

## Use of mesoscale model MM5 forecasts as proxies for surface meteorological and agroclimatic variables

Daniel O. Silva<sup>1</sup>, Francisco J. Meza<sup>1,2</sup>, and Eduardo Varas<sup>3</sup>

<sup>1</sup>Facultad de Agronomía and Ingeniería Forestal, <sup>2</sup>Centro Interdisciplinario de Cambio Global, and <sup>3</sup>Facultad de Ingeniería, Pontificia Universidad Católica de Chile, Santiago, Chile.

### Abstract

**D.O. Silva, F.J. Meza, and E. Varas. 2009. Use of mesoscale model MM5 forecasts as proxies for surface meteorological and agroclimatic variables. Cien. Inv. Agr. 36(3):369-380.** There is increasing interest in meteorological information and its application to strategic planning at the farm as well as regional level. Although we have recently seen significant improvements to strengthen and enlarge networks of weather observations, their density is still insufficient to cover large extents at the desired spatial and temporal resolution. Climate scientists have developed and used mesoscale models to understand and predict future atmospheric conditions. These models represent a major contribution to objective weather forecasts throughout numerical simulations. They use global circulation outputs as boundary conditions and can be run in a nested manner so as to increase their spatial resolution. Because of this, we can obtain information about weather variables in grid cells spaced 15 km apart covering important areas and providing information in places where analog or automatic stations are not available. The objective of this work is to evaluate the use of raw data from the MM5 mesoscale model as well as MOS-corrected information (a statistical post-processing of MM5 outputs) as a proxy for surface meteorological data. Temperature, wind speed, relative humidity, and daily solar radiation forecasts were evaluated for eleven stations in the Maipo river basin. In all cases, the MOS forecast produced better results than the raw MM5 data. Determination coefficients reached values near 0.9, and the RMSE was usually smaller for MOS-corrected data. The small variability of the MOS parameters allows their use as regional values to estimate meteorological data for the whole region, particularly at a weekly time step.

**Key words:** Agroclimatological variables, Maipo river basin, MM5 data, MOS.

### Introduction

Currently, except by greenhouse gases, atmospheric conditions cannot be controlled or modified to a great extent, so monitoring climatic variables is very important for managing cultivation. Unfortunately, measurements are limit-

ed to the use of instruments focused on specific sites, making their widespread use difficult and expensive, as more instruments (sometimes in nets) are needed to obtain significant estimates of climatic variables at the regional level.

Recently, efforts have been made to develop meteorological nets in Chile. One example is the "Agroclima" net, a joint initiative of the Fundación para el Desarrollo Frutícola (FDF), the Dirección Meteorológica de Chile (DMC) and

the Instituto Nacional de Investigación Agraria (INIA). Beside this net, there are initiatives from private companies and consultants providing climatic information that is applied to agricultural production. However, their coverage is insufficient, as it is difficult to develop an efficient system for medium and small producers, especially when the extent and variability among localities makes extrapolating the results obtained from one spot to another more difficult.

Thus, the shortage of meteorological stations nets and accessible instruments makes spatially quantifying meteorological variables difficult, so we must search for alternative methods that can be applied at the regional level and with relatively low cost, to have a database at an appropriate and accessible resolution. The possible use of information from models of atmospheric simulations for agronomic applications in Chile is thus interesting and underdeveloped as an alternative to direct measurement methods.

Currently, the use of prediction models of meteorological conditions (Numerical Weather Prediction, NWP) is widely accepted at the scientific level, and even developed at the public and commercial levels (Allen *et al.*, 2007; Bastiaanssen *et al.*, 1998). Due to the need for estimates in zones with different topographies, which are more specific for each region, Mesoscale Models (MM) were developed, with special estimation systems to provide localized forecasts (Oncley and Dudhia, 1995; Dudhia *et al.*, 2005).

In this regard, the MM5 model, developed by the National Center for Atmospheric Research and Pennsylvania State University (NCAR/PSU), is one of the most-used mesoscale models for forecasts at the world level and is used widely at the scientific as well as commercial levels (Falvey, 2007; Wilks, 1995). The model, which was developed based on the investigations of Anthes (Anthes and Warner, 1978), includes the estimation of more than 15 different variables, at several height levels, and the use of multiple nestings, which makes their resolution vary among certain limits, as necessary.

Due to the chaotic nature of the atmospheric system, numerical weather prediction models

tend to present significant deviations with respect to the real conditions of the prediction zone. In addition, there is a second source of error in predictions related to the discrete representation of differential equations, to allow their numerical resolution (Wilks, 1995). The most common solution to this problem is to carry out a statistical output postprocessing of a numerical weather prediction model (NWP), adjusting the data obtained from the model to a series of real data from the zone under study.

The objective of this investigation was to fit and validate the MM5 outputs, and adjust them through MOS (Model Output Statistics), for the Metropolitan Region of Chile, to obtain direct and indirect estimates of meteorological variables of agricultural value. In addition, we attempted to establish a basis for a regionalization of the model variables, using the MM5 as an interpolator, in the area, to increase the density of meteorological information. The objective of this article is to prove that the information contained in the MM5 forecasts can potentially be used to estimate variables such as temperature, wind speed, relative humidity and radiation, once the statistical procedures are done. In addition, we explored the use of this information to calculate agroclimatic indices and/or to produce secondary information such as the accumulation of degree days, chill hours, or frost occurrence.

