > = \frac{1}{p_s} \int_0^{p_s} dp\mathbf{V} \tag{22}$$

is the horizontal mean wind velocity. Figure 1 illustrates the case.

2.4 Ocean equations

The hydrostatic equations describing the dynamics of an incompressible ocean in the Boussinesq approximation and using depth coordinates are (Marshal et al., 2004)

$$D_t\mathbf{V} + f\mathbf{k} \times \mathbf{V} + \nabla_z \frac{p}{\rho_0} = \mathbf{F} \tag{23}$$

$$\partial_z \frac{p}{\rho_0} + g\frac{p}{\rho_0} = 0 \tag{24}$$

$$\nabla_z \cdot \mathbf{V} + \partial_z \omega = 0 \tag{25}$$

$$\rho = \rho(\theta, S, p) \tag{26}$$

$$D_t\theta = \frac{Q_\theta}{\Pi} \tag{27}$$

$$D_tS = Q_s \tag{28}$$

where $\omega = D_t z$ is the vertical velocity, p denotes pressure, ρ the density, ρ_0 a constant density reference, S the salinity and $D_t = \partial_t + \mathbf{V} \cdot \nabla_z + \omega(\partial_z)$ corresponds to (12) but in depth coordinates.
Again, $\mathbf{F}$, Q_θ and Q_s represent sources and sinks of momentum, heat and salinity, respectively, and are parametrised in the numerical models.
The boundary conditions are

$$\omega = \begin{cases} -\mathbf{V} \cdot \nabla H, & z = -H \\ D_t\eta = -(P - E), & z = \eta \end{cases} \tag{29}$$

with H being a scalar field describing the bathymetry ($-H$ is the sea bottom), η being the ocean surface and $(P - E)$ being the net difference between precipitation and evaporation.
The boundary condition for the hydrostatic equation in this case reads

$$p = p_s \tag{30}$$

at $z = \eta$, where we recall that p_s is the sea level pressure.
The surface elevation evolves following

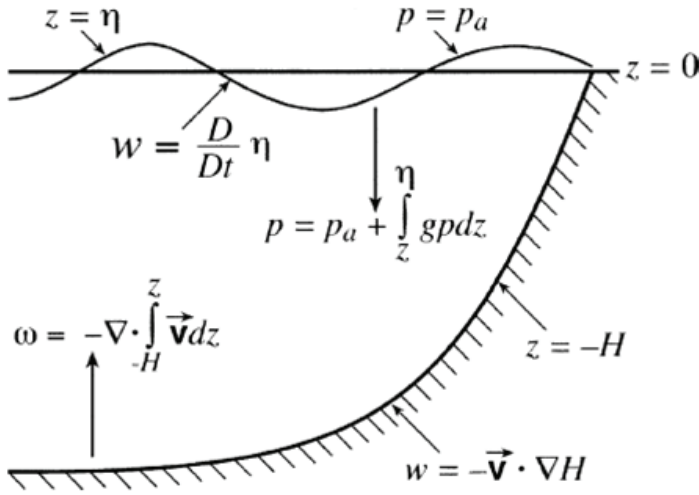$$\partial_t\eta + \nabla \cdot (H + \eta) < \mathbf{V} > = P - E \tag{31}$$

Fig. 2. Mathematical description of the ocean boundary conditions. Source: (Marshal et al., 2004)

where

$$< \mathbf{V} >= \frac{1}{H+\eta} \int_{-H}^{\eta} dz \mathbf{V} \tag{32}$$

is the horizontal mean wind velocity. See figure 2 for illustration.

2.5 Isomorphisms and physical coupling

By taking a look at the previous subsections it is easy to recognize the change in variables for writing the isomorphism. Choosing r as the new isomorphic variable, such that (Marshal et al., 2004)

$$
\begin{array}{ccc}
ocean & \longleftrightarrow isomorphism \longleftrightarrow & atmosphere \\
z & r & p \\
\omega & \dot{r} & \omega \\
\frac{p}{\rho_0} & \phi & \Phi \\
-g\frac{\rho}{\rho_0} & b & -\alpha \\
\theta & \theta & \theta \\
S & s & q \\
\eta & r'_s & p_s \\
P-E & P_r & 0
\end{array}
\tag{33}
$$

the boundary conditions (19,20,29,30) automatically remain isomorphic too. An illustration is depicted in Figure 3.

Then, the general equations are

$$D_t \mathbf{V} + f\mathbf{k} \times \mathbf{V} + \nabla_r \phi = \mathbf{F} \tag{34}$$

$$\partial_r \phi - b = 0 \tag{35}$$

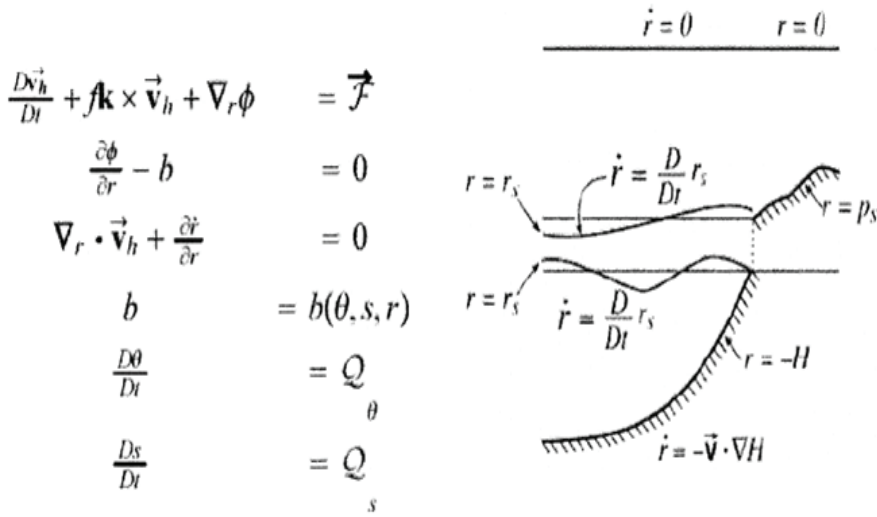$$\nabla_r \cdot \mathbf{V} + \partial_r \dot{r} = 0 \tag{36}$$

$$\frac{D\vec{v}_h}{Dt} + f\mathbf{k} \times \vec{v}_h + \nabla_r \phi = \vec{\mathcal{F}}$$

