EXPLORING MERGING TECHNIQUES FOR RAIN GAUGE MEASUREMENTS AND SATELLITE RAINFALL ESTIMATES FOR URUGUAY

Proyecto DACC

Desarrollo y Adaptación al Cambio Climático

Contrato de Préstamo MGAP-Banco Mundial 8099-UY

November 2013

Alfaro P. (palfaro@motionsoft.com.uy),

De Vera A. (adevera@fing.edu.uy),

Dinku T. (tufa@iri.columbia.edu).

GENERAL INDEX

ACRONYMS

CMORPH	CPC MORPHing technique
CPC	Climate Prediction Center, NOAA
DACC	Desarrollo y Adaptación al Cambio Climático
DNM	Dirección Nacional de Meteorología
FING	Facultad de Ingeniería, Universidad de la República, Uruguay
IMFIA	Instituto de Mecánica de los Fluidos e Ingeniería Ambiental, FING
INUMET	Instituto Uruguayo de Meteorología (formerly DNM)
IRI	International Research Institute, Earth Institute, Columbia University
NOAA	National Oceanic and Atmospheric Administration
SNIA	Sistema Nacional de Información Agropecuaria
TMPA	TRMM Multisatellite Precipitation Analysis
TRMM	Tropical Rainfall Measuring Mission

RESUMEN

El presente trabajo fue llevado a cabo en el marco del proyecto DACC "Desarrollo y Adaptación al Cambio Climático". Su propósito es mejorar el monitoreo del tiempo de Uruguay, en particular de la precipitación, mediante la combinación de observaciones pluviométricas y estimaciones satelitales de precipitación.

Dos productos satelitales, en una escala temporal diaria y con una resolución especial de 0.25º lat/long, fueron evaluados para determiner cual es más representativo de la precipitación en Uruguay. Los datos de referencia provienen de una red de estaciones, relativamente densa y uniformemente distribuída, de 144 pluviómetros provistas por el INUMET (Instituto Uruguayo de Meteorología, anteriormente DNM) cuyos registros comienzan en enero de 1998 y llegan hasta el presente. Los productos de precipitación satelital evaluados son el Tropical Rainfall Measuring Mission (TRMM) en su version cercana-al-tiempo-real (3B42RT), y el NOAA/Climate Prediction Center MORPHing technique (CMORPH).

Se computaron estadísticos de validación para todo el país para el período entre 2007 y 2009. Los estadísticos utilizados para la comparación fueron los siguientes: probabilidad de detección (POD), tasa de falsas alarmas (FAR), sesgo (FBS), puntuación de habilidad de Heidke (HSS), coeficiente de correlación lineal (Corr), sesgo multiplicativo (BIAS), error medio (ME), error absoluto medio (MAE), y la función de distribución acumulada empírica (CDF). Todas las comparaciones se hicieron contra un conjunto de datos grillados -en la misma grilla que las estimaciones satelitales- de la información de las estaciones, el cual fue obtenido eligiendo el mejor de varios métodos de interpolación: Kriging, Suma Ponderada por el Inverso de la Distancia (IDW por sus siglas en inglés) e IDW en bloque (interpolar a 0.05º y luego promediar a 0.25º).

La validación y comparación de los distintos productos satelitales mostró un mejor desempeño del CMORPH, con valores inusualmente altos (favorables) de algunos de los estadísticos, en particular Corr y POD. Por lo tanto, en vista del mejor desempeño en los estadísticos considerados más relevantes para los propósitos del estudio, CMORPH fue elegido para explorar las distintas técnicas de combinación.

Dado que hay dos versiones disponibles de CMORPH (la Versión 0.x disponible desde el 2002 hasta el presente y la Versión 1.0, un reprocesamiento de los datos de CMORPH utilizando un algorítmo fijo, disponible desde 1998 hasta el presente), se comparó el rendimiento de éstas dos versiones para diferentes períodos de tiempo. Los resultados confirman las diferencias menores entre ambas versiones y muestran un rendimiento creciente del CMORPH v1.0, con la peor performance durante los primeros años (1998-2000) y los mejores valores obtenidos para el ultimo período (2010-2012), para la mayoría de los estadísticos. De esta forma, en vista del rendimiento similar y la mayor longitud temporal de la serie de datos disponible, se trabajó con CMORPH v1.0 para explorar las distintas técnicas de combinación.

En este trabajo, se exploraron diversos métodos para la combinación de las observaciones pluviométricas y las estimaciones satelitales de precipitación, comenzando con algunos de los más simples y progresivamente incluyendo y combinando técnicas más elaboradas.

La metodología seguida para la combinación comprende cuatro pasos básicos:
i) Eliminación del sesgo de la estimación satelital por medio de una remoción de sesgo simple o una igualación de CDF.
ii) Regresión de la información de las estaciones en la estimación satelital, bruta o insesgada, usando un modelo lineal generalizado.
iii) Interpolación de los residuos de la regresión en las ubicaciones de las estaciones a la grilla completa usando un esquema de grillado universal (Kriging o IDW + tendencia) e interpolación en bloque.
iv) Aplicación de una máscara de lluvia/no lluvia (en base a información únicamente de estaciones, únicamente de satélite o una combinación de ambas) para evitar la sobreestimación de la ocurrencia de precipitación.

Siguiendo este esquema, varios productos combinados fueron implementados utilizando las siguientes configuraciones:

Tabla 1: Productos combinados implementados y probados

Producto Combinado	Técnica de Eliminación de Sesgo	Método de Interpolación	Máscara de Lluvia/No Lluvia
IDW_Raw	No	IDW en Bloque	Solo Satélite
IDW_CDF_1	Igualación de CDF	IDW en Bloque	Solo Satélite
IDW_CDF_2	Igualación de CDF	IDW en Bloque	Solo Estaciones
Kriging_CDF	Igualación de CDF	Kriging en Bloque	Solo Satélite

Para evaluar las distintas aproximaciones, la información de las estaciones se dividió aleatoriamente en un conjunto de entrenamiento, con un tercio de las observaciones, y un conjunto de validación, con los restantes dos tercios. Luego, los estadísticos de validación se calcularon para todo el país usando el conjunto de validación como información de referencia.

En la comparación se incluyó un conjunto grillado por medio de IDW en bloque de la información de las estaciones de entrenamiento (representando la habilidad actual de la red pluviométrica) y la estimación bruta del satélite (para evaluar la mejora en la habilidad de la información satelital cuando se combina con observaciones en tierra).

Los resultados muestran un aumento general de la habilidad, en comparación con la estimación satelital bruta, al usar las técnicas de combinación propuestas, indicando una mejora en la precisión de la estimación debido a la incorporación de las observaciones en tierra. Comparando la información grillada de las estaciones con el producto combinado, la información de las estaciones muestra un desempeño un poco mejor en términos de cantidades de precipitación, mientras que el producto combinado es mejor en los estadísticos de detección de precipitación. Aún así, la interpolación del conjunto de entrenamiento por medio de IDW en bloque (usando sólo las observaciones) obtuvo los mejores resultados para la mayoría de los estadísticos a través de todos los períodos evaluados.

De esta manera se puede concluir para el caso particular de Uruguay, donde se dispone de una densidad relativamente alta de observaciones en superficie, que la información de las estaciones es de mayor calidad que las estimaciones satelitales (y tan buena como los productos combinados). Países con redes pluviométricas de menor densidad y con topografías más complejas podrían beneficiarse más de la información satelital y usarla como un conjunto de datos complementario de libre acceso cuando no se disponga de observaciones, pero para obtener la mayor precisión la información observada en las estaciones sigue siendo preferida. Estos resultados son consistentes con los obtenidos en trabajos previos realizados para Uruguay por De Vera y Terra, 2012 (Combining CMORPH and Rain Gauges Observations over the Rio Negro Basin. *J. Hydrometeor.*, 13, 1799-1809).