## Materials and methods

### *Study area*

The study area corresponds to the region between the latitudes of 32°55' S and 43°15' S and the longitudes of 70° W and 72° W, an area covered mainly by the Maipo river basin, located between the regions of Metropolitan and Valparaíso. The Maipo river basin has a total surface area of 15,157 km<sup>2</sup>, of which 33% corresponds to high mountains, 45% to forest, 15% to agricultural areas and the remaining 5% is divided between 524 km<sup>2</sup> of urban areas, 206 km<sup>2</sup> of riverbeds, and 32 km<sup>2</sup> of lakes and reservoirs (DGA, 2007).

The geography of this zone varies from east to west, from the slopes of the Andes mountains to the coast of the Pacific Ocean. Despite the geographic variability, the atmospheric conditions tend to show similar values and correlate spatially and could be considered as belonging to the same geoclimatic zone. However, a climatic change from the Andean foothill zone to the valleys and then the coastal zone is clearly visible.

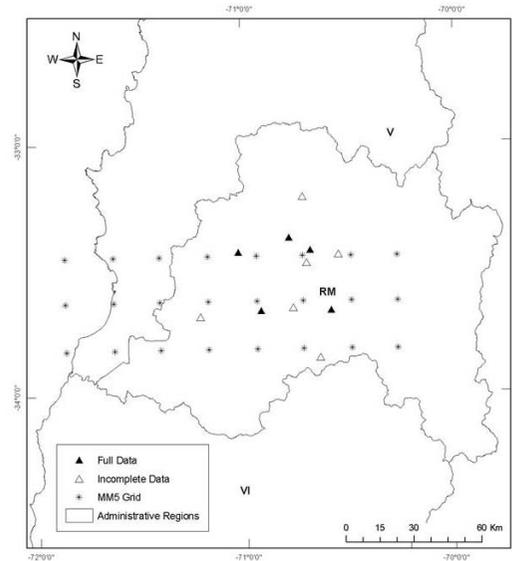
The Maipo river valley contains most of the population of the country, and one of the largest cities, Santiago. In addition, it is one of the important agricultural valleys, due to its climate and closeness to commercial spots, airports and ports. The climate of the valley corresponds mainly to a Mediterranean warm temperate climate, with clear seasons and winter rains. Eighty percent of the precipitation falls between May and August, reaching almost 350 mm annually and transforming into solid precipitation above 1,500 m elevation. The annual average relative humidity is slightly above 70%, while the temperature reaches 7.7°C in the winter and 22.1°C on average in summer, with an annual mean of 13.9°C. The distribution of precipitation in the region defines the summer cultivation; agricultural systems depend on the existing water to supplement the evapotranspiration water demand through irrigation.

### *Meteorological information*

Meteorological information is divided into two types: (i) observed data collected by automatic and manual meteorological stations and (ii) data generated by the MM5, obtained by the Department of Geophysics (DGF), Universidad de Chile, Santiago, Chile.

The real data were obtained from a network of meteorological stations available in the study area belonging to the Agroclima project (FDF-INIA and DMC), in addition to some stations not included in this DMC net, the Dirección General de Aguas (DGA) and the meteorological station of Pirque (Pontificia Universidad Católica de Chile). The locations of these stations is shown in Figure 1. The stations of Pirque

and Talagante correspond to automatic stations providing a comprehensive record, while the rest present fewer variables and are semiautomatic or analog (manual).



**Figure 1.** Geographic distribution of weather stations in the basin. Stations with complete data have records of temperature, relative humidity, wind speed and solar radiation. Incomplete ones have only some of these records.

The continuity and frequency of data collection from automatic stations compared to analog (manual) stations and the inherent error in the sensor readings by the observer make the data from the automatic stations more reliable than those from analog stations.

Both data groups were subject to a process of quality control where incoherent data were eliminated, which generally correspond to errors in sensors or problems in measurement. Thus, all data exceeding three standard deviations from the distribution mean from each variable were eliminated.

Regarding the MM5 data, the DMC evaluates the use of the results (called outputs) from this model to support their daily regional forecasts. Likewise, the DGF has implemented the MM5 since 2002 and has run it periodically since then, balancing their parameters and using them in diverse investigations (Garreaud and Muñoz, 2004; Garreaud and Rutlland, 2003; Falvey, 2006).

The DGF starts the MM5 based on information from the Global Forecast System (GFS), a global simulation for the atmospheric conditions of NCAR. Three nested dominions of 135, 45, and 15 km<sup>2</sup> have been implemented for their execution, with 29 height levels and 17 variables per level. The main variables included in the model are: temperature, specific humidity, short wave sun radiation, long wave radiation, sensible heat flow, latent heat flow, soil heat flow, and direction and intensity of wind.

The model data were obtained from the MM5 model-DGF for each point corresponding to a station, by extrapolating the closest points by the inverse of the tridimensional distance from each point. The whole process of data extraction from the MM5 and the interpolation to the geographic points of the stations was done using MATLAB (Mathworks, 2007).

The MM5-DGF provides predictions from a grid in which the points are spaced 15 km apart

(in the smallest region), as shown in Figure 1. In addition, the model provides information on eight variables at each point (height, pressure, temperature, zone wind, south wind, vertical wind, specific humidity, cloud water and precipitation water) at each of the 29 vertical levels. It also provides nine diagnostic surface variables (precipitation, temperature, short wave and long wave radiation, sensible and latent heat flow) (Falvey, 2006). The MM5 model is executed daily at an initial hour of 00:00 UTC (20:00 local time) and simulations supplying a schedule forecast until 72 h in the future are performed.