$$\frac{\partial \phi}{\partial r} - b = 0$$

$$\nabla_r \cdot \vec{v}_h + \frac{\partial \dot{r}}{\partial r} = 0$$

$$b = b(\theta, s, r)$$

$$\frac{D\theta}{Dt} = Q_\theta$$

$$\frac{Ds}{Dt} = Q_s$$

Fig. 3. Isomorphic relations and coupling. Source: (Marshal et al., 2004)

$$b = b(\theta, s, r) \tag{37}$$

$$D_t\theta = Q_\theta \tag{38}$$

$$D_t s = Q_s \tag{39}$$

where r is the vertical, isomorphic coordinate. The total derivative takes the form

$$D_t = \partial_t + \mathbf{V} \cdot \nabla A + \dot{r}\partial_r, \tag{40}$$

and

$$\nabla = \nabla_h + \mathbf{k}\partial_r, \tag{41}$$

with ∇_h being applied on the horizontal surfaces (constant r) and $\mathbf{k}\partial_r$ acting along the vertical coordinate. ϕ is associated with pressure or geopotential, b is the buoyancy, s is the especific humidity for the atmosphere or the salinity for the ocean, $\mathbf{F}$ represents the external forcings and dissipation on $\mathbf{V}$, Q_θ the external forcings and dissipation on θ and Q_s the ones on s.

The terms $\mathbf{F}$, Q_θ and Q_s are provided by physical parameterizations of the subgrid scale atmosphere and ocean fluxes.

The corresponding boundary conditions are

$$\dot{r} = \begin{cases} -\mathbf{V} \cdot \nabla R_{fixed}, & z = R_{fixed} \\ D_t r_s - P_r, & r = R_s \end{cases} \tag{42}$$

where

$$R_{fixed} = \begin{cases} -H, & \text{at sea bottom} \\ 0, & \text{at atmosphere top} \end{cases} \tag{43}$$

defines the location of the fixed boundary surface, while

$$R_s = R_0 + r_s \tag{44}$$

defines the location of the moving boundary surface, being

$$R_0 = \begin{cases} 0, & \text{at sea surface} \\ R_0(x,y) = p_s(x,y), & \text{at terrain surface} \end{cases} \tag{45}$$

The reference location of the moving boundary surface and its deviations are described by

$$r_s = \begin{cases} \eta, & \text{at sea surface} \\ p_s, & \text{at terrain surface} \end{cases} \tag{46}$$

Moreover,

$$P_r = \begin{cases} P - E, & \text{at sea surface} \\ 0, & \text{in the atmosphere} \end{cases} \tag{47}$$

is the volumetric flux through r_s.

It is important to remark at this level how the system of equations presented are so complex that they need to be solved numerically in the general case. The last section shows a natural way to couple the needed quantities at the boundaries (see figure 3), which sketchs the coupling between ocean and atmosphere, the two most dynamic components of the Climate System. Bearing this review of the fundamental equations in mind, it will be easy to consider the particular expressions used in the different numerical models.

3. An integrated, multi-scale approach

The rationale behind climate modeling is to mathematically express the different interactions present in the Climate System. This is ussually accomplished by means of

- Dynamical climate models
- Statistical climate models.
- Hybrid climate models.

The dynamical approach tries to describe the observed phenomenology using physics (Anderson, 2008), with sets of equations like the ones presented in the previous section. On the other hand, the statistical approach looks for the same goal, but it tries to recognize spatio-temporal patterns in the historical dataset of at least two different variables (usually known as the *predictor* and the *predictand*); it employs these patterns to build a mathematical model (Mason & Baddour, 2008), capable of reproducing the historical behaviour and to forecast the future, in analogous way to the dynamical approach. Finally, mixtures of both methodologies are also possible, entailing different kinds of hybrid models. The climate numerical simulations are the execution of these models. Executions reproducing the past are called *retrospective simulations* or *hindcasts*. Weather and seasonal forecasts, and climate change scenarios are examples of climate simulations looking at the future.

Each of the aforementioned approaches has its own pros and cons. For instance, the dynamical models need to solve the set of equations for each cell in the selected spatial domain, and thus they consume a lot of computational resources and are slow compared with the common statistical models. The latter, in turn, need homogeneous and quality datasets for the variables under study, an important issue in several regions of the planet. However, once the temporal

domain has been selected, the speed of the computations for all the methods is strongly dominated by the spatial resolution, even when in general the statistical methods will run many times faster than the dynamical ones. A coarse resolution implies less cells where to compute the models (physical, statistical or hybrid). The coarser the domain, the faster the computations, but the less the ability to accurately represent in a point the state of the Climate System.

In this section we review some of these aspects and discuss how to construct a multi-scale approach for climate research and decision-making.

3.1 General Circulation Models: Coupled vs non-coupled

The planetary-scale, dynamical numerical models are called General Circulation Models (GCMs), and they can be coupled (CGCMs) or non coupled (e.g. atmospheric GCMs - AGCMs, and oceanic GCMs - OGCMs), with typical spatial resolutions in the order of 2.5^o to 1°. The coupling makes reference to linking different models for mimetizing the various interactions of the Climate System components, following the ideas presented in section 2.1. Nonetheless, as it has been discussed previously, S can be simplified (see equation (2)) when considering certain temporal and spatial scales. Such cases are usually simulated by means of non-coupled GCMs, and due to their shorter execution times and some key advantages (see Mason (2008) and references therein), they have been the global dynamical models most widely utilised.

Coupled and non-coupled models may or may not coincide in their results, depending on several aspects. One of the main differences, as can be deduced from the previous discussion, is the inclusion or not of the various feedbacks. As an example, consider the simulation of the interactions between the atmosphere and the ocean on a seasonal (e.g. 3 months) scale. The coupled models are designed to guarantee the feedbacks, but the stand-alone atmospheric models consider the ocean effects via the boundary conditions (which in general evolve in time), without permiting any updates to the sea surface coming from the new atmospheric states. In other words, for these non-coupled atmospheric models to run it is first necessary to subscribe the evolution of the boundaries in a separate tier. This is why sometimes the stand-alone models are also called *two-tiered* models. One way to account for the differences among one-tiered and two-tiered simulations is to look at the energy balance between the surface and the top of the atmosphere. The radiative imbalance is usually less than 0.5 Wm^{-2}, which is considered acceptable for seasonal timescales. However, the Climate System's slow component interactions can provide severe bias that must be considered in detail (Hazeleger et al., 2010).