Una conclusión adicional es que la alta congruencia entre la información del conjunto de validación y la información del conjunto de entrenamiento, coeficientes de correlación cercanos a 0.9 y valores de POD cercanos a 0.95, muestran un alto grado de homogeneidad en la información disponible en las estaciones y una alta representatividad de la información contenida en las series de datos del INUMET.

Durante el transcurso de la investigación, surgieron algunos problemas e ideas que precisan de una mayor exploración. Posibles líneas de trabajo futuro incluyen: i) realizar análisis diferenciados para el semestre Abril-Setiembre (estación fría) y Octubre-Marzo (estación cálida), ii) probar otras configuraciones de combinación y en particular otros métodos de interpolación como interpolación óptima o modelado de la estructura espacial mediante cópulas, iii) interpolación puntual (no en bloque) de alta resolución para mitigar los problemas encontrados frente a eventos extremos, iv) usar la distribución gamma para el modelo lineal generalizado de la etapa de regresión, v) mejorar el remuestreo de la grilla satélital, usando interpoladores más avanzados como el filtro de Lanczos y vi) explorar la combinación con distintos números de estaciones.

ABSTRACT

The present work was conducted in the framework of the DACC project. Its purpose is to improve the climate monitoring of Uruguay, in particular of rainfall, by using combinations of rain gauge observations and satellite rainfall estimates.

Two different satellite products, at daily time scale and a spatial resolution of 0.25º lat/long, were evaluated in order to determine which was the most representative of Uruguay's rainfall. The reference data come from a relatively dense and uniformly distributed station network of 144 rain gauges provided by INUMET, with records starting from 1998 up to the present. The evaluated satellite rainfall products are the Tropical Rainfall Measuring Mission (TRMM) in its near-real-time version (3B42RT), and the NOAA/Climate Prediction Center morphing technique (CMORPH).

Validation statistics were computed for the whole country for the period between 2007 and 2009. The statistics used for the comparison are the following: probability of detection (POD), false alarm ratio (FAR), frequency bias (FBS), Heidke skill score (HSS), linear correlation coefficient (Corr), multiplicative bias (BIAS), mean error (ME), mean absolute error (MAE), and the empirical cumulative distribution function (CDF). All comparisons were made against a gridded set of the station data -on the same grid as the satellite data- which was obtained by choosing the best from several gridding methods: Kriging, Inverse Distance Weighting (IDW), and Block IDW (interpolate at 0.05-deg using IDW and then average at 0.25-deg). The validation and intercomparison of the different satellite products shows that CMORPH has a better performance with unusually good values of some of the statistics, particularly Corr and POD. Therefore, in view of the higher skill in the parameters considered most relevant to the purpose of the study, CMORPH was selected to explore the different merging techniques.

Since there are two versions available of CMORPH (Version 0.x available since 2002 up to present, and Version 1.0 a reprocessing of the CMORPH data with a fixed algorithm available since 1998 up to present) we compared the performance of these two versions for different time periods. The results confirm the minor differences between them and show an increasing performance of CMORPH v1.0 over time, showing the worst performance during the first years (period 1998-2000) and the best values for the last period 2010-2012, for most of the statistics. Hence, in view of the similar performance and the larger data period available, we worked with CMORPH v1.0 to explore the different merging techniques.

Some approaches to merging station data with satellite estimates were explored in this work, starting with some of the simplest ones and progressively including and combining more elaborate techniques.

The methodology used for the merging can be split in four basic steps: i) bias removal from satellite grid using a simple bias removal or the CDF matching, ii) regression of the station data on the raw or unbiased satellite data using a generalized linear model, iii) interpolation of the regression residuals at station locations to the entire grid using an universal gridding (Kriging or IDW + trend) schema and block interpolation, and iv) application of a rain/no rain mask (station data only, satellite data only or a combination of both) to prevent the overestimation of the occurrence of rainfall.

Then, several merged products were generated by combining these steps (Table 2):

Table 2: Merged products implemented and tested

Merged Product	Bias Removal Technique	Interpolation Method	Rain/No Rain Mask
IDW_Raw	No bias removal	Block IDW	Satellite only
IDW_CDF_1	CDF Matching	Block IDW	Satellite only
IDW_CDF_2	CDF Matching	Block IDW	Station only
Kriging_CDF	CDF Matching	Block Kriging	Satellite only

To test the different approaches, station data was divided randomly into a training set, having one third of the observations, and a validation set, having the remaining two thirds. Next, validation statistics were computed for the whole country using the validation set as the reference data.

In the comparisons, gridded station data from the training set was also included. The station data was gridded using Block IDW interpolation. The satellite estimates were also included in the comparisons.The results show an overall increase of skill achieved by using the proposed merging techniques when compared to the satellite-only only product, indicating an improvement in the accuracy of the estimation owing to incorporation of station data. Comparing gridded station data to the merged product, gridded station exhibits a slightly better performance in terms of rainfall amounts while the merged products is better in rainfall detection statistics. Therefore, we can conclude that in the particular case of Uruguay, where a relatively high density of surface observations is available, station data is of higher quality than the satellite estimates and as good as the merged products. Countries with lower rain gauge densities and with more complicated topography might benefit more from satellite data and use it as a readily available, complementary dataset when no station data is available, but to obtain the best accuracy station data is still preferred. These results are consistent with the results obtained by a previous work for Uruguay done by De Vera and Terra (2012).

A side conclusion is that the high congruence between the data in the validation set and the data in the training set, correlation coefficients close to 0.9 and POD values close to 0.95, shows a high degree of homogeneity in station data and a high representativeness of the data in INUMET's database.

During the investigation, several issues and ideas arose which would need further exploration. Possible lines of further work include: i) perform separate analyses for the April-September semester (cold season) and October-March semester (warm season), ii) try other combinations of merging techniques and also other interpolation methods like optimal interpolation or spatial modelling with copulas, iii) try high resolution point interpolation to cope with extreme events, iv) use the Gamma distribution for the generalized linear model of the regression step, v) improve the re-gridding of the satellite grid by using more advanced interpolators like the Lanczos filter, and vi) explore merging with different number of stations.

1. INTRODUCTION

The present work was carried out by Pablo Alfaro (MotionSoft Consulting S.R.L. in representation of INUMET) and Alejandra De Vera (IMFIA, FING) during a 2-week internship at IRI, supervised by Tufa Dinku (IRI, Earth Institute, Columbia University). It was conducted in the framework of the DACC project.

The purpose of the visit was to work on the improvement of climate monitoring, in particular of rainfall values, using combinations of rain gauge observations and satellite rainfall estimates. This is one of the fundamental issues being addressed by the creation of the SNIA ("Sistema Nacional de Información Agropecuaria").

2. SATELLITE PRODUCT SELECTION: TRMM VS CMORPH

2.1 DATA PREPARATION

2.1.1 SATELLITE DATA

Two different satellite products were evaluated to determine which was the most representative of Uruguay's rainfall. The evaluated satellite rainfall products are:

- Tropical Rainfall Measuring Mission (TRMM) multisatellite precipitation analysis (TMPA; Huffman et al. 2007), it's near-real-time version 3B42RT.
- National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA/CPC) morphing technique (CMORPH; Joyce et al. 2004).

Both products are available at a 3-hourly frequency and a spatial resolution of 0.25º lat/long.

The TRMM and CMORPH data for the period between 2007 and 2009 were downloaded from the following websites.

