Modeling of meteorological variables in localities without meteorological data, from the data recorded in neighboring locations that have topomesoclimatically similar conditions for the installation of renewable energy projects

E Mera 1, P Pacheco 1, C Parodi 2 and L Gutiérrez 3

1 Universidad Tecnológica Metropolitana, Physics Department, Box 9845, Santiago, Chile
2 Universidad Tecnológica Metropolitana, Engineering Department, Box 9845, Santiago, Chile
3 Corporación Chile Ambiente, Box 1372, Santiago, Chile

E-mail: emera@utem.cl

Abstract. The present study models meteorological variables for a sector with nonexistent data through the technique of data imputation [2,3] for the installation of renewable energy projects, which in a first step seek for data generated under similar conditions, and in a second step applies a ratio [2] to these data based on an expert consultation. To validate the methodology, a meteorological station representative of the sector under study was installed and the modeled and real values were contrasted, registering that the lowest correlation coefficient is in the variable of Wind Speed (0.246) and high in the variable Temperature (0.657).

1. Introduction

It is still common to find territories in which there are no meteorological stations. For the generation of meteorological data in sectors where there are no meteorological stations for the installation of renewable energy projects, using data imputation techniques, which can be organized into two groups. The first group are the methods based on the Nearest Neighbor. The second group are the methods based on Statistical Data supported by experts, among which are the Ratio Method [2], Random Hot Deck Method, and the Regional Vector Method [1]. These last methods give origin to the present investigation which proposes to model meteorological variables for a sector with nonexistent data through the method of data imputation, which consists of the following processes:

- Environmental characterization of the place under study.
- Search of neighboring sectors topomesoclimatically similar to the area studied [2,3], which presents meteorological stations with complete data series for various meteorological variables.
- Consult experts for the application of a Reason or condition of equivalence, based on expert criteria to experienced meteorologists (> 20 years) in the sector studied, in order to proceed to apply on each of the data series of the variables understudy, the Reason Method [2], which modulates a signal, based on a normalized scalar valid for a certain time interval.
- Analysis of the quality of the data imputation, contrasting in a data dispersion diagram, the values obtained in the previous point, with respect to the data obtained from a reference meteorological station installed in thesector understudy.

1 emera@utem.cl
The proposed methodology was applied in the sector of Tres Quebradas (3Q), located on the border of Chile with Argentina, in which meteorological data were recreated at the hourly level for the meteorological variables of Temperature (ºC), Relative Humidity (%), Wind Speed (m/s), Pressure (hPa), Solar Radiation (W/m²).

2. Results

Next, the aforementioned processes are developed, in the present analysis case:

2.1. Characterization of the place under study:
The sector under study is a Salar, which is at an altitude of 4100 m.a.s.l, located in the meso climatic zone of northern Chile, in the Andes. The sector under study is a lithium production basin, for which it is already a singularity

2.2. Search for Neighbor Sectors Topomesoclimatically Similar:
The topomesoclimatically similar sector found, with meteorological data available, is the meteorological station of the Salar del Olaroz (O) (Susques, Argentina), which is a micro-basin producing Lithium. It is located at an altitude of 3900 m.a.s.l, in the northern meso climatic zone of Chile, in the Andes, the data used were captured in 2017.

2.3. Consult Experts for the Application of a Reason or Condition of Equivalence:
The following are the data imputation algorithms, assumed at the hourly level, for each meteorological variable under study:

2.4. Temperature (ºC)
The temperature modeling is based on the expert meteorological criterion which applies a normalized ratio, the temperature in Tres Quebradas is 5% lower than in the Salar de Olaroz, since it is located further south (both are contained in the same mesoclimatic zone), the imputed data are obtained with the following expression:

\[
T(3Q) \, ^{\circ}C = (T(O) \times 0.95) \, ^{\circ}C
\]

In Equation 1, 3Q is imputed data in the sector of Tres Quebradas, and O is data observed in the sector of Salar de Olaroz. The series of temperature data (ºC) imputed are:

![Imputed Temperature](Image)

**Figure 1.** Imputed Temperature

2.4.1. Relative Humidity (%)
Relative humidity modeling is based on the expert criteria of a meteorologist, which establishes that the Relative Humidity of the sector under study is never less than 3%, and that this is 3% higher than that registered in the Salar de Olaroz. The imputation equations of meteorological data are:
If $\text{RH}(3Q)\% > 3\%$ then $\text{RH}(3Q)\% = (\text{RH}(O) \times 1.03)\%$ otherwise $\text{RH}(3Q)\% = 3\%$ \hspace{1cm} (2)

Considering equation 2, the series of imputed data of Relative Humidity ($\%$), is (Figure 2).

![Figure 2. Imputed Relative Humidity.](image)

2.4.2. Wind speed ($\text{m/s}$)

The expert criterion of the meteorologist concluded that the wind speed is 30% higher in the sector of Tres Quebradas, than in the sector of the Salar de Olaroz (principle of the reason):

$$V(3Q)\left(\frac{\text{m}}{\text{s}}\right) = (V(0) \times 1.3)\left(\frac{\text{m}}{\text{s}}\right)$$ \hspace{1cm} (3)

From Equation 3, the present series of imputed data for Wind Speed ($\text{m/s}$);

![Figure 3. Imputed Wind Speed](image)

2.4.3. Pressure ($\text{hPa}$)

The data imputation algorithm is based on the principle of reason reviewed by an expert meteorologist, it is defined that the atmospheric pressure in the Tres Quebradas sector experiences 7% less atmospheric pressure than the Salar de Olaroz sector:

$$P(3Q)\left(\text{hPa}\right) = (P(O) \times 0.93)(\text{hPa}).$$ \hspace{1cm} (4)

According to Equation 4, this is the series of imputed data for Pressure ($\text{hPa}$);
2.4.4. Solar Radiation (W/m²)

The imputation of data from this sector is based on the fact that the radiation observed at the surface level in the Tres Quebradas sector is 30% lower on average than what is observed in the Olaroz sector, since it receives less solar radiation. The imputation algorithm based on the principle of ratio is:

\[ R(3Q) \ (W/m²) = (R(O) \times 0.7)(W/m²) \] (5)

When applying Equation 5, on the existing data in Olaroz, we obtain the present data series for Solar Radiation (W / m²) for the sector of Tres Quebradas, see Figure 5.

Table 1. Correlation value

| Variable | Temperature (ºC) | Relative Humidity (%) | Wind Speed (m/s) | Pressure (hPa) | Solar Radiation (W/m²) |
|----------|------------------|-----------------------|------------------|----------------|----------------------|
| R2       | 0.657            | 0.529                 | 0.246            | 0.270          | 0.637                |

In Table 1, it is observed that the imputed variable that has the lowest level of correlation is the variable of Wind Speed (0.246) and the highest is the Temperature (0.657).
3. Discussion
The meteorological data generated improves as the associated areas are more similarly
topomesoclimatically and a higher level of experience has the meteorologist associated with the study.

Acknowledgments
The authors thank the support of the Fondef ID18I10189 project.

References
[1] Espinoza J 2005 Utilización del vector regional con HYDRACCES SENAMHI, Perú
[2] Rao J 1996 On variance estimation with imputed survey data (with discussion), *Journal of the American Statistical Association* 91 499–520
[3] Shao J, & Sitte R 2012 Bootstrap for Imputed Survey Data, *Journal of the American Statistical Association* 91:435 1278-1288