Amongst the main CGCMs used today are the CCSM-CESM (National Center for Atmospheric Research-NCAR Community Earth System Model (Blackmon et al., 2001)), COSMOS (Max Planck Institute-MPI Community Earth System Models (Roeckner et al., 2006)), HadGEM (Hadley Centre Global Environmental Model (Johns et al., 2006)) and CFS (NCEP Climate Forecast System (Saha et al., 2006)). Older and newer versions are available at a number of research and forecast centers around the world.

3.2 Seasonal forecast methodologies for GCMs

Both hindcast and forecast are important for climate simulations in research and decision-making. Retrospective simulations make use of grids filled with observations (e.g. gauge stations, satellite) for providing the GCMs' boundary conditions along the past period of interest while, obviously, there is no direct analog for a simulation of the future. The key idea here is that the GCM integrates exactly the same set of equations for the past or the future, the difference being on how the boundary evolves: in a hindcast, the observations are used for

describing its evolution, while in a forecast there are several ways of predicting the behaviour of the boundaries.

Clearly, the sea surface temperature (SST) is one of the main modulators of the Climate System (see for example Peixoto & Oort (1992) and references therein). Indeed, modern GCMs need both the SST and the ice cover as boundary conditions, whose spatial patterns are known to vary on a monthly to seasonal basis. The way they are constructed to evolve defines the forecast methodology.

In the case of modern CGCMs, atmosphere, ocean and land-surface models run synchronously and interactively to describe not only the different component states, but also the boundary conditions evolution, simultaneously.

On the other hand, in two-tiered forecasts the tier-1 is associated to the SST and ice cover forecast, whereas the tier-2 is the AGCM. There are different ways to forecast in the tier-1. Here we study some principal methodologies for SST.

The simplest method -but not very reliable- is to prescribe the SST to behave as its climatological value, i.e. its monthly mean value over the last 30 to 50 years. Thus, the climatological January corresponds to all the 30 to 50 January values averaged over. This is done for each ocean cell (or ocean basin cell) each month. Naturally, this prescription is most probably destined to fail whenever years or seasons exhibit extreme events, like ENSO for instance.

The method most widely used is called *persistence* (Li et al., 2008). Also known as 'serial dependence' (Wilks, 2006), it assumes that the SST anomalies (aSST, from now on) in the preceding month or season will persist along the forecast period. Thus the *persisted SST* or *pSST* forecast for the following months is constructed as their corresponding climatological value plus the preceding ($t = 0$) month observed anomaly $aSST_0$. Over the extratropical oceans, it is customary to include a damping coefficient with an e-folding time of 3 months,

$$aSST(t) = aSST_0 e^{-t/3}, \qquad t = 0, ..., 3 \qquad (48)$$

Another methodology is called *constructed analogues* or *caSST* (Van den Dool, 1994). Suggested in 1994, the rationale is to write the preceding month (base) aSST as a linear combination of the same month aSST for a long past period (at least 30 years, excluding the year in course). The coefficients of the linear combination are computed via classical least-squares minimization; the resulting weights are used to forecast the target months in terms of the subsequent base months in the past period. It is important to note that the constructed analogue is the same linear combination for all leads, i.e the weights persist.

Finally, we may find the aSST as tier-1, forecasted by a CGCM (see (Li et al., 2008)). For example, the Andean Observatory (Muñoz et al., 2010), a regional initiative which produces seasonal forecasts in the Andean countries, employs aSST from the CFS. Let us refer to this case as *cfsSST*.

Figure 4 compares two of the foresaid *cfsSST* and *pSST* methods.

Similar approaches can be followed for the ice cover, but very frequently it is prescribed to follow the climatological monthly variation. Figure 4 compares two of the foresaid cfsSST and pSST methods. Both realisations provide similar distributions and magnitudes of the accumulated seasonal precipitation. However, the differences between them are important, for they allow probable alternative precipitation behaviours in the final ensemble product (see Section 4).

Once the tier-1 is provided, the AGCM (tier-2) integrates the associated equations and gives all the required atmospheric variables.
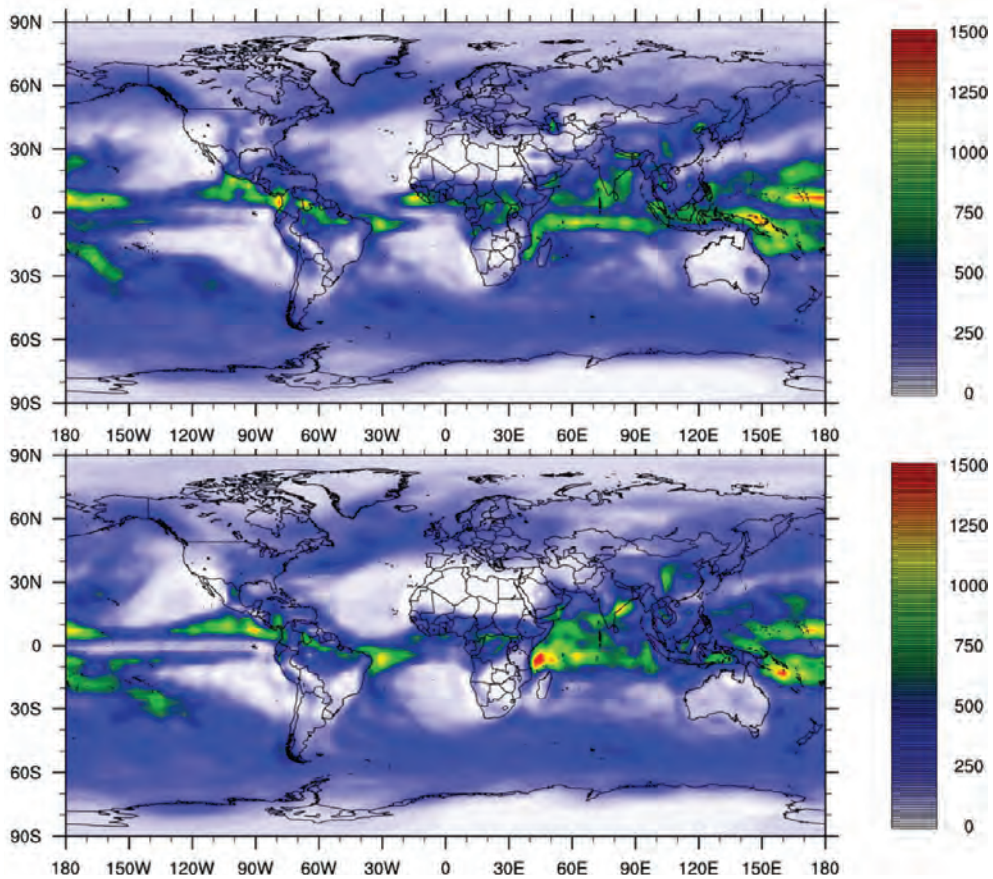
Fig. 4. Accumulated precipitation (in mm) for Apr-May-Jun 2011 using CAM and two forecast methodologies discussed for two-tiered models: *cfsSST* (above) and *caSST* (below). See main text.

