



Geostatistical integration to improve representativeness of satellite precipitation estimates and field measurements

Felipe Omar Tapia Silva^{1,2}

¹Laboratorio de Geomática Aplicada-Depto. Hidrobiología. Universidad Autónoma Metropolitana-Iztapalapa, CDMX, México.

²CentroGeo Campus CDMX, México (sabbatical stay).

otapia@xanum.uam.mx, <http://www.izt.uam.mx/>,
otapia@centrogeo.edu.mx

Abstract.

Since the network of rainfall gauges and ground radars is generally not dense enough, satellite data have been used to estimate Precipitation (P). These data have the ability to capture the spatial variability pattern of the parameter, but are often inaccurate in relation to the value of the field measured parameter. Therefore, geostatistical methods were evaluated to improve the spatial representativeness of field measurements (FM) and satellite estimates. The work has been made for a hydrological sub region in the Mexican tropic. The geostatistical methods used to interpolate P-FM were ordinary kriging (KO), universal kriging (KU) and regression kriging (RK) as well as the Inverse Distance Weighted (IDW) mechanical interpolator for comparison purposes. Furthermore, the values at the pixel centers of the Tropical Rainfall Monitoring Mission (TRMM) images were interpolated using OK and evaluated using leave-one-out cross validation (LOO-CV). The best LOO-CV evaluated method consisted of the RK interpolation of the point FM taking as auxiliary variable the OK interpolation of the TRMM cell centers. It is concluded that the geostatistical integration between rainfall estimates from satellite data and FM data is promising because satellite information has the ability to capture spatial variability and the point FM add accuracy to the results. These characteristics combined can produce a P product useful for modeling activities and environmental management.

1 INTRODUCTION

Sustainable planning and environmental management require reliable estimates of P and other elements of the hydrological cycle. P represents a life-critical energy and hydrologic exchange between the Earth's atmosphere and its surface and that knowledge of where, when and how much rain falls is essential for scientific research and societal applications (Skofronick-Jackson et al. 2018). However, a better understanding of the spatial and temporal P patterns is still necessary in order to quantify risks and design suitable mitigation measures in a context of P pattern change (Agou et al. 2019).

Gauges that measure P at a point remain the most common approach to ground-based measurement and are the ultimate reference and the only measurement method available in many regions of the world (New et al. 2001). However, the direction and magnitude of regional climatic trends cannot be reliably inferred from single-site records, even over relative homogeneous terrain (Pielke et al. 2000). The representativeness of an observation is the degree to which it accurately describes the value of the variable needed for a specific purpose (WMO, 2008). Synoptic observations should be representative of an area up to 100 km around the station, but for small-scale or local applications the considered area may have dimensions of 10 km or less (WMO, 2008). WMO (2008) also indicates that achieving good location of stations for data collection is difficult. Unfortunately, the situation in Mexico regarding the representativeness of field values taken in the stations is unfavorable. The information measured in ground stations does not meet the appropriate representativeness criteria that allow the generation of data, information and knowledge to support better decision making. An example is the case mentioned by Tapia-Silva et al. (2013) about the National Meteorological Service (SMN) stations that include approx. 3,300 observation sites. Assuming that each station is representative of an area of 100 km², as defined by WMO (2008) for flat areas, a total area of 330,000 km² would be covered, which represents only 17% on the territorial extension of Mexico. Therefore, it is necessary to use other information sources and to apply spatial analysis methods in order to better capture the spatial variability of this parameter at regional level.

Geostatistical interpolation (kriging) has become very important to environmental studies (Matheron, 1969), since it has been defined as the best unbiased linear estimator (Cressie, 1990; Hengl, 2007). These techniques have been successfully explored to generate more representative P spatial layers from point measurements (Holawe & Dutter 1999, Goovaerts, 2000, Keblouti et al. 2012, Agou et al. 2019). However, its application in countries such as Mexico presents problems in capturing spatial variability due to the aforementioned problem regarding the quantity and location of field stations. Most of these were installed in accessible areas without making a correct geographic sampling design. An additional problem is that many of them are not functioning anymore.