Table 1 gives the variables considered for each station, and indicates which were used as estimators obtained from the model. We note that the MM5 model does not simulate relative humidity directly, but specific humidity, so the data from specific humidity and temperature were used to obtain values of relative humidity for the comparisons.

**Table 1.** Meteorological data available at each station and in the MM5 outputs.

| Station          | Temperature    |                | Relative humidity % | Solar radiation MJ m <sup>-2</sup> day <sup>-1</sup> | Wind speed m s <sup>-1</sup> | Specific Humidity g kg <sup>-1</sup> | Pressure Pa    |
|------------------|----------------|----------------|---------------------|------------------------------------------------------|------------------------------|--------------------------------------|----------------|
|                  | Maximum °C     | Minimum °C     |                     |                                                      |                              |                                      |                |
| Cerrillos        | X <sup>1</sup> | X <sup>1</sup> | X <sup>1</sup>      |                                                      | X <sup>1</sup>               |                                      |                |
| Colina           | X              | X              | X                   |                                                      |                              |                                      |                |
| Huelquén         | X              | X              | X                   |                                                      |                              |                                      |                |
| Lo Herrera       | X              | X              | X                   |                                                      |                              |                                      |                |
| Melipilla        | X              | X              | X                   |                                                      |                              |                                      |                |
| Panguiles        | X              | X              |                     | X                                                    |                              |                                      |                |
| Pirque           | X              | X              | X                   | X                                                    | X                            |                                      |                |
| Pudahuel         | X              | X              | X                   |                                                      | X                            |                                      |                |
| Quinta Normal    | X              | X              | X                   |                                                      | X                            |                                      |                |
| Talagante        | X              | X              | X                   | X                                                    | X                            |                                      |                |
| Tobalaba         | X              | X              |                     |                                                      | X                            |                                      |                |
| MM5 <sup>1</sup> | X              | X              |                     | X                                                    | X                            | X <sup>1</sup>                       | X <sup>1</sup> |

<sup>1</sup>X = Data available. MM5= Data obtained from MM5 outputs.

The MM5 model-DGF presents errors in the estimation of the variables; the main sources are from the initial or edge conditions (depending on the GFS model), the parameterization of

subgrid processes (inherent to the model), the low spatial resolution (especially in zones with complex topography), and numerical errors. The sum of these errors tends to produce devia-

tions between the prediction and the predictor, which in many cases may be systematic, that is, repeated through time. When this occurs, an MOS application to correct these forecasts becomes necessary, and they are adjusted to the real values (Sokol and Rezacova, 2000; Hart *et al.*, 2002).

*Model output statistics, MOS*

Glahn and Lowry (1972) developed a forecast tuning method for Numerical Weather Prediction (NWP), through the determination of statistical relations between the prediction and the variables of the model at specific times, determined by an analysis of regression (Stepwise Regression or Screening) (Antolik, 2000; Wilks, 1995; Sokol and Rezacova, 2000; Hart *et al.*, 2004). The procedure has become popular, and the method as a whole is called Model Output Statistics (MOS; Glahn and Lowry, 1978; Wilks, 1995).

The first stage of an MOS involves eliminating incoherent or wrong data. Then, the predictors are selected to explain the variability of the parameter, relating the prediction with real data from each point of study. At the annual level, some meteorological variables present cyclical tendencies (Wilks, 1995), related to the translation and rotation of the Earth, which makes their variability mostly explainable by periodic functions. Most of the observed variability is explained when obtaining the necessary coefficients from the real data, to reproduce this seasonal behavior, becoming one of the predictors that perform better.

The reproduction of this seasonability is achieved through a multiple regression, where the predictors are sine and cosine functions operating with an angle with an angular frequency that varies according to the periodicity and the coefficients are determined by the regression with observed data. These functions, known as Fourier series (Ec 1), consist of orthogonal combinations of sine and cosine functions, with their amplitudes and phases adjusted to a determined cycle, as follows:

Ec 1:

$$F_t = b_0 + b_1 \sin\left(\frac{2\pi t}{\tau}\right) + b_2 \cos\left(\frac{2\pi t}{\tau}\right)$$

where F corresponds to the seasonal component or Fourier,  $b_0$ ,  $b_1$ , and  $b_2$  are the coefficients associated with the Fourier series of a harmonic, t is the Julian day, and  $\tau$  is the total period of the cycle, in this case 365 days. Therefore, when the specific value for each prediction given by the Fourier series is obtained, the average distribution of a specific variable is available. This value is a good estimator for the temporary variability of a meteorological variable influenced by the Earth cycle, and is generally the first predictor used in an MOS (Antolik, 2000; Clark and Hay, 2004).

Once the seasonal components of the variables to be predicted have been determined, the anomalies of these variable may be obtained, subtracting the observed value from the value determined by the Fourier series. In addition to helping to explain a significant percentage of the variability observed, this procedure reduces the probability of incorporating predictors whose correlation is only based on seasonability.