3.3 Interpolation and Downscaling

Normally, the GCM output does not describe all the physical processes in one especific region of the planet. As pointed out before, different physical mechanisms are associated to different spatio-temporal scales, and so we normally need to increase either the spatial or the temporal resolution of the GCM output, or both. This resolution increase can be achieved by means of interpolations, statistical estimators or downscaling.

Interpolations change the resolution using a wide variety of mathematical expressions to provide new points on an array already in existence. For example, if the GCM output grid possesses a spatial resolution of 1°, this means that the mean distance between the main node points (i.e. the intersection of latitude and longitude lines) is about 111.11 km at the equator. An interpolation will provide more new node points to the grid, say each one at 56 km, thus increasing the spatial resolution. The main concern is that at this new resolution, different interpolation rules provide different values of the physical variable at the new points. Which rule is better? In addition, there is no guarantee that the interpolated fields will satisfy the

original set of physical equations that provided the variables (Haltiner & Williams, 1980). A common interpolation rule nowadays is the bilinear interpolator, a 2D generalization of the standard linear interpolator.

Commonly included as interpolators (the difference is not always clear in the literature), the estimators use statistics (e.g. variance/covariance matrix, stochastic models with spatial dependence, variograms, unbiased linear regressions, see (Cressie, 1993)) to increase the spatial resolution of the fields. One of the most employed statistical estimators is the Krigging method (Cressie, 1993).

Finally, there are downscaling methods. In this case the spatial and temporal resolution increase is done through physical equations (dynamic downscaling, see for example (Murphy, 1999)) or statistical methods like principal component analysis or canonical correlation analysis (statistical downscaling, see (Mason & Baddour, 2008; Murphy, 1999) and references thererin).

The general consensus is to proceed with downscaling methods instead of interpolations/estimators, even when the latter are in general many times faster than the former. The downscaling needs analysis (observed data on grids) or GCMs fields to produce the higher resolution variables.

In statistical downscaling, models are constructed recognizing relationships between sets of variables through statistical analyses of time series, from historical observations or GCM output. For example, one can downscale the precipitation provided by a global model (the predictor) using the historical precipitation reported by rainfall gauge stations (the predictand). This process is accomplished, for instance, by constructing statistical models with principal component regressions or, more often, canonical correlation analysis (for details see (Mason & Baddour, 2008)). Sometimes this particular downscaling aimed at correcting the dynamical model output using observed data is known as Model Output Statistics (MOS), but indeed the MOS involves a more general set of processess (Wilks, 2006).

Another possibility is to downscale GCM outputs using higher resolution dynamical models, known as Regional Climate Models (RCMs). In this case, the physical equations are solved using as boundary and initial conditions the fields provided by the global model, ensuring physical consistency among the variables at the new spatial and temporal resolutions. However, it is important to bear in mind that at different scales the dynamical models require different *physical parametrisations* in order to represent sub-grid scale phenomena. There are often several parametrisations for the same spatial scale, and several possible configurations for the dynamical downscaling models, aimed at resolving the physics for the region of interest.

In the case of multi-scalar phenomena, a nesting procedure is used for RCMs, where the mother (or father) domain is fed by the GCM output, provides the necessary information for their sons and so on. Due to scarcity of computational resources, in the past a one-way nesting approach has been used. This means that the information flux goes only from mother to son. Today, a two-way nesting feeds back the mothers with the information of higher resolution fields. Several technical issues (e.g. the domain's location and size and its *buffering zone* configuration) must be taken into account in all these cases to ensure that the mass and energy fluxes through the lateral boundaries are adequately considered. For details see (Liang & Kunkel, 2001).

To illustrate the differences among the precipitation spatial patterns provided by the CAM (approximately at 2.5^o resolution) and the WRF (Skamarock et al., 2005) at 0.27^o resolution, see Figure 5. Here, their specific configurations correspond to those in operation at the Andean Observatory (Muñoz et al., 2010). The RCMs seem to be more adequate than the GCMs to simulate and forecast climate events because they recognise multi-scale patterns and higher
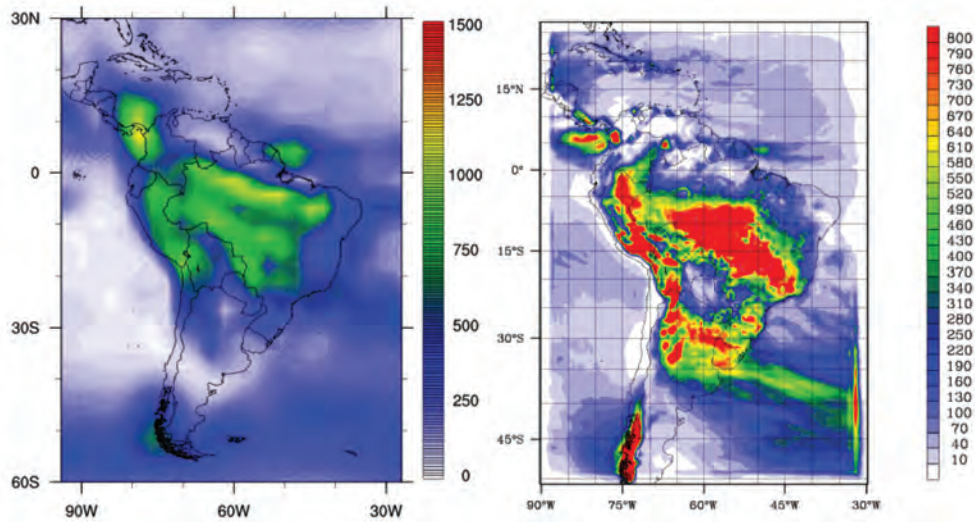
Fig. 5. Mean rainfall fields for season Jan-Feb-Mar 2011 provided by CAM (left) and the downscaling model WRF in its climate version (right). Source: the Andean Observatory (Muñoz et al., 2010)

.

resolution, but especial care must be exercised since the downscaling process may increase the inherent uncertainties.