- TRMM-3B42-RT:

  ftp://trmmopen.nascom.nasa.gov/pub/merged/mergeIRMicro/

  Available since 2002 up to the present, in a binary format that is compatible with GrADS binary format.

- CMORPH:

  http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.CPC/.CMORPH/.3-hourly/.mean/.morphed/.cmorph/

  CMORPH Version 0.x, available since Dec.2002 up to the present, in NetCDF format.

  ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/

  CMORPH Version 1.0, available since Jan.1998 up to the present, in GrADS binary format.

The 3-hourly data were then converted into daily totals. The present study is limited to daily precipitation totals since this is the information available at the rain gauges (daily rainfall totals are taken at 1000 UTC). The daily accumulation is obtained by adding the individual 3-h amounts from 0900 UTC of one day to 0900 UTC of the next. There is, therefore, an inevitable 1 hour lag between the satellite estimates and the rain gauge records.

2.1.2 STATION DATA

Rainfall data from 144 stations was provided by INUMET, with records starting from 1998 up to the present.

Gridded station data on the same grid as the satellite data was needed in order to perform the evaluations, for this we tested several gridding methods:

1) Interpolate at 0.25-deg using Kriging.

2) Interpolate at 0.25-deg using Inverse Distance Weighting (called IDW).
3) Interpolate at 0.05-deg using Inverse Distance Weighting and then average at 0.25-deg (called Block IDW).

2.2 VALIDATION OF SATELLITE ESTIMATES

Satellite data was compared with the interpolated gauge values, using only those pixels that contained a gauge observation for the specific day. This is meant to ensure the high quality of the reference data.

The statistics used for comparison are the following (Wilks 2006; Dinku et al. 2007; Dinku et al. 2010):

- For rainfall detection:
    - The probability of detection (**POD**), gives the fraction of events that were correctly detected. It ranges from 0 to 1 with a perfect score of 1.
    - The false alarm ratio (**FAR**), is the fraction of yes forecasts that turn out to be wrong, that proportion of the forecast events that fail to materialize. It thus ranges from 0 (best possible scenario) to 1.
    - The frequency bias (**FBS**), compares the number of events forecasted against the ones observed. Unbiased forecasts exhibit FBS = 1, indicating that the event was forecasted the same number of times that it was observed. However, it provides no information about the correspondence between the forecasts and observations.
    - The Heidke skill score (**HSS**), measures the fraction of correct forecasts after eliminating those forecasts which would be correct due purely to random chance. It ranges from -∞ to 1, 0 indicates no skill, with a perfect score of 1.
- For rainfall amount:
    - The linear correlation coefficient (**Corr**), gives a good measure of linear association or phase error. It ranges from -1 to 1 with the extremes being the best possible scenario and 0 the worst.
    - The multiplicative bias (**BIAS**), compares the average forecast magnitude to the average observed magnitude. It does not measure the correspondence between forecasts and observations. It ranges from -∞ to ∞, with a perfect score of 1.
    - The mean error (**ME**), gives the average forecast error. It ranges from -∞ to ∞, with a perfect score of 0. It does not measure the magnitude of the errors nor the correspondence between forecasts and observations, i.e., it is possible to get a perfect score for a bad forecast if there are compensating errors.
    - The mean absolute error (**MAE**), gives the average magnitude of the forecast errors, it does not indicate the direction of the deviations. It ranges from 0 to ∞, with a perfect score of 0.

- <u>For rainfall distribution</u>:
  - Empirical cumulative distribution function (**CDF**).

The following table summarizes the values of these statistics for each satellite product. We also included the values of the statistics as calculated by Tufa in his methodology.

Table 3: Validation statistics comparing the performance of daily satellite rainfall estimates over the whole of Uruguay in the period 2007-2009

Satellite Product	Gridding Method	Rainfall Amount				Rainfall Detection			
		Corr	BIAS	MAE	ME	POD	FAR	FBS	HSS
CMORPH	Kriging	**0.79**	1.29	**2.85**	1.05	0.82	**0.26**	1.11	0.71
	IDW	**0.79**	1.29	**2.85**	1.05	0.82	**0.26**	1.11	0.71
	Block IDW	**0.79**	1.29	2.86	1.05	0.82	**0.26**	1.11	**0.72**
	Tufa	**0.79**	1.35	2.86	1.19	**0.84**	0.28	1.18	0.71
TRMM	Kriging	0.72	**1.18**	3.23	**0.63**	0.74	0.31	**1.07**	0.63
	IDW	0.72	**1.18**	3.24	**0.63**	0.74	0.31	**1.07**	0.63
	Block IDW	0.72	**1.18**	3.24	**0.63**	0.74	0.31	**1.07**	0.63
	Tufa	0.71	1.22	3.22	0.77	0.76	0.33	1.14	0.63

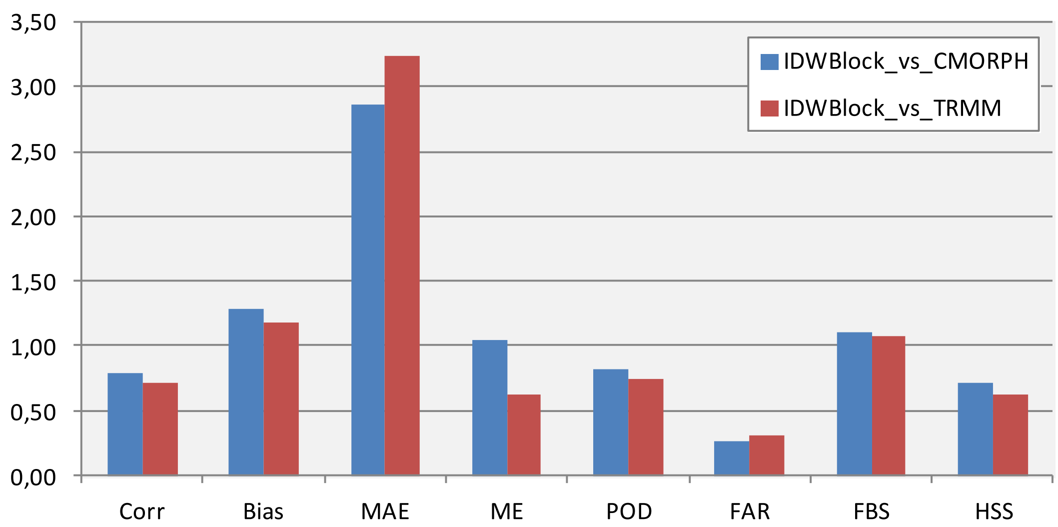


Figure 1: Comparison of the performance of satellite rainfall estimates at the daily time scale in the period 2007-2009

The scatterplot of the satellite rainfall estimates against the observed values was also used, to measure how well the satellite estimates correspond to the observed values (an accurate forecast will have points clustering on or near the diagonal).