After including the first predictor, seasonality, the following MM5 predictors are selected to explain the existing variability and the difference between the prediction and the real value of the variable. Thus, the stepwisefit process is used, which consists of searching for the predictor that best explains the prediction variability, from a multiple regression, reducing the sum of the squares, or obtaining their higher correlation index. Once the first predictor has been selected, a second predictor is searched for that forms the best pair and minimizes the sum of the squares. This process is repeated until the minimum mean-square error decreases to a satisfactory level. Thus, the prediction from MM5 variables (X) for an observed meteorological variable (Y) is as follows:

Ec 2:

$$(Y_t - F_{yt}) = a_0 + a_1(X_{1t} - F_{x1t}) + a_2(X_{2t} - F_{x2t}) + \dots + a_n(X_{nt} - F_{xnt})$$

where  $Y_p$  is the model prediction corrected by MOS,  $a_i$  are the regression coefficients, and  $X_i$  are the estimators obtained from the model that allow us to correct the prediction (Antolik, 2000).

This process considers the correlations between each of the predictors, avoiding the problem of overfitting the observed values and including unnecessary variables. This new prediction, obtained after the MOS, is compared with the real data, to check how it fits in relation to real factors.

The data set used correspond to information from the years 2004 to 2007. They were divided into two groups, the first from 2004 to 2006 and the second corresponding to 2007, to balance the equations with the initial conjoint and validate their pertinence with a group of independent observations, corresponding to the year 2007. Finally, the different localities with data from meteorological stations were compared, proving the existence of a common pattern allowing the predictions to be regionalized, to estimate unbalanced points, such as where there are no stations with real data, and estimate variables at points not found in the MM5 grid. It is convenient to observe the behaviors of the different patterns characterizing each locality, such as regression coefficients (like seasonability for the MM5 predictors; Ec 2) and the predictors chosen by the stepwise-fit process. The similarity in zones close to the meteorological stations allows the use of the coefficients obtained at points that were not directly balanced.

## Results and discussion

### *Meteorological variables*

As the meteorological variables generally present an outstanding pattern of seasonability, an important proportion of the variability observed was captured by the equations representing the periodic cycles (Fourier series). The model parameters (estimated through multiple regres-

sion), the determination coefficients and the root mean square of the error of each variable are shown in Table 2.

Table 2 indicates that seasonality explains most of the variability of the meteorological observations, representing almost 80% of the data variance. Similar results have been found in other studies whose objective was the stochastic modeling of the meteorological data variability (Meza *et al.*, 2003). There was significant evidence ( $p < 0.001$ ) to reject the hypothesis that the regression coefficients are equal to zero at all the localities and for all the variables. According to the determination coefficients, the better fits are observed for temperature and radiation, possibly due to the seasonal nature of these parameters, with rather lower values for relative humidity and wind. The best fit corresponds to the variable mean temperature, as the daily and monthly variability were lower than at extreme temperatures as they are averages, being very well represented by the sinusoidal cycle of the Fourier series.

On the other hand, the regression coefficients show certain levels of similarity among the different localities, which allows us to considering the option of using coefficient values from nearby localities to estimate the seasonal fluctuations of meteorological variables in localities without surface information.

Similarly, Table 3 shows the MOS contribution from variables of the MM5 model to the variability of the data surface. In this case, the coefficients correspond to the regression between the corresponding MM5 variable with the surface observation (without MOS), and the prediction from postprocessing of the MM5 data versus the surface observation (with MOS). It is clear that the existing direct effect from using MOS improves the predictions in all the cases, not only on the correlation level, but also with respect to bias and variability (RMSE). These results agree with the reports by Hart *et al.* (2004). These authors observed significant improvements in temperature forecasts, relative humidity and wind speed when MOS-type statistical processing was used.

**Table 2.** Regression coefficients, determination coefficients and root mean square error (RMSE) of Fourier series representing the seasonality of the meteorological variables (P-values not included).