3.4 Hierarchical flux approach

Hitherto, we have discussed how to perform climate simulations and forecasts using different tools, and how they can be related with each other. It is clear that in order to have an integrated, multi-scale approach for research and decision-making, especially if dealing with basin scale applications or the atmospheric dynamics over complex terrain, we need a system that considers all the representative interactions. Such a technique is described below.

The simplest way to take into account the different phenomena at the various scales in the Climate System is through a unified, fully coupled GCM at very high resolutions, such that there is no need to use any parametrisations because the physics can be resolved explicitly by the model equations. Despite some good efforts (Hazeleger et al., 2010), the computational (e.g. time execution and infrastructure) costs for such a seamless Earth System Model are so high that it will take several years before it becomes an operational standard. This is why it is necessary to stablish a hierarchical flux of information between the different models, and to execute them in a sequential mode to built up a multi-scale simulation system.

A three-level hierarchy can thus be defined (see Figure 6). Level I involves the various GCMs that can be used, at a coarse resolution, to provide the initial and boundary conditions for the RCMs and the predictors for the statistical downscaling models that make up the Level II. Tailored application models (e.g. Malaria (MacDonald, 1957; Recalde, 2010), ecodynamical (Tapias, 2010), fire (Chandler et al., 1983), drought (Palmer, 1968; Svoboda et. al, 2002) or off-line hydrological models (Liang & Xie, 2001)) using the Level II output as part of their own input, belongs to the hierarchy's Level III.
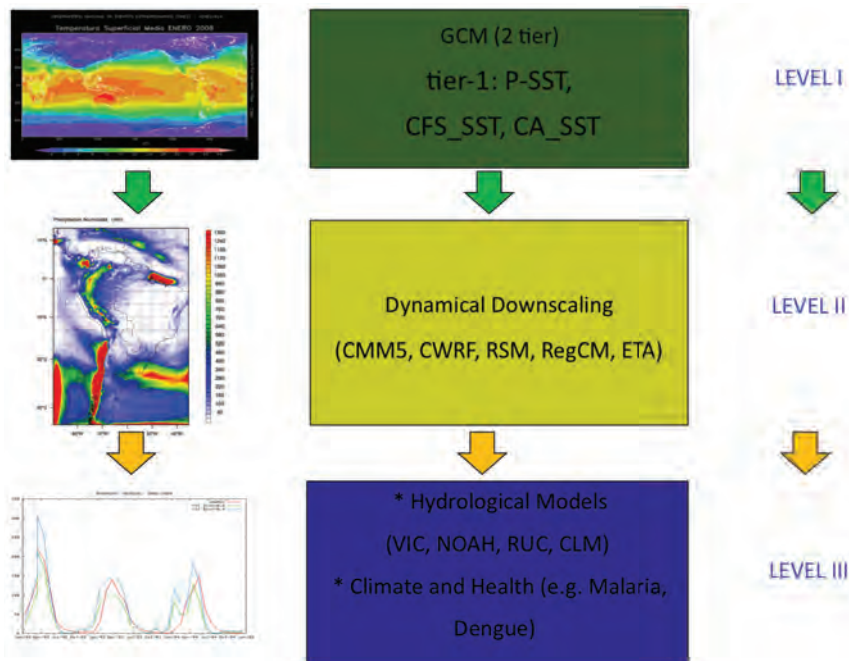
Fig. 6. The three level, two-way hierarchical information flux between climate related models. Source: (Muñoz et al., 2010)

.

We emphasise that the same hierarchical information flux works for both weather (short-term) and climate (seasonal to long-term) simulations. Statistical, dynamical or hybrid models can be present at any level.

This simple and efficient structure allows higher-level (i.e. levels II and III) models to run using as *first guess* the output of the preceding level, but the opposite must be also considered. Lower-level models should also be updated as in *upscaling* applications, providing a two-way flux that enables a feedback among the different models.

4. Multiple feasible futures

Clearly, forecasts are model dependent. Different models and parametrisations will determine different probable futures, all of them physically acceptable if using dynamical models. Some will behave better than others when compared with observations, depending on several factors (e.g. the adequacy of the physical phenomena description in a dynamical model or the boundary conditions employed). Therefore, useful products for decision-makers should include, aside from the forecast maps themselves, additional information about their confidence and uncertainty.

The uncertainty of the final products can be decreased if several *realizations* are used, each one corresponding to a different and independent model execution (i.e. a feasible future). The idea is to *ensemble* in a final product as many independent realizations as possible, using statistical weights for each member that are defined in terms of how well each one represents the observed values in a certain period. It is a common practice, however, to

employ equally weighted members in the final ensemble. For weather applications of the ensemble methodology the reader can review (Toth & Kalnay, 1993), and for seasonal forecast applications (Li et al., 2008) and references therein.

The different realizations can be produced in several ways. For example, they can be constructed in terms of perturbations to some initial state, using different methodologies (like the ones explained in section 3.2) for the tier-1 in seasonal forecast, or even employing diverse combinations of physical parametrisations. Figure 7 sketches a two member behaviour for the SST evolution in an AGCM equatorial grid cell. Each member originally differed in the SST cell by only 10^{-3} K. Due to the *butterfly effect*, after a few weeks we will likely see important differences in the temporal behaviour of the variable, and not only for that cell. Climate models fed with these two SST fields as boundary conditions will provide different members in the ensemble forecast.

For climate simulations, it has been shown (Li et al., 2008) that the use of different methodologies for the tier-1 offers better results (fewer uncertainties) in the final ensemble than the simulations associated with only one methodology.