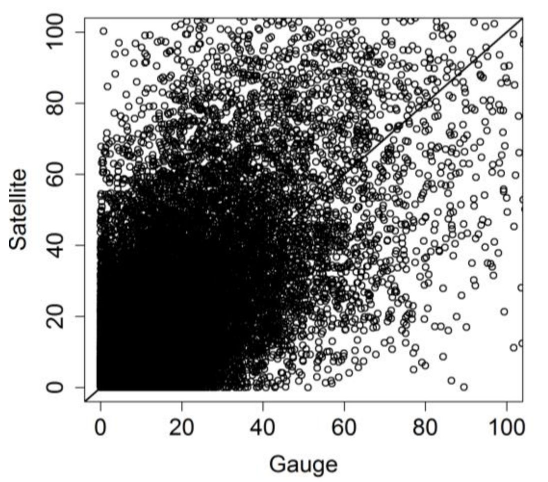
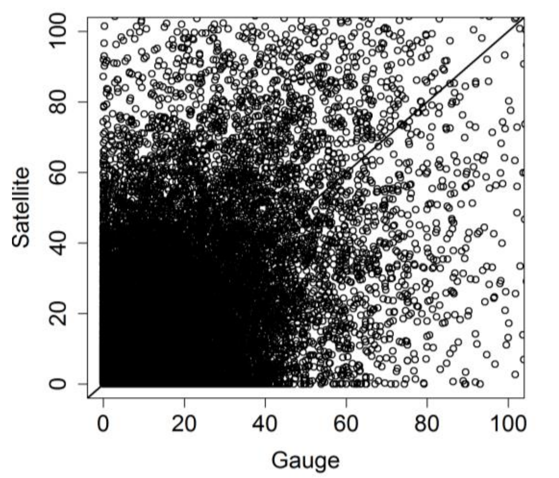
Scatterplot	
CMORPH	**TRMM**
	

Figure 2: Scatterplots of the satellite rainfall estimates against the observed values at the daily time scale in the period 2007-2009

The following figures show the maps with the stations' values, the interpolated gauge values (using Block IDW) and the satellite rainfall estimates (TRMM and CMORPH), for some particular days.