| Location                                           | $b_0$ | $b_1$ | $b_2$  | RMSE  | R <sup>2</sup> |
|----------------------------------------------------|-------|-------|--------|-------|----------------|
| Maximum temperature, °C:                           |       |       |        |       |                |
| Colina                                             | 24.21 | 1.24  | 6.71   | 3.90  | 0.61           |
| Huelquén                                           | 24.35 | 1.26  | 8.63   | 4.37  | 0.67           |
| Lo Herrera                                         | 22.02 | 1.18  | 6.32   | 3.98  | 0.57           |
| Panguiles                                          | 22.00 | 2.21  | 6.92   | 2.96  | 0.74           |
| Pirque                                             | 22.29 | 2.28  | 6.46   | 3.98  | 0.58           |
| Pudahuel                                           | 22.41 | 2.20  | 7.45   | 3.72  | 0.66           |
| Quinta Normal                                      | 22.44 | 2.46  | 7.88   | 3.35  | 0.73           |
| Talagante                                          | 22.10 | 1.72  | 7.03   | 3.58  | 0.65           |
| Minimum temperature, °C:                           |       |       |        |       |                |
| Colina                                             | 6.54  | 0.11  | 5.02   | 3.80  | 0.47           |
| Huelquén                                           | 7.65  | 0.35  | 3.91   | 2.98  | 0.46           |
| Lo Herrera                                         | 6.30  | 0.62  | 3.93   | 2.64  | 0.53           |
| Panguiles                                          | 9.25  | 1.55  | 3.60   | 2.29  | 0.59           |
| Pirque                                             | 5.17  | 0.51  | 3.71   | 2.83  | 0.43           |
| Pudahuel                                           | 7.72  | 1.19  | 4.50   | 2.63  | 0.58           |
| Quinta Normal                                      | 8.96  | 1.56  | 4.67   | 2.21  | 0.69           |
| Talagante                                          | 7.96  | 1.08  | 3.53   | 2.48  | 0.51           |
| Mean temperature, °C:                              |       |       |        |       |                |
| Colina                                             | 14.69 | 0.38  | 6.50   | 2.29  | 0.81           |
| Huelquén                                           | 15.06 | 0.76  | 5.99   | 2.38  | 0.76           |
| Lo Herrera                                         | 13.33 | 0.77  | 5.32   | 2.07  | 0.77           |
| Panguiles                                          | 15.63 | 1.88  | 5.26   | 1.84  | 0.82           |
| Pirque                                             | 13.17 | 1.19  | 5.68   | 2.11  | 0.77           |
| Pudahuel                                           | 15.06 | 1.70  | 5.98   | 2.02  | 0.81           |
| Quinta Normal                                      | 15.70 | 2.01  | 6.28   | 2.05  | 0.82           |
| Talagante                                          | 14.06 | 1.19  | 5.29   | 1.79  | 0.81           |
| Relative humidity, %:                              |       |       |        |       |                |
| Colina                                             | 76.02 | 3.27  | -11.89 | 9.70  | 0.45           |
| Huelquén                                           | 76.57 | 0.66  | -9.69  | 11.13 | 0.28           |
| Lo Herrera                                         | 84.07 | 1.06  | -6.24  | 7.62  | 0.26           |
| Panguiles                                          | 83.30 | -0.77 | -6.50  | 7.37  | 0.25           |
| Pirque                                             | 73.50 | -2.57 | -10.88 | 9.00  | 0.41           |
| Pudahuel                                           | 66.38 | -1.70 | -15.17 | 9.28  | 0.53           |
| Quinta Normal                                      | 70.91 | -0.12 | -15.47 | 8.12  | 0.58           |
| Talagante                                          | 83.43 | -0.48 | -6.45  | 7.16  | 0.26           |
| Radiation, MJ·m <sup>-2</sup> ·day <sup>-1</sup> : |       |       |        |       |                |
| Pirque                                             | 11.30 | -1.83 | 6.62   | 3.13  | 0.68           |
| Pudahuel                                           | 11.82 | -0.83 | 7.20   | 3.09  | 0.69           |
| Quinta Normal                                      | 11.88 | -0.58 | 7.57   | 2.98  | 0.71           |
| Talagante                                          | 11.85 | -1.21 | 7.41   | 2.98  | 0.72           |
| Wind speed, m·s <sup>-1</sup> :                    |       |       |        |       |                |
| Pirque                                             | 0.94  | -0.11 | 0.52   | 0.28  | 0.59           |
| Pudahuel                                           | 3.31  | 0.08  | 1.08   | 0.92  | 0.37           |
| Quinta Normal                                      | 2.12  | 0.16  | 0.62   | 0.68  | 0.28           |
| Talagante                                          | 2.76  | -0.01 | 1.81   | 1.02  | 0.57           |

**Table 3.** Goodness of fit between observed values and MM5 data, with and without MOS<sup>1</sup>.

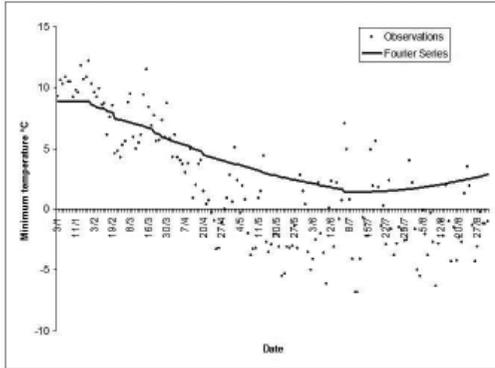
| Location                      | Without MOS    |       |       |       |       | With MOS       |      |       |      |       |
|-------------------------------|----------------|-------|-------|-------|-------|----------------|------|-------|------|-------|
|                               | R <sup>2</sup> | RMSE  | n     | m     | MAD   | R <sup>2</sup> | RMSE | n     | m    | MAD   |
| <i>Maximum temperature</i>    |                |       |       |       |       |                |      |       |      |       |
| Pirque                        | 0.74           | 2.90  | -0.85 | 1.02  | 13.70 | 0.79           | 2.64 | 0.00  | 1.00 | 12.52 |
| Talagante                     | 0.63           | 3.69  | 2.09  | 0.85  | 13.20 | 0.69           | 3.38 | 0.00  | 1.00 | 8.80  |
| Quinta Normal                 | 0.80           | 2.86  | -1.37 | 0.99  | 12.60 | 0.84           | 2.59 | 0.00  | 1.00 | 8.87  |
| <i>Minimum temperature</i>    |                |       |       |       |       |                |      |       |      |       |
| Pirque                        | 0.48           | 2.49  | 0.59  | 0.63  | 10.30 | 0.74           | 1.76 | 0.00  | 1.00 | 5.89  |
| Talagante                     | 0.46           | 2.59  | 3.18  | 0.55  | 11.50 | 0.57           | 2.31 | 0.00  | 1.00 | 6.75  |
| Quinta Normal                 | 0.63           | 2.41  | 3.85  | 0.70  | 10.40 | 0.84           | 1.60 | 0.00  | 1.00 | 6.33  |
| <i>Mean temperature</i>       |                |       |       |       |       |                |      |       |      |       |
| Pirque                        | 0.85           | 1.57  | 1.45  | 0.84  | 10.11 | 0.88           | 1.41 | 0.00  | 1.00 | 7.35  |
| Talagante                     | 0.75           | 2.05  | 3.58  | 0.71  | 10.74 | 0.84           | 1.65 | 0.00  | 1.00 | 5.20  |
| Quinta Normal                 | 0.83           | 1.98  | 3.13  | 0.85  | 6.45  | 0.90           | 1.51 | 0.00  | 1.00 | 4.90  |
| <i>Mean relative humidity</i> |                |       |       |       |       |                |      |       |      |       |
| Pirque                        | 0.31           | 9.62  | 41.94 | 0.57  | 49.20 | 0.52           | 8.02 | 0.00  | 1.00 | 38.25 |
| Talagante                     | 0.08           | 7.97  | 68.08 | 0.24  | 58.31 | 0.28           | 7.03 | 0.00  | 1.00 | 21.65 |
| Quinta Normal                 | 0.12           | 11.84 | 47.04 | 0.39  | 52.50 | 0.67           | 7.24 | 0.00  | 1.00 | 22.17 |
| <i>Solar radiation</i>        |                |       |       |       |       |                |      |       |      |       |
| Pirque                        | 0.69           | 2.94  | -0.36 | 0.71  | 15.70 | 0.75           | 2.65 | 0.00  | 1.00 | 8.78  |
| Talagante                     | 0.68           | 3.16  | -0.30 | 0.73  | 16.85 | 0.76           | 2.77 | -0.04 | 1.01 | 8.97  |
| Quinta Normal                 | 0.68           | 3.16  | -0.31 | 0.74  | 16.08 | 0.75           | 2.78 | 0.00  | 1.00 | 9.51  |
| <i>Wind speed</i>             |                |       |       |       |       |                |      |       |      |       |
| Pirque                        | 0.19           | 0.38  | 1.74  | -0.24 | 5.25  | 0.64           | 0.25 | 0.00  | 1.00 | 0.98  |
| Talagante                     | 0.01           | 1.54  | 3.14  | -0.18 | 6.62  | 0.58           | 1.00 | 0.00  | 1.00 | 3.11  |
| Quinta Normal                 | 0.05           | 0.77  | 2.73  | -0.19 | 5.52  | 0.28           | 0.67 | 0.00  | 1.00 | 1.88  |