5. The Latin American Observatory: An operational research and forecast system

To illustrate an operational research and forecast system which provides useful tools for decision-makers and stake-holders, in this section the *Latin America Observatory for Climate Events* structure will be discussed briefly. Its goals are similar to those of the Andean Observatory (Muñoz et al., 2010), but in this case the participation of all interested institutions in the Latin American countries is fully brought forth and supported. The idea is to facilitate scientific tools for the decision-makers, thus enabling the continuous interaction between research (universities and centers in the region) and operational activities (basically the National Weather Services and related institutions). The present coordinator of this project is the Centre for Scientific Modelling (Centro de Modelado Científico - CMC, in Spanish) at the University of Zulia, Venezuela.

The Observatory, known as OLE^2, currently has got a number of methodologies:

- Dynamical Weather Forecast

  At present, OLE^2 offers 72-hour weather forecasts on a daily basis using the high resolution downscaling models MM5 (Michalakes, 2000) and WRF (Skamarock et al., 2005). The GFS (Kalnay et. al, 1990) 3-hourly outputs and assimilation of SYNOP, METAR and TEMP reports are used as initial conditions. Each country determines the best set of model parametrisations, typically running at resolutions of 30 km and higher. The model outputs are valuable for the forecasting processes in countries where the Andes Mountain Chain provides complex disturbances that frequently GFS and other global models cannot resolve.

- Dynamical Seasonal Forecast

  The NCAR Community Atmospheric Model version 3.1 (CAM3) (Collins et al., 2006) has been configured at T42L26 resolution at CMC by the Atmospheric Model Intercomparison Project (AMIP); it runs through the Green House Gases (GHGs) with monthly variability from 1966 to present. The first 5 years have been discarded for spin-up reasons. The selected climatology corresponds to the 1971-2000 period.

  The current seasonal forecast methodology is sketched in Figure 8. On a monthly basis, the CAM runs 6 ensemble members, where as tier-1: (a) two of them follow the persisted SST e-folding methodology (psst, see for example, (Li et al., 2008)), (b) two members use the SST forecast of the CFS model (cfssst, (Saha et al., 2006)), and (c) two realizations are
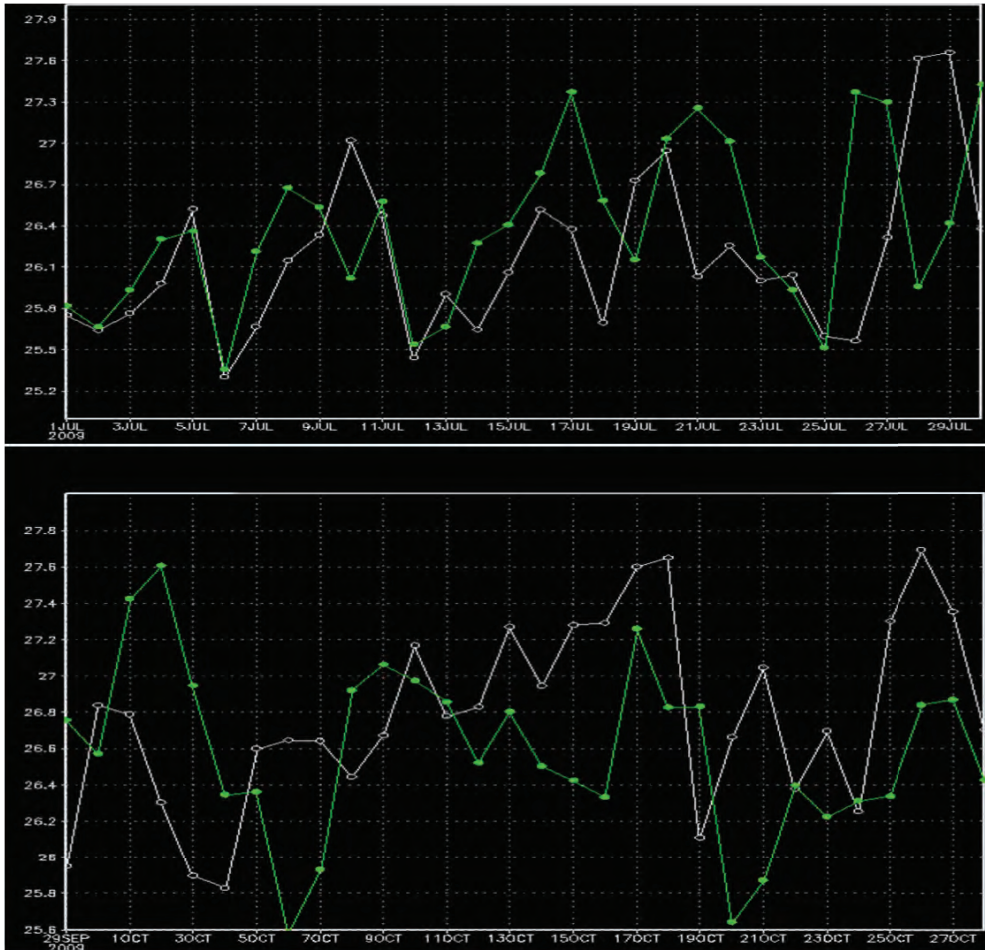
Fig. 7. SST behaviour for Jul 2009 (above) and Oct 2009 (below) using two CAM members for a cell located in the equator and the dateline (180W). The green lines corresponds to a 10^{-3} K perturbation of the climatological value (in white). After three months the differences are notable.

obtained following the constructed analog (casst, (Van den Dool, 1994)) methodology. For all members the lead's monthly ice fraction coverage is described by the climatological values. For each member's output, the necessary initial and boundary conditions are extracted and written in the special (intermediate) format requested by the climatic versions of MM5 and WRF (CMM5 and CWRF, from now on), and are then available for the Andean NWSs through the OLE[2] web portal (http://ole2.org), which has been totally built with Open Source resources by CMC developers.