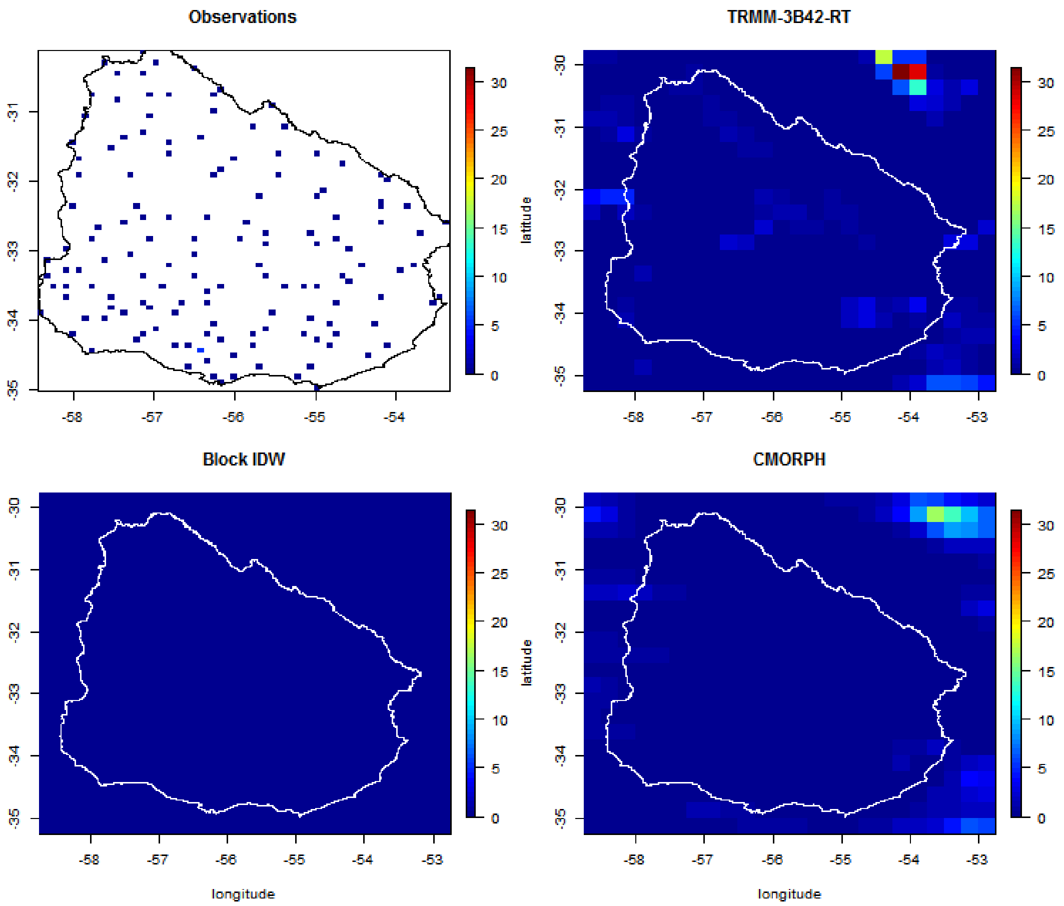


Figure 3: Observations, Block IDW, TRMM and CMORPH for 27/01/2007

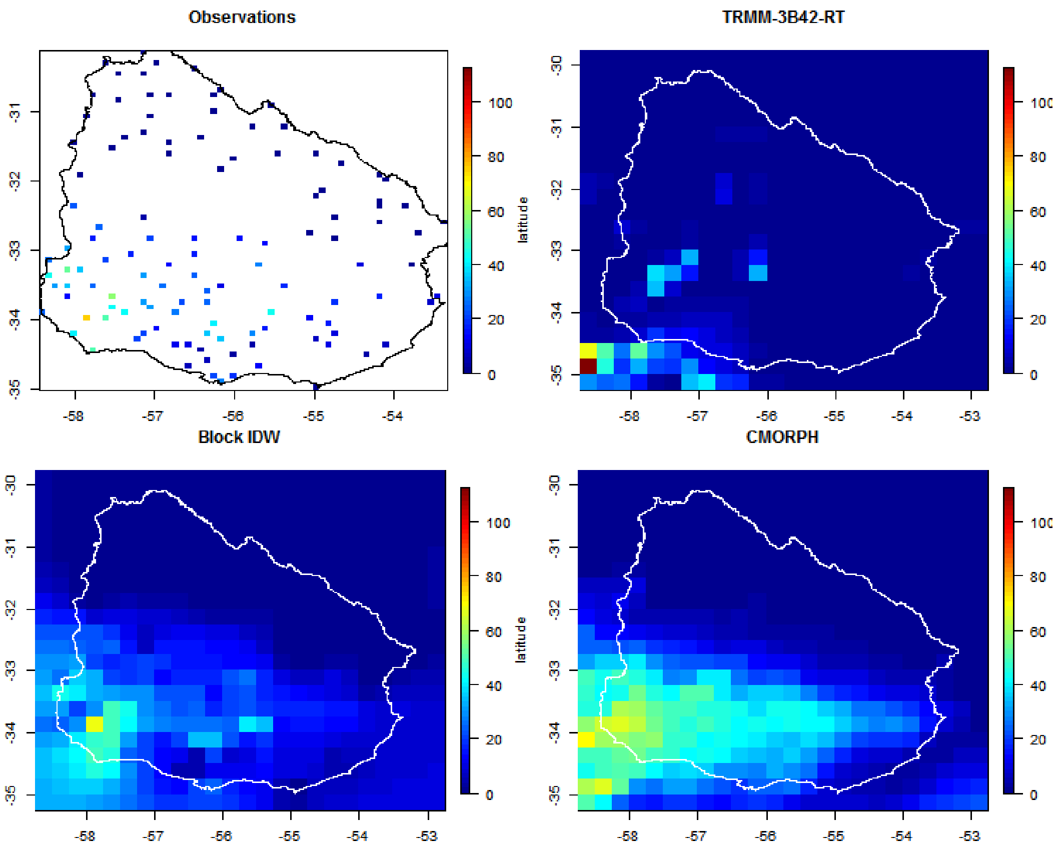


Figure 4: Observations, Block IDW, TRMM and CMORPH for 15/01/2008

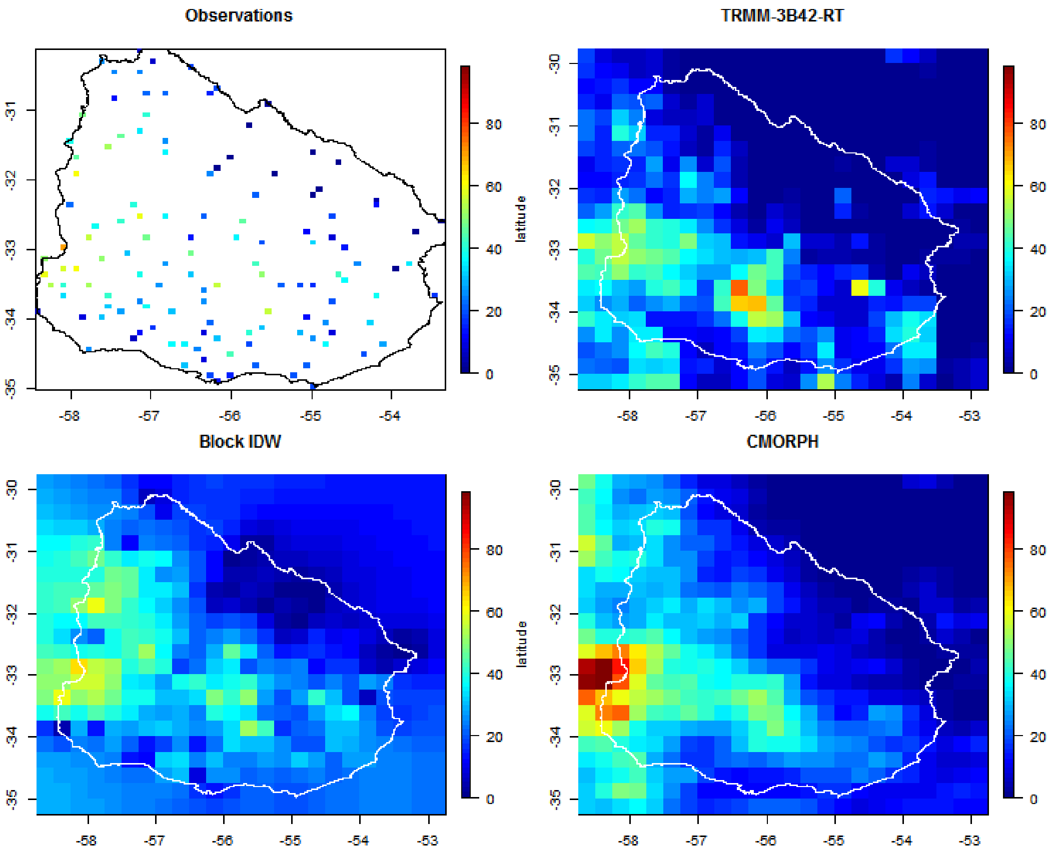


Figure 5: Observations, Block IDW, TRMM and CMORPH for 19/12/2009

2.3 SELECTION

Both satellite products have exhibited good performance over Uruguay. In fact, these results are unusually high. Comparisons of the two satellite products have shown that CMORPH has a slightly better performance, particularly for the correlation coefficient and POD.

Therefore, CMORPH has been selected hereafter in view of the higher skill in the parameters considered most relevant to the purpose of the study.

2.4 CMORPH V0.X VERSUS CMORPH V1.0

There are two versions of CMORPH:

- CMORPH Version 0.x: available since 2002 up to present.
- CMORPH Version 1.0: a reprocessing of the CMORPH data with a fixed algorithm, that is similar to the one being used by CMORPH v0.x in the present, and using inputs of the same versions, available since 1998 up to present.

Once the CPC finishes releasing the reprocessed CMORPH to cover the entire data period from Jan 1998 to Dec 2012, they will terminate the production of Version 0.x. However, since the fixed algorithm used to generate the Version 1.0 is very close to the currently operational algorithm being used to produce the CMORPH Version 0.x for recent years, the differences between Version 0.x and Version 1.0 should be minor (ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/CMORPH_V1.0_README.txt).

The following table and figure shows the values of the validation statistics comparing the performance of the two CMORPH versions for different time periods.

Table 4: Validation statistics comparing the performance of the two CMORPH versions (Version 0.x versus Version 1.0) for different time periods

Version	Period	Rainfall Amount				Rainfall Detection			
		Corr	Bias	MAE	ME	POD	FAR	FBS	HSS
CMORPH v0.x	2007-2009	0.79	1.29	2.86	1.05	0.82	0.26	1.11	0.72
CMORPH v1.0	2007-2009	0.76	1.28	2.91	1.00	0.82	0.27	1.13	0.70
CMORPH v1.0	2010-2012	0.79	1.25	2.70	0.87	0.85	0.29	1.20	0.70
CMORPH v1.0	1998-2000	0.73	1.37	3.39	1.38	0.83	0.38	1.35	0.61
CMORPH v1.0	1998-2012	0.77	1.34	3.05	1.25	0.84	0.31	1.22	0.69

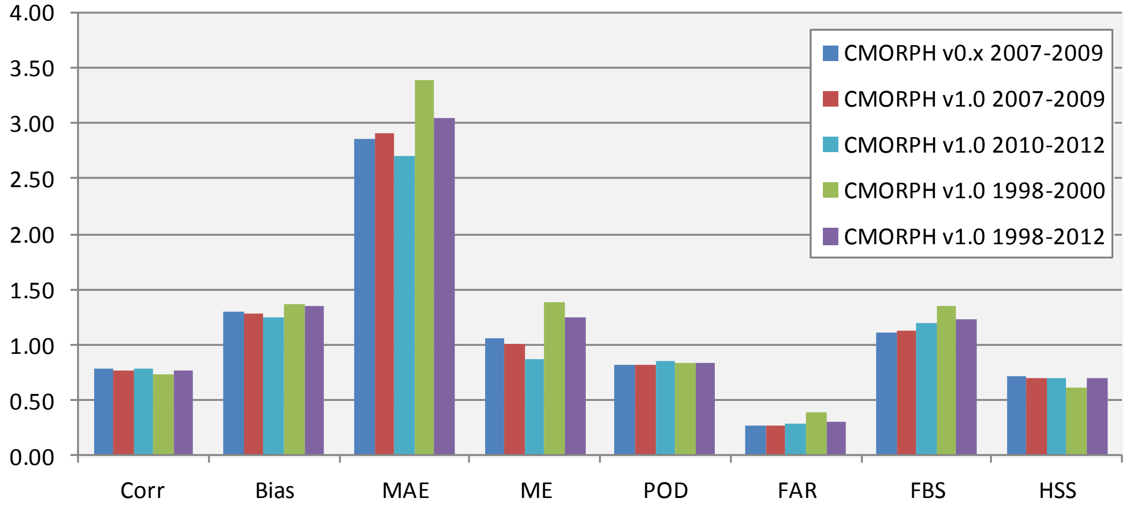


Figure 6: Validation statistics comparing the performance of the two CMORPH versions (Version 0.x versus Version 1.0) for different time periods

These results confirm there are no significant differences among the two versions (period 2007-2009) and the different periods. Hereafter, in view of the larger data period available, we will work with CMORPH v1.0 to explore the different merging techniques.

3. MERGING TECHNIQUES

Interpolation of daily rainfall data is challenging for several reasons. Two of the potential challenges include overestimation of the occurrence or spatial extent of rainfall and underestimation of high rainfall values.

Satellite rainfall estimates, with appropriate techniques, could help to alleviate these problems. Though their accuracy is not very reliable at a daily timescale, they can provide information about the spatial structure of rainfall, including where rainfall did or did not occur. These characteristics are used to improve on the interpolation of station measurements.