<sup>1</sup>R<sup>2</sup> is the determination coefficient, RMSE corresponds to the root mean square error, MAD corresponds to the maximum absolute deviation, and m and n are the slope and intercept of the observed vs. predicted regression, respectively.

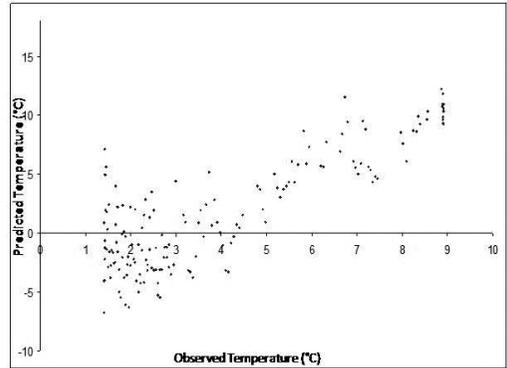
In addition, when comparing the information contained in Tables 2 and 3, the contributions of the model representing seasonability (Table 2), the MM5 variable without transformation (Table 3 without MOS) and the further statistical processing (Table 3 with MOS) are quantitatively noticeable. This is clearly exemplified in Figures 2, 3, 4 and 5. In the first case (Figure 2), the seasonal variability of minimum temperature in Pirque and the tendency of the mean were calculated based on the regression coefficients of the Fourier series. We note that the coefficients were obtained based on the balance data conjoint (years 2004 to 2006), while Figure 2 represents the minimum temperatures for the year 2007. Figure 3 shows the contrast between the minimum observed temperature and the values predicted by the Fourier series. A clear dispersion is observed, especially around the minimum and maximum values,

which may be predicted by the seasonal component because this statistical model is limited between determined values (1.3 and 8.9°C); in these conditions, the Fourier series model can reproduce up to 43% of the observed variability.

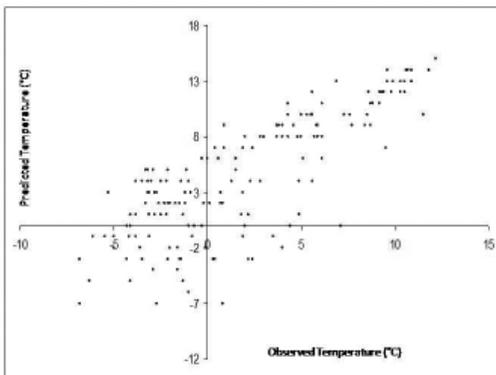
The contribution of the prediction from the MM5 model to the variable minimum temperature in Pirque is seen in Figure 4. The data dispersion is reduced and flexibility is achieved as it can better represent the temperature anomalies (deviations from the central tendency). In this case, the coefficient of determination is substantially higher, and reproduces almost 48% of the total variance. The highest contribution is achieved using MM5 estimations corrected by statistical processing (MOS). Figure 5 shows that as the observation dispersion decreases substantially, the bias disappears almost completely, reaching



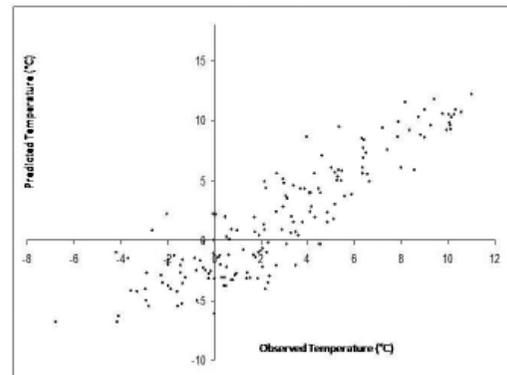
**Figure 2.** Seasonal variation of minimum temperature registered in Pirque, Chile in 2007.