Each NWS downloads the required files to execute the models in their own computational infrastructures and, since January 2010, using two different sets of physical parametrisations per model. Thus, a multi-parametric multi-model ensemble is produced
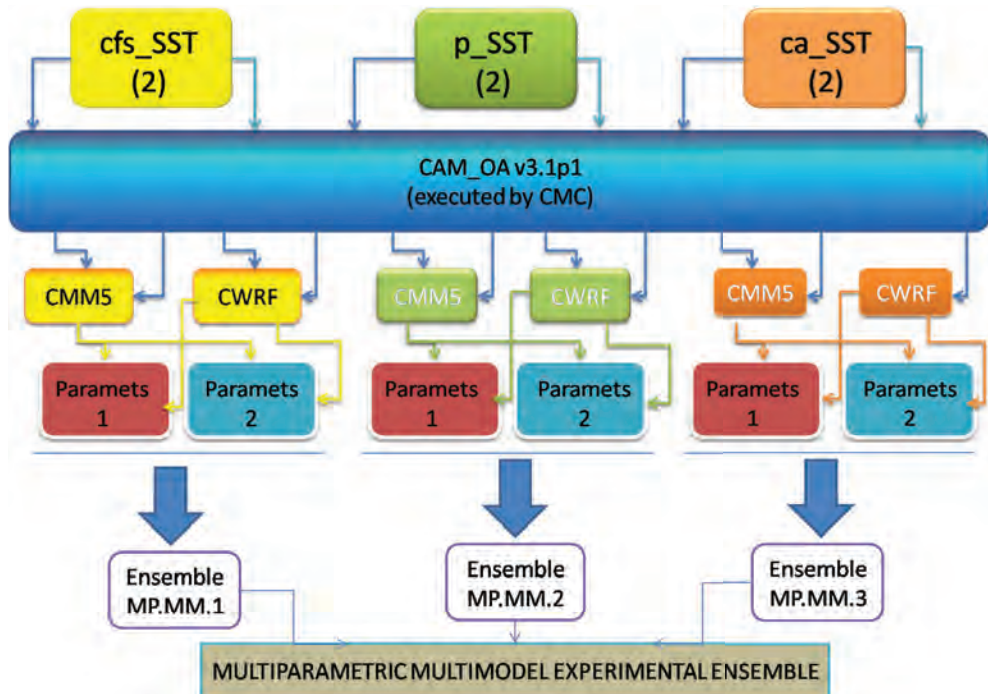
Fig. 8. Multi-parametric, multi-model ensemble employed for the seasonal forecast in the Latin American Observatory.

for each country, and then uploaded to the OLE[2] web portal for its publication after internal filters and discussion. Figure 2 depicts an OLE[2] seasonal precipitation anomaly map for South America, with the corresponding observed rainfall anomaly for comparison. These products are also used in each NWS for the generation of agricultural risk maps as well as products and tools for decision makers.

• Oceanographic High Resolution Forecast

The Regional Oceanic Modeling System (ROMS-AGRIF, (Penven et al, 2007)) has been configured for a computational domain in the Eastern Pacific. At present, the boundary and initial conditions are provided by the ECCO Consortium (Estimating the Circulation and Climate of the Ocean, (Stammer et al., 1999)) and the GFS (Kalnay et. al, 1990). The OLE[2] runs daily ROMS for a 5-day high resolution (30 km) forecast, sharing in the web portal products like SST, surface salinity, vertical velocities (upwelling and downwelling) and marine currents. This kind of product is very useful as a basis for fishery maps, indicating locations of nutrient-rich areas due to upwelling processes.

This OLE[2] component has been developed at CMC in collaboration with the *Comisión Permanente del Pacífico Sur* (CPPS) in order to set up the same methodology developed by the NWS for the Marine and Coastal Services of Colombia, Ecuador, Peru, and Chile.

• Dynamical Hydrological Forecast

The Dynamical Hydrological Forecast (Level III) process is carried out at OLE[2] by coupling the NOAH Land Surface Model (Schaake et al., 1996) with the Level II models, or directly

using the latter's forecast precipitation, temperatures and wind outputs into the Variable Infiltration Capacity (VIC) Model of (Liang & Xie, 2001). VIC is a macroscale (typical cell resolution > 1 km), semi-distributed hydrologic model that solves full water and energy balances. At OLE^2 the VIC is specifically configured for each basin of interest (the resolution depends on the selected basin) with the corresponding soil and vegetation type data.

For both procedures (coupled LSMs or uncoupled VIC Model), a bias correcting calibration procedure is applied to the raw output using historical, local streamflow data as reference. After the calibration stage, the final outputs can be considered as a main tool for the corresponding Early Warning System in the countries involved.

- Other Applications

Other applications include products related with droughts, floods, fires and ecosystem dynamics. In the case of droughts (Palmer, 1968), indices are employed, while a composite map between runoff and hydrologic capacity of model cells are used to forecast possible floods. Likewise, the (Chandler et al., 1983) index is utilised as a measure of joint probability of fire occurrence and propagation.

Climate and Health applications are focused mainly on malaria seasonal predictability for northwestern South America using the model studied in (MacDonald, 1957). Given the necessary entomological and epidemiological parameters, the high resolution output at OLE^2 supplies the climate information for running this epidemiological tool.

Finally, a new framework is related to Ecosystem Dynamics, especially Lemna (duckweed) population dynamics. In 2004 an important duckweed bloom took place in Maracaibo Lake (Tapias, 2010), the South American largest lake, bringing economic (e.g. fisheries) and health related (e.g. necrotic Lemna at lake shores produce an increase of diseases) problems to human populations in those coastal zones. Recently, the CMC provided an application known as CAVEL ((Tapias, 2010)) that makes use of MODIS VIS and IR data (Barnes et al., 2002) for providing normalized vegetation index (NDVI) maps, and time series of total surface coverage.

6. Concluding remarks and future research

In this article we have studied a general methodology for research and forecasting. It utilises a hierarchical flux of information between different models, aimed at providing useful, easy to understand scientific tools for decision makers and stake-holders. It can be decomposed into three levels. The first involves either coupled or non-coupled General Circulation Models which supply the initial and boundary conditions for the second level, namely, Regional Models, which employ statistical or dynamical downscaling. The third level feeds on the information given by level 1, offering tailor-made applications for decision makers, ranging from hydrological resources availability in a basin to ecosystem dynamics or vector borne diseases related to climate.

The various applications take into consideration different climate phenomena occurring at several spatial and temporal scales. The goal is to develop really useful tools for policy making, where high spatial resolutions are often needed for short-term, seasonal, decadal and climate-change scales. These efforts require heavy computational resources which, fortunately, are becoming more commonly available nowadays in Climate Centres, and sometimes through regional collaborations like the Latin American Observatory of Climate Events, as has been stressed throughout this article.

In the near future the present methodology is likely to change through the use of a seamless Earth System Simulator, capable of executing hundreds of petaflops which support the

numerous time scales and integrate the physical equations resorting to new meshes and more powerful numerica1l schemes. With the help of such systems, climate science should advance significantly over the next few years.