There are many different approaches to merging station data with satellite estimates. In this work some of them were explored, starting with some of the simplest ones and progressively including and combining more elaborate techniques.

After the merging was complete, the same approach used to evaluate TRMM and CMORPH was used to evaluate the results.

The products we tested were a station-only interpolation, the raw satellite estimation and several merged products.

3.1 GENERAL CONSIDERATIONS

The following considerations are prior to the merging and apply for the whole procedure.

3.1.1 SPLITTING STATION DATA INTO TRAINING AND VALIDATION SETS

To test the different spatial prediction approaches, station data was divided randomly into a training and a validation sets. One third of the stations were used for training and two thirds for validation.

The rationale behind this was to try to reproduce ‘reality’ as faithfully as possible by having a large number of stations in the validation set. This methodology aims at gaining an insight of how much improvement can be gained by incorporating satellite data, rather than evaluating the absolute performance of one or another method.

The actual performance of the final product is expected to be better than this evaluation suggests since it will incorporate the complete station data set, however this evaluation will provide a means to selecting the best method among the possibilities taken into account.

3.1.2 GRIDDING OF STATION DATA

All comparisons made between satellite (or merged) estimates data and station rainfall data were made against a gridded set of the validation station data, which were obtained through a block interpolation to the same grid as the satellite data. Under this block schema, each of the grid cells is subdivided into smaller sub cells, on which station data is interpolated, and then the values of the sub cells are averaged over the entire original cell to obtain it’s value.

This is to achieve a more fair comparison with satellite data, since the latter is an areal average for each of the pixels which comprise it rather than a point observation.

3.1.3 TIME WINDOW

Some of the steps used in the merging require a statistical calibration for their use (ECDF, calculation of regression parameters, variogram modelling). In order to provide more robust estimations of their parameters a moving time window was used to include in the fitting process data from 5 days before and after the actual day being calculated.

3.2 MERGING METHODOLOGIES

The merging technique which was used in this work involves four basic steps:

- Bias removal from satellite grid.
- Regression of the station data on the (unbiased) satellite data.
- Interpolation of the regression residuals at station locations to the entire grid.
- Application of a rain/no rain mask (RNR) to prevent the overestimation of the occurrence of rainfall.

These steps will be discussed further in this section. The following variables will be used throughout the section to clarify explanations:

- Point Data Sets
    - Validation Set: VS
    - Training Set: TS
- Gridded Data Sets
    - Gridded Station Data: Gr(set) where set is one of VS, TS or CS
    - Satellite Estimates: SE
- Point Values Extracted from Gridded Data Sets
    - Gridded Data at Station Locations: StnLoc(grid) where grid is one of the previously defined gridded data sets

3.2.1 BIAS REMOVAL

Two approaches were used to remove bias on the gridded satellite rainfall estimations.

<u>Simple bias removal</u>

The first one, which we named 'Simple bias removal' consisted of calculating the difference between the station data and the satellite estimation at the station locations, then interpolating the differences. These differences (biases) are then added to the original satellite estimate to obtain the unbiased satellite estimates.

A highly simplified pseudo-code for this would be:

$$residuals = TS - StnLoc(SE)$$
$$UnbiasedSE = Gr(residuals) + SE$$

CDF matching

The second approach, CDF matching, aims at matching the CDF of the satellite estimates with that of the gridded station data.

On one hand, the ECDF of the gridded station data is calculated and in the other the quantiles in the satellite estimation's ECDF of the current satellite observation, then each satellite observation is substitued by the value of it's quantile in the ECDF of the actual rainfall data to obtain the unbiased satellite estimation.

$$UnbiasedSE = ECDF^{-1}(TS, ECDF(SE, SE))$$

Where:

$ECDF(X, Y)$ is the Empirical Cumulative Distribution Function of sample X evaluated at sample Y.

$ECDF^{-1}(X, Y)$ is the Inverse Empirical Cumulative Distribution Function of sample X evaluated at sample Y.

3.2.2 REGRESSION (GENERALIZED LINEAR MODEL)

Given the biased or unbiased satellite estimation grid, a regression was carried out in order to fit the satellite estimates to the observed rainfall values.

The regression method used was a Generalized Linear Model (GLN) with the following formula:

$$StnLoc\big(Gr(TS)\big) = a * StnLoc\big(\widetilde{SE}\big) + b + E$$

Where:

$\widetilde{SE}$ are the biased or unbiased satellite estimates.

a and b are the regression coefficients.

E is an error term following a random distribution from the exponential family and having zero mean. In this case the Gaussian distribution was used for simplicity but the evaluation of other distributions like the Gamma distribution remains a future work.

Once the regression model was fit, it was used to predict regressed values for the whole grid by means of the following formula:

$$Reg\big(\widetilde{SE}\big) = a * \widetilde{SE} + b$$

Where:

$Reg\big(\widetilde{SE}\big)$ are the regressed satellite estimations (biased or unbiased) for the whole grid.

3.2.3 UNIVERSAL GRIDDING

After the regression, interpolation of the residuals was carried out. In order to do this Universal gridding[1] was used.

In the Universal Gridding schema, data at location (x, y) can be modelled using the following equation:

$$Z(x, y) = U(x, y) + E(x, y)$$

Where:

$Z(x, y)$ is the value of the variable being modelled at location (x, y); rainfall in our case.

$U(x, y)$ is a spatial trend function, which is known for every (x, y) and gives the expected value of $Z(x, y)$.

$E(x, y)$ is a random error term with zero mean.

For $U(x, y)$ we used the regressed satellite estimation $Reg\left(\widetilde{SE}\right)$ and since we know the values of Z for the station locations (x_{stn}, y_{stn}), the error term at the station locations can be calculated:

$$E(x_{stn}, y_{stn}) = U(x_{stn}, y_{stn}) - Z(x_{stn}, y_{stn})$$

Once the error terms are obtained, they can be interpolated to the whole grid using any interpolation technique to obtain $E^{*}(x, y)$, an estimate of the error term for any (x, y). In this work both Kriging and Inverse Distance Weighting were used for the interpolation, but others such as triangular interpolation and optimal interpolation could have been used as well. Using these interpolated error terms the value of $Z^{*}(x, y)$ for any (x, y) can be obtained by:

$$Z^{*}(x, y) = U(x, y) + E^{*}(x, y)$$

This is carried out for the whole satellite grid to obtain the quasi-final, unmasked merged product. As before the block interpolation schema explained in section 3.1.2 was used since the values being predicted are area averages.

3.2.4 RNR MASK

The final step in merging the data consisted of estimating a rain/no rain mask and applying it to the unmasked merged product.

All the interpolation methods used in this work model their output as a weighted average of the values observed at the known locations, differing only in the means of obtaining these weights. Typically, the longer the distance from the observation the smaller it's weight. Given this, these weighted averages tend to give very small but positive values around zero observations. In order to overcome this problem a Boolean rain mask is estimated to determine the locations where rain is believed to have occurred and were it is not. After the mask's (0/1 values) were determined, the

[1] The actual standard method is Universal Kriging, but here Universal "Gridding" is used because other interpolation methods besides Kriging were used even when the Universal component was used with all of them.

unmasked merged product was multiplied by the mask to obtain the final masked merged product.

Three approaches were taken to estimate the RNR mask:

- Stations only

  The first approach used station only data and consisted of creating a binary data set from the observations of the training set, where values of 1 correspond to positive rainfall values and values of 0 correspond to zero rainfall values. This binary data set was interpolated using Inverse Distance Weighting to obtain a gridded data set in the [0, 1] range for the whole area. Finally, to obtain the RNR mask, grid cells whose interpolated value was below a given threshold were assigned a value of 0 (no rain) and those above the threshold a value of 1.