**Figure 3.** Scatter plot between the observed minimum temperature and forecasted values using Fourier series obtained in Pirque, Chile, in 2007.



**Figure 4.** Scatter plot between the observed minimum temperature and predicted values using MM5 for data obtained in Pirque, Chile, in 2007.



**Figure 5.** Scatter plot between the observed minimum temperature and MOS corrected MM5 predictions in Pirque, Chile in 2007.

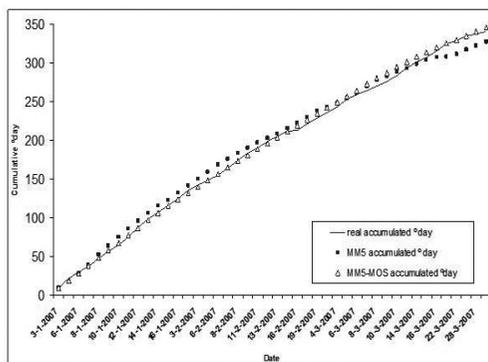
a level of explanation of 74% of the variability observed. The capacity of the statistical processing (MOS) for reducing the bias errors represents one of the most attractive attributes of this type of procedure; it is generally superior to other similar statistical treatments (Cheng and Steenburgh, 2007).

*Variables of agricultural interest*

One interesting aspect observed in this investigation is that the meteorological observations present higher variability than that explained by the model considering the Fourier series and the MOS processing to the data of the MM5 model. However, this discrepancy is compensated for,

as the errors in the estimation by excess or defect are similar through time and tend to cancel. This allows that, even when the fit between the model predictions and the observations is not perfect, estimations with higher levels of confidence are achieved, as the temporary resolution decreases (by increasing the time intervals to scales of weeks or months). This becomes useful for some variables of agricultural common use that are estimated at the week, month, or even season level (for example, from October to April). Examples of these variables are the accumulation of degree days (thermal sums over a threshold value), chill hours defined as the number of hours under a certain temperature threshold with the property for inducing changes in the development states, and accrued evapotranspiration.

As seen in Figure 6, the behavior of the MM5 estimations with MOS to predict the accumulation of degree days in the Pudahuel station is acceptable and tends to reproduce the evolution of this variable through time. Although the MM5 values without correction also present this property when added, the best results are obtained based on the corrected estimations.



**Figure 6.** Accumulated degree days from January to March 2007 for Pudahuel station, Chile.

The surface meteorological information is relevant for developing season-specific strategies for productive management. One of the main limitations for productively using meteorological information is the relative scarcity of meteorological monitoring stations. We evaluated the use of results from the numerical MM5 forecast model as a substitute for surface information.

From the results of this investigation, we conclude that: (i). The estimates of the MM5 model have acceptable fits in the case of temperature and radiation. However, they tend to provide uncertain forecasts in the case of humidity and wind speed. (ii). The use of MOS applied to the data from the MM5 model provides good results in virtually all cases. In the case of the maximum and the mean, a significant increase is observed in the coefficient of determination between the observed and simulated data. The best results, in relative terms, are for the prediction of wind and relative humidity; however, they do not reach high levels of correlation. An important portion of variability is explained, allowing estimations with limited uses, but higher

than the estimations from the model without processing. (iii). At the week and month levels, the estimations of variables such as degree days or hours chill have a good fit, as the temporary resolution of the model compensates for the overestimations with the subestimates at the day level. (iv). The use of MOS improves not only the coefficients of correlation between the prediction and the predictors, but also reduces the variability and the bias.

The results obtained in this investigation are therefore promising and can be used in different areas of applied meteorology and bias. For instance, it may be interesting to evaluate their MM5 estimations corrected by MOS as a substitute for surface observations in estimation models of real evapotranspiration based on satellite images (for example, the SEBAL model described by Bastiaansen *et al.*, 1998). They can also be used directly as predictors of meteorological variables, or to predict derivative variables, such as the accumulation of degree days, hours of chill, or reference evapotranspiration. In the case of reference evapotranspiration, this could allow studies to be carried out on water balance at the basin level.

### Acknowledgements

We acknowledge support from the CONICYT, Fondecyt 1060544 project, and from the Inter-American Institute for Global Change Research, project SGP-HD 003, which is part of Grant GEO-0642841 of the US National Science Foundation. The main author thanks the support from the Dirección de Investigación and Postgrado of Pontificia Universidad Católica. The primary database was configured thanks to the collaboration of the Agroclima net (FDF, INIA, DMC) through Alan Hernández and Enrique Garrido, and Carlos Gálvez from the experimental station in Pirque, Pontificia Universidad Católica de Chile. Finally, we thank René Garreaud and Mark Falvey, Departamento de Geofísica, Universidad de Chile, for providing access to the MM5 database.