7. References

Anderson, D. (2008). Overview of Seasonal Forecasting, in *Seasonal Climate: Forecasting and Managing Risk*, Springer Science+Business Media B.V., pp 45-66.

Adcroft, A.; Campin, J.M.; Dutkiewicz, S.; et al. (2011). MITGCM USer Manual. Available at `http://mitgcm.org/public/r2_manual/latest/online_documents/manual.pdf`

Barnes, W. L., and V. V. Salomonson, 1992: A global imaging spectroradiometer for the Earth Observing System. Critical Review, *Optical Technologies for Aerospace Sensing*, Vol. CR 47, 1?23.

Blackmon, M. B., Boville, B., Bryan, F. et al. (2001) The Community Climate System Model. *BAMS*, 82, 11, 2357-2376.

Chandler C., P. Cheney, P. Thomas, L. Trabaud & D. Williams (1983) Fire in Forestry, in Vol. I: *Forest Fire Behavior and Effects*. Jhon Wiley & Sons, New York, NY. 450 pp.

Cressie, N. (1993) *Statistics for Spatial Data*, Wiley, New York.

Collins, W. D., and Coauthors, 2006: The formulation and atmospheric simulation of the Community Atmosphere Model Version 3 (CAM3). *J. Climate*, 19, 2144?2161.

Haltiner, G. & Williams, R. (1980) *Numerical Prediction and Dynamic Meteorology*, 2nd Edition, John Wiley & Sons, NY.

Hazeleger, W., Severijns, C., Semmler, T., et al., 2010: EC-Earth: A seamless Earth System Prediction Approach in Action, *BAMS*, 9, 10, 1357-1363.

Johns, T.C, Durman, C.F., Banks, H.T. et al. (2006) The new Hadley Centre climate model HadGEM1: Evaluation of coupled simulations. *Journal of Climate*, 19, pp. 1327-1353

Kalnay, E., M. Kanamitsu, and W. E. Baker, 1990: Global numerical weather prediction at the National Meteorological *Center. Bull. Amer. Meteor. Soc.*, 71, 1410 - 1428.

Li, Sh., Goddard, L., Dewitt, D. (2008). Predictive Skill of AGCM Seasonal Climate Forecasts Subject to Different SST Prediction Methodologies, *Journal of Climate*, 21, 10, 2169.

Liang, X-Z & Kunkel, K. (2001). Development of a Regional Climate Model for U.S. Midwest Applications. Part I: Sensitivity to Buffer Zone Treatment, *Journal of Climate*, 14, pp. 4363-4378.

Liang, X., & Xie, Z. (2001) A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models, *Advances in Water Resources*, 24(9-10), pp. 1173-1193.

Marshall, J.; Adcroft, A.; Campin, J. M.; Hill, C. & White, A. (2004). Atmosphere Ocean Modeling Exploiting Fluid Isomorphisms. *Monthly Weather Review*. 132, 2882-2894.

Mason, S. (2008). From Dynamical Model Predictions to Seasonal Climate Forecasts, in *Seasonal Climate: Forecasting and Managing Risk*, Springer Science+Business Media B.V., pp 163-201.

Mason, S. and O. Baddour (2008). Statistical Modelling, in *Seasonal Climate: Forecasting and Managing Risk*, Springer Science+Business Media B.V., pp 163-201.

MacDonald, G. (1957) The Epidemiology and Control of Malaria. London, U.K. Oxford University Press.

Michalakes, J. (2000) The same-source parallel MM5. in *Sci. Comput.*, 8, 5?12

Muï£¡oz, ï£¡. G., Liï£¡pez, P., Veliï£¡squez, R., et al., 2010: An Environmental Watch System for the Andes Countries: El Observatorio Andino. *BAMS*, 91, 1645?1652.

Murphy, J. (1999). An evaluation of statistical and dynamical techniques for downscaling local climate, *Journal of Climate*, 12, pp. 2256-2284.

Palmer, W. C. (1968) Keeping track of crop moisture conditions nationwide: The new crop moisture index. *Weatherwise*, 21, pp. 156-161.

Peixoto, J. & Oort, A. (1992) *Physics of Climate*, American Institute of Physics, NY.

Penven, P., P. Marchesiello, L. Debreu, and J. Lefevre, 2007: Software tools for pre- and post-processing of oceanic regional simulations. *Environ. Modell. Software*, 23, 660?662.

Recalde, G. C. (2010). Malaria predictability experiment for the Ecuadorian coast using dynamical models (Thesis). Environmental Enginering. Universidad Agraria del Ecuador. 72 pp. Available at `http://cmc.org.ve/descargas/Recalde2010.pdf`

Roeckner, E., R. Brokopf, M. Esch, et al. (2006). Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, *Journal of Climate*, 19, 3771-3791.

Saha, S., Nadiga, S., Thiaw, C. et al. (2006). The NCEP Climate Forecast System. *Journal of Climate*, 19, 15, 3483-3517.

Schaake, J. C., V. I. Koren, Q.-Y. Duan, K. Mitchell, and F. Chen, 1996: Simple water balance model for estimat- ing runoff at different spatial and temporal scales. J. Geophys. Res., 101 (D3), 7461?7475.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers (2005). A Description of the Advanced Research WRF Version 2, NCAR Tech. Note NCAR/TN-468&STR, 88 pp.

Stammer, D., and Coauthors, 1999: The consortium for estimating the circulation and climate of the ocean (ECCO)?Science goals and task plan. ECCO Consortium Rep. 1, 24 pp.

Svoboda, M. D., et al. (2002) The Drought Monitor, *BAMS*, 83, pp. 1181-1190.

Tapias, E. (2010) CAVEL: An automated system for Lemna recognition and analysis in the Lake Maracaibo.(Thesis). Facultad de Ciencias. La Universidad del Zulia, 104 pp. Available at `http://cmc.org.ve/descargas/Tapias2010.pdf`

Toth, Z. & Kalnay, E. (1993) Ensemble Forecasting at NMC: The generation of perturbations, *BAMS*, 74, pp. 2317-2330

Van Den Dool, H. (1994). Searching for analogues, how long must we wait?, *Tellus*, 46 A, 314-324.

Wilks, D., (2006). *Statistical Methods in the Atmospheric Sciences*, 2nd Edition, Elsevier, Amsterdam.