  In this work we used a fixed value of 0.5 for the threshold, however as a future work it could be estimated optimally from the observed data, perhaps using a conditional probability approach.

- Satellite only

  The second approach used only satellite data and involved determining the 0/1 values applying the threshold but to the satellite estimation directly rather than the interpolated 0/1 values.

- Combine both

  The last approach combined the first two methods by determining the station and satellite masks separately, and then for those pixels whose centroid was closer than a given distance to any station use the station mask's value and for those who weren't the satellite mask's value.

3.3 RE-GRIDDING (10 KM)

In a nonformalized experiment, we tried resampling the satellite estimates from 0.25x0.25 degree to 0.1x0.1 degree using bilinear interpolation. Then the resampled satellite estimate was used for merging.

This method produced "nicer looking" maps in the sense that the variations in the rainfall field were smoother and the pixels were not as evident, but the values of the evaluated statistics didn't differ significantly from those obtained using the 0.25x0.25 grid.

3.4 EXPERIMENTS AND RESULTS

To evaluate the different products the same validation statistics as discussed in section 2.2 were used.

The station data set used to estimate the different merged products was the training set, the data set used to evaluate them was the block-gridded validation set, and the satellite data set used was CMORPH Version 1.0.

The following combinations of merging methodologies were tested for the 2005-2009 period to determine which combination worked best:

i. **IDW_Raw**: Raw satellite estimate + Regression + Block IDW Interpolation + Satellite-only mask.

ii. **IDW_CDF_1**: CDF matched satellite estimate + Regression + Block IDW Interpolation + Satellite-only mask.

iii. **Kriging_CDF**: CDF matched satellite estimate + Regression + Block Kriging Interpolation + Satellite-only mask.

Table 5 summarizes the obtained results. The values of the statistics for the Raw CMORPH estimation are included for reference.

Table 5: Validation statistics for the different merged data sets for the for the 2005-2009 period

Version	Rainfall Amount				Rainfall Detection			
	Corr	Bias	MAE	ME	POD	FAR	FBS	HSS
CMORPH V1.0	0.78	1.41	3.61	1.69	0.83	**0.27**	**1.14**	0.69
IDW_Raw	**0.87**	1.06	1.84	0.19	**0.89**	0.30	1.28	0.72
IDW_CDF_1	**0.87**	**1.03**	**1.79**	**0.12**	**0.89**	**0.27**	1.22	**0.74**
Kriging_CDF	0.86	1.04	1.86	0.16	**0.89**	**0.27**	1.22	**0.74**

As can be seen, **IDW_CDF_1** shows a slightly better result for most of the statistics and thus it will be the merging schema used. All the rainfall amount statistics, the POD and the HSS, showed improvement on all of the merged techniques over the raw satellite estimates. The FAR statistic didn't show a significant reduction with any of the merging techniques and the FBS showed a small performance loss.

The final experiment compared the gridded training set, the raw satellite estimate and the merged product generated by combining CDF matching, regression, block IDW interpolation and the station-only mask. The compared data sets are:

i. **Gridded Training Set**: Gridded set of the training station data using IDW Interpolation.

ii. **CMORPH V1.0**: Raw satellite estimates.

iii. **IDW_CDF_2**: CDF matched satellite estimation + Regression + Block IDW Interpolation + Station only mask.

Figure 7 shows, as an example, a set of maps with the training and validation sets observations', the gridded training set, the gridded validation set, the raw satellite rainfall estimates and the merged set for 23/05/2012.

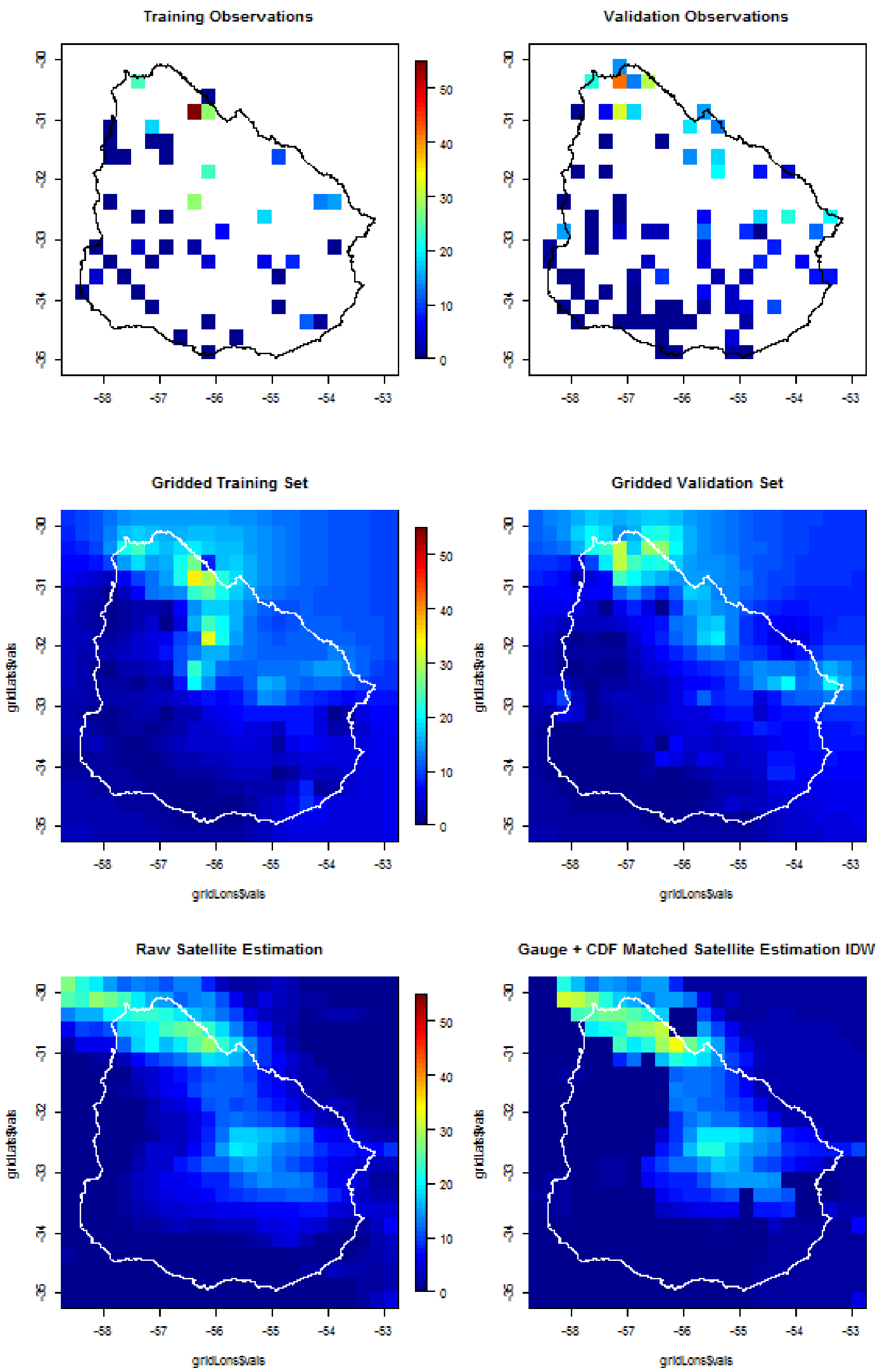


Figure 7: Maps with the training and validation observations, the gridded training set, the gridded validation set, the raw satellite rainfall estimates and merged set for 23/05/2012

Table 6 summarizes the validation statistics obtained for each period for each data set.