## Resumen

**D.O. Silva, F.J. Meza y E. Varas. 2009. Uso del modelo MM5 como predictor de variables meteorológicas y de interés agroclimático. Cien. Inv. Agr. 36(3):369-380.** En la actualidad existe un interés creciente por contar con información meteorológica que permita la gestión de la producción agrícola y el manejo de recursos naturales tanto a nivel de predio como a escala regional. A pesar que se han hecho esfuerzos considerables por fortalecer y expandir las redes de monitoreo meteorológico, todavía su densidad es insuficiente para proveer de información regional con el nivel de frecuencia y representatividad espacial deseado. Climatólogos y meteorólogos han desarrollado y empleado los modelos de simulación de meso-escala con el fin de mejorar la comprensión de los fenómenos meteorológicos y predecir el comportamiento de la atmósfera. Dichos modelos son alimentados por los resultados de modelos de mayor envergadura (Modelos de Circulación Global) quienes entregan las condiciones de inicio para que los modelos de mesoescala aumenten la resolución espacial mediante un sistema de ejecución anidado. Como resultado se logra obtener información climática en celdas que están espaciadas por 15 km. Esto es promisorio ya que tal densidad espacial tiene la capacidad de generar información en lugares donde no existen sistemas de monitoreo activos. El objetivo de este trabajo fue evaluar el uso de los resultados del modelo de meso escala MM5 a través de cuantificar la capacidad de reproducir condiciones meteorológicas observadas, tanto los resultados directos del modelo MM5 como también valores corregidos en base a un procesamiento estadístico conocido como MOS. Se seleccionaron 11 estaciones meteorológicas pertenecientes a la cuenca del Maipo y se evaluaron variables como temperatura, velocidad del viento, humedad relativa y radiación solar. Los resultados muestran que los procesamientos estadísticos posteriores mejoran notablemente las estimaciones iniciales llegando a mostrar coeficientes de determinación de hasta 0,9 ( $p = 0,01$ ). Además se observa que los parámetros de las ecuaciones de regresión son similares entre estaciones, lo que abre la posibilidad de usar valores únicos de ellos para generar estimaciones regionales de variables meteorológicas y agroclimáticas particularmente a escalas de tiempo semanal.

**Palabras clave:** Cuenca del río Maipo, datos MM5, MOS, variables agroclimáticas.

## References

- Allen, R.G., M. Tasumi, A. Morse., R. Trezza, J.L. Wright, W. Bastiaansen, W. Kramber, I. Lorite, and C.W. Robison. 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-applications. *Journal of Irrigation and Drainage Engineering* 133:395-406.
- Anthes, R., and T. Warner. 1978. Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. *Monthly Weather Review* 106:1045-1078.
- Antolik, M.S. 2000. An overview of the national weather service's centralized statistical quantitative precipitation forecasts. *Journal of Hydrology* 239:306-337.
- Bastiaansen, W., M. Menenti, R. Feddes, and A. Holtslag. 1998. A remote sensing surface energy algorithm for land (SEBAL) formulation. *Journal of Hydrology* 212-213:198-212.
- Cheng, W., and J. Steenburgh. 2007. Strengths and weaknesses of MOS, running-mean bias removal, and Kalman filter techniques for improving model forecasts over the Western United States. *Weather and Forecasting* 22:1304-1318.
- Clark, M.P., and L.E. Hay. 2004. Use of medium-range numerical weather prediction model output to produce forecasts of streamflow. *Journal of Hydrometeorology* 5:15-32.
- DGA. 2007. Bases para la Formulación de un Plan

- Director para la Gestión de los Recursos Hídricos Cuenca del Río Maipo, Etapa I Diagnóstico/ Ministerio de Obras Públicas, Dirección General de Aguas (DGA), Departamento de Estudios y Planificación, CONIC BF Ingenieros Civiles Consultores. Santiago, Chile. 620 pp.
- Dudhia, J., D. Gill, K. Manning, W. Wang, and C. Bruyere. 2005. PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modelling System Version 3. <http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html> (Accessed: June, 2008).
- Falvey, M. 2006. Desarrollo de Sistema de Pronóstico de Temperatura Utilizando el Modelo MM5. Informe Final preparado para el Departamento de Geofísica Facultad de Ciencias Físicas y Matemáticas Universidad de Chile. Santiago, Chile. 127 pp.
- Garreaud, R., and J. Rutllant. 2003. Coastal Lows along the subtropical West COSAT of South America: Numerical Simulation of a Typical Case. *Monthly Weather Review* 131:891-908.
- Garreaud R., and R. Muñoz. 2004. The diurnal cycle in circulation and cloudiness over the subtropical southeast Pacific: a modeling study. *Journal of Climatology* 17:1699-1710.
- Glahn, H.R., and D.A. Lowry. 1972. The Use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology* 11:203-211.
- Hart, K.A., W.J. Steenburgh, D.J. Onton, and A.J. Siffert. 2004. An evaluation of mesoscale-model-based model output statistics (MOS) during the 2002 olympic and paralympic winter games. *Weather and Forecasting* 19:200-218.
- Meza, F.J., D.S. Wilks, S.J. Riha, and J.R. Stedinger. 2003. Value of perfect forecasts of sea surface temperature anomalies, for selected rain-fed agricultural locations of Chile. *Agricultural and Forest Meteorology* 116 (3-4): 117-135.
- Oncley, S., and J. Dudhia. 1995. Evaluation of surface fluxes from MM5 using observations. *Monthly Weather Review* 123:3344-3357.
- Sokol, Z., and D. Rezacova. 2000. Improvement of local categorical precipitation forecasts from NWP model by various statistical post processing methods. *Studia Geophysica et Geodesia* 44:38-56.
- Wilks, D.S. 1995: *Statistical Methods in the Atmospheric Sciences*. First Ed. Academic Press. NY, USA. 467 pp.