Rain estimates by satellite media have been intensely investigated since the 70's (New et al. 2001). One of the emblematic missions has been TRMM (Kummerow et al., 2000) which has a satellite-borne active microwave system as well as a passive microwave radiometer. However, rainfall satellite products are not error-free (Anagnostou et al., 2010; Zulkafli et al., 2014). This is due to factors such as spatial resolution and the used estimation methods. The spatial resolution produces generalization of the value for a geographical area, for example, TRMM offers rain products with information for surfaces of 4 km². About the estimation methods implemented, Greene and Morrysey (2000) mention that a large uncertainty is associated with satellite P estimates, stemming from unknown variations in space and time of the physical and statistical relationships between P and satellite-sensed radiance.

FM have been frequently used to validate or compare P information obtained by satellites (e.g. Laurent et al. 1998, Bowman, 2005, Bell, 2003). Alternatively, some integration schemes of FM and satellite estimates have been developed as this reported by New et al. (2001) consisting in a procedure that weights the individual input components by the inverse of the random error to produce a final merged product. Other merging method was reported by Wu et al. (2018). In this method, the P from rain gauges

is added to the first-guess field when combining the P estimates of TRMM Multi-Satellite P Analysis (TMPA) 3B42 with rain gauges. More recently, Sivasubramaniam et al. (2019) merged a nonparametric k-nearest neighbor (k-*nn*) estimation of radar P with thin plate spline interpolated gauge P.

Nevertheless, better integration procedures can be developed. The integration can be based on the fact that the information obtained by satellite allows capturing the spatial variability of parameters that can be determined based on the interaction of electromagnetic energy with rain-laden clouds and that FM capture their accuracy (Nerini et al. 2015, Wu et al. 2018). Hengl et al. (2007) reports the possibility to use RK to interpolate FM using estimations from satellite images as auxiliary variable. This type of geostatistical integration schemes using kriging are innovative and only a few publications such as Wang and Lin (2015) and Lin and Wang (2011) have reported its implementation.

Considering the above, in this investigation geostatistical methods are applied with the aim of improving the spatial representativeness of the P geographical layers. The methods investigated are OK, UK and RK. The IDW mechanical interpolator is also included for comparison purposes. As a main result, a geostatistical integration scheme between the FM and P-estimates from TRMM is developed. The objective of this work is the answer to the following questions: Is it possible to propose a geostatistical integration method between FM and TRMM satellite P estimates? How do the results of this method compare with those of commonly used methods as IDW, OK and UK?

1.1 Theory of Kriging

Kriging accounts for local variations in the mean by limiting the domain of stationarity to a local neighborhood, Ω , around the position, x , where the variable is to be estimated. Let $Z(x) = Y(x) + m(x)$ be a stochastic process with a variable mean that is determined by $m(x)$ and the covariance function $C(h)$. As such, $Y(x)$ is a stochastic process with a null mean. A linear estimator is a linear combination of measurements $Z(x_1), Z(x_2), \dots, Z(x_n)$ at positions $x_1, x_2, \dots, x_n \in \Omega$. Specifically:

$$\hat{Y}(X) = \sum_{k=1}^n \lambda_k(X) Y(X_k) \quad (1)$$

or:

$$\hat{Z}(X) = m(X) + \sum_{k=1}^n \lambda_k(X) [Z(X_k) - m(X_k)] \quad (2)$$

If the mean is constant in domain Ω , then it can be eliminated from the equation above by forcing the kriging weights to sum to one, in which case, the estimator is called OK, and is expressed as:

$$\hat{Z}_{KO}(X) = \sum_{k=1}^n \lambda_k(X) Z(X_k) \quad (3)$$

with:

$$\sum_{k=1}^n \lambda_k(x) = 1 \quad (4)$$

The Lagrange multiplier method (Goovaerts, 1997) is used to obtain the optimal weights that minimize the estimation of error variance, which results in the following system of equations:

$$\begin{cases} \sum_{k=1}^n \lambda_k(X) C(X_j - X_k) + \mu(x) = C(X_j - X), & j=1, \dots, n \\ \sum_{k=1}^n \lambda_k(X) = 1 \end{cases} \quad (5)$$

where μ denotes the Lagrange multiplier. Alternatively, when considering the relation between the covariance function and the semivariogram function $\gamma(h)$, i.e., $C(h) = C(0) - \gamma(h)$, then the above system can be written as:

$$\begin{cases} \sum_{k=1}^n \lambda_k(X) \gamma(X_j - X_k) + \mu(x) = \gamma(X_j - X), & j=1, \dots, n \\ \sum_{k=1}^n \lambda_k(X) = 1 \end{cases} \quad (6)$$

OK assumes a stationary mean, that is, it is a constant of the random function, $Z(x)$, of the real underlying value. But it is often not constant throughout the entire study area. When that is the case, a non-stationary regionalized variable has two components (Davis, 1973): drift (average or expected

value of the regionalized variable) and residual (difference between the values of the parameter that are considered to be real and drift).

UK divides the random function into a linear combination of deterministic functions: the smooth and non-stationary trend (drift or mean) $\mu(x) \in \mathbb{R}$, and the residual random function $Y(x) = Z(x) - \mu(x)$ (Wackernagel, 2003). UK assumes that $\mu(x)$ is a function of the spatial location, and it can be approximated by a model, for example, that of Kumar (2007):

$$\mu(X) = \sum_{i=1}^n \alpha_i f_i(X) \quad (7)$$

where: α_i = coefficient to be estimated based on the data
 f_i = basic function of drift as a function of the spatial coordinates
 n = number of functions used in the drift model

As with OK, the weights in UK are obtained by minimizing the variance of the prediction error subject to the unbiasedness restriction. The Lagrange multiplier is applied once again, taking into consideration the spatial autocorrelation structure in order to obtain the optimal weights.

RK is the best Linear Unbiased Prediction (BLUP) model for spatial data and all other techniques as OK, IDW, etc. can be seen as its special cases (Hengl, 2007, Hengl, 2009). In matrix notation, regression-kriging is commonly written as:

$$\hat{Z}_{RK}(x_o) = q_o^T \cdot \hat{\beta}_{GLS} + \lambda_o^T \cdot (Z - q \cdot \hat{\beta}_{GLS}) \quad (8)$$

$\hat{Z}_{RK}(x_o)$ is the predicted value at location x_o , q_o the vector of $p + 1$ predictors, $\hat{\beta}_{GLS}$ are regression coefficients estimated with OLS (Ordinary Least Squares) or optimally with GLS (Generalized Least Squares) and λ_o is the vector of n kriging weights used to interpolate the residual. It has a prediction variance that reflects the position of new locations (extrapolation) in both geographical and feature space (Hengl, 2007, Hengl, 2009):

$$\hat{\sigma}_{RK}^2(x_o) = (C_o - C_1) - c_o^T \cdot C^{-1} \cdot c_o + (q_o - q^t \cdot C^{-1} \cdot c_o)^T \cdot (q^t \cdot C^{-1} \cdot q)^{-1} \cdot (q_o - q^t \cdot C^{-1} \cdot c_o) \quad (9)$$

Where $C_o + C_1$ is the sill variation and c_o is the vector of covariances of residuals at the unvisited location.

According to Hengl (2007), if the residuals show no spatial auto-correlation (pure nugget effect) RK converges to pure multiple linear regression because the covariance matrix (C) becomes identity matrix. Hengl (2007) also points out that if the target variable shows no correlation with the auxiliary predictors, the RK model reduces to OK because the deterministic part equals the (global) mean value.

2 METHODOLOGY

2.1 Study Region

The Bajo Grijalva basin (fig. 1) is located southeast of Mexico and covers an area of 9,830 km². It includes 12 municipalities of the State of Tabasco and 20 of the State of Chiapas. The largest river system in the country, the Grijalva-Usumacinta, converges in this area. The runoff of its rivers is the largest in the Mexican Republic, on the order of 3,700 m³/s on annual average (Conagua, 2015). The plains of this basin have recurrent flooding due to runoff generated by heavy rains, mainly conducted by the Sierra river system that does not have any flood control (Cepal and Cenapred, 2008). They are formed rivers, streams, lagoons and swamps. In the upper part of the basin is located the Peñitas dam. The plains are covered by six types of plant associations: tropical rainforest, tropical savanna, medium and low forest, low formations of the beach, mangrove forest and swamp vegetation (West, et al., 1985). The diverse vegetation existing in the plains is a result of climatic conditions and is conserved despite alterations caused by man.

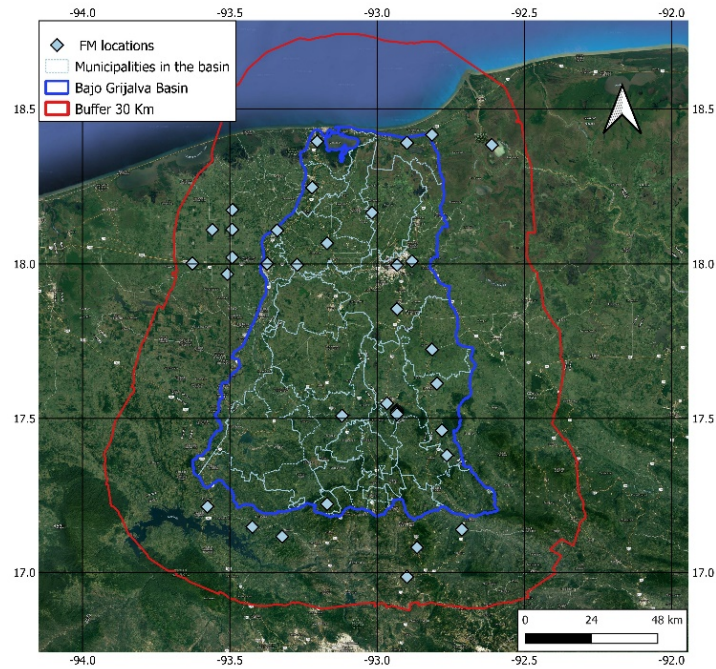


Figure 1: Bajo Grijalva Basin and its municipalities, 30 km buffer and FM location.

2.2 Data

To delimit the study area, a 30 km buffer was generated on the basin polygon (fig. 1), taking this from its definition provided by the INEGI hydrographic network (downloadable at <https://www.inegi.org.mx/temas/hidrografia/default.html#Descargas>). 2001 P data of the stations located in the buffer (fig. 1) were obtained from the database provided by Mexican Meteorological Service (SMN). These data were revised to avoid missing records and were grouped by year. On the other hand, the annual aggregates for the same year of TRMM were downloaded from <https://pmm.nasa.gov/data-access/downloads/trmm>.

2.3 Geostatistical Interpolations

The FM were interpolated by IDW, KO and KU. The values of the TRMM cell centers were also interpolated using OK to generate a product that better capture the spatial variability pattern of the parameter in comparison to the original TRMM product. RK was initially tested using TRMM with a resolution of 5 km as an auxiliary variable and the FM as the variable to be interpolated. RK was also applied to interpolate FM using as auxiliary variable the resulted layer of the OK interpolation of the values of the TRMM cell centers. In case of RK, the mean was estimated from the linear relationship with the auxiliary variable and the spatially autocorrelated error by fitting a semivariogram model using the FM. Since no linear relationship between altitude and P-values was found, this variable was not included in the RK interpolation. LOO-CV was performed to evaluate the interpolations results that shown a suitable spatial pattern.

2.4 LOO-CV

Interpolation results were evaluated using LOO-CV. The evaluation parameters were correlation coefficients (r) and z-scores (z_i) calculated as (Bivand et al. 2008):

$$z_i = \frac{z(s_i) - \hat{z}_{[i]}(s_i)}{\sigma_{[i]}(s_i)} \quad (10)$$

With $\hat{z}_{[i]}(s_i)$ the cross validation prediction for s_i and $\sigma_{[i]}(s_i)$ the corresponding kriging standard error. According to Bivand et