Table 6: Validation statistics for the gridded training set, the raw satellite estimation and the merged set (CDF matched + Regression + Block IDW + Station only mask)

Period	Data set	Rainfall Amount				Rainfall Detection			
		Corr	Bias	MAE	ME	POD	FAR	FBS	HSS
1998-2000	Gridded Training Set	0.88	1.04	1.69	0.13	0.94	0.24	1.23	0.79
	CMORPH v1.0	0.73	1.38	3.37	1.40	0.84	0.39	1.39	0.61
	IDW_CDF_2	0.87	1.08	1.83	0.29	0.83	0.08	0.91	0.84
2005-2009	Gridded Training Set	0.88	1.01	1.64	0.04	0.93	0.22	1.19	0.80
	CMORPH v1.0	0.78	1.41	3.61	1.69	0.83	0.27	1.14	0.69
	IDW_CDF_2	0.86	1.02	1.69	0.06	0.78	0.07	0.84	0.81
2010-2012	Gridded Training Set	0.89	1.02	1.57	0.07	0.93	0.24	1.22	0.79
	CMORPH v1.0	0.78	1.24	2.69	0.84	0.85	0.30	1.20	0.70
	IDW_CDF_2	0.87	1.03	1.64	0.11	0.78	0.08	0.85	0.81
Overall	Gridded Training Set	0.88	1.02	1.69	0.06	0.94	0.23	1.22	0.80
	CMORPH v1.0	0.77	1.34	3.05	1.23	0.85	0.31	1.23	0.69
	IDW_CDF_2	0.80	1.05	1.77	0.17	0.81	0.08	0.88	0.83

As can be seen in Table 6, there are no significant differences between gridded station data and the merged product. Even though the gridded station data has a better POD, the merged data exhibits a better FAR. This may be ascribed to a reasonably well-distributed station network and relatively flat topography. It also seems that the spatial variation of rainfall is relatively low. All these factors contribute to the better-than-expected performance of the gauge-only product

The merged **IDW_CDF_2** product got better results for the FAR, FBS and HSS statistics. This could largely be due to the RNR mask since it was not incorporated in the gridding of the training set. The small positive values around zero observations are likely to be zeros and that would be the reason behind the decrease in FAR and FBS. This suggests the incorporation of the RNR mask in the current rainfall interpolation methods to get rid of these false positives. In addition, the merged set shows the worst POD in every tested situation. The reasons behind this are not entirely clear and should be investigated further but it’s possible that the same RNR mask that improves the FAR also hinders the POD, discarding real positive rainfall values in some cases. Investigating into better methods for determining the RNR mask’s threshold could help alleviate this.

The **IDW_CDF_2** set shows an improvement among all the statistics except POD when compared to the raw satellite estimation. This shows an overall increase of skill -which is particularly large for the MAE and ME statistics- achieved by using the proposed merging techniques and station data, indicating an improvement in the accuracy of the satellite estimate owing to incorporation of station data.

4. CONCLUSIONS

Two different satellite products, the real-time version of TRMM (3B42RT) and CMORPH, were evaluated to determine which one is the most suitable for Uruguay. Evaluation was done at daily time scale and a spatial resolution of 0.25º lat/long. Validation statistics were computed for the whole country.

The comparisons between the two satellite products have shown that CMORPH has a relatively better performance with high (good) values of some of the statistics, particularly correlation and POD. As a result, CMORPH was selected to explore the different merging techniques.

Several merging techniques were explored by implementing a series of merging steps, namely: bias removal (simple bias removal or CDF matching), regression analysis (generalized linear model), universal gridding (kriging or inverse distance weighting) and block interpolation, and the application of a rain/no rain mask (station data only, satellite data only or a combination of both). To test the different approaches, station data was divided randomly into a training set having one third of the observations and a validation set having the remaining two third. Next, validation statistics were computed for the whole country using the validation set as the reference data.

The results show an overall increase of skill when satellite data was merged with station measurements. Still, the block inverse distance weighted interpolation from the training set (using station only data) exhibited a similar result to that of the merged product.

Therefore, we can conclude that in the particular case of Uruguay, where a relatively high density of surface observations is present, merging does not offer a better option than station-only gridded data. These results are consistent with results obtained in a previous work for Uruguay done by De Vera and Terra (2012).

A side conclusion is that the high agreement (correlation values close to 0.9 and POD values close to 0.95) between the data in the validation set and the data in the training set shows a high degree of homogeneity in station data and a good quality of the station data in INUMET’s archives.

The simple inverse distance weighting method used to create the gridded training set shows some problems regarding the FAR statistic, but, in order to improve the ability of the rain gauge network, incorporating the RNR mask could help to avoid the overestimation of rainfall occurrence. The underestimation of high rainfall values still remains to be a problem and has to be dealt with.

5. FUTURE WORK

Along the investigation several issues and ideas arose which could be addressed with further exploration. Possible lines of further work mentioned are:

- Seasonal analysis

  Rainfall in Uruguay follows two major distinct patterns according to seasons. In the cold season (April-September) rainfall is mostly stratiform with smaller rainfall amounts spread relatively uniformly over space, whereas in the warm season (October-March) rainfall is mostly convective, having larger amounts over shorter times and concentrated into smaller, more punctual areas. These two patterns should be analized separately to determine the performance and the gain of the merged products for each of them.

- Other combinations of merging techniques

  For reasons of time, some other merging techniques were not tested. In particular, the combined station-satellite rain/no-rain mask may be applied to improve the merging thechniques to evaluate its performance. Other interpolation methods like optimal interpolation or the spatial modelling with copulas proposed by (Kazianka et al. 2010), which is designed speciffically for dealing with extreme value distributions, may need to be tested.

- High resolution point interpolation for extreme events

  As was explained earlier, the block-gridding technique was used to produce areal averages over the grid cells. This imposes a penalty on the detection of extreme rainfall values as these values may be smoothed out by the averaging. To overcome this issue, a method similar to the simple bias removal can be used; but using the merged product as the base grid and using a higher resolution grid instead of the block scheme for the interpolation of the residuals.

- GLM using gamma distribution

  Ten-daily and monthly rainfall data have been shown to fit quite well to the Gamma distribution. For daily rainfall data this fit isn't as good but still it's a much better fit than the Gaussian distribution. Using the Gamma distribution for the GLM model of the regression step may yield better results.

- Re-gridding of the satellite grid

  For the re-gridding of the satellite grid a bilinear interpolation was used. This could be improved upon by using more advanced interpolators like the Lanczos filter. This filter provides very good results in the field of image resampling and could help improve the resampling.

6. REFERENCES

De Vera A. and Terra R., 2012: Combining CMORPH and Rain Gauges Observations over the Rio Negro Basin. *J. Hydrometeor.*, **13**, 1799-1809.

Dinku T., P. Ceccato, E. Grover-Kopec, M. Lemma, S. J. Connor and C. F. Ropelewski, 2007: Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.*, **28**, 1503–1526.

Dinku T., Ruiz F., Connor S. J. and Ceccato P., 2010: Validation and Intercomparison of Satellite Rainfall Estimates over Colombia. *J. Appl. Meteor. Climatol.*, **49**, 1004–1014.

Huffman, G.J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeor.*, **8**, 38–55.

Joyce, R.J., Janowiak J.E., Arkin P.A., and Xie P., 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487-503.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 627 pp.

Kazianka, H. & J. Pilz, 2010: Copula-based geostatistical modeling of continuous and discrete data including covariates, Stochastic Environmental Research and Risk Assessment 24(5), 661 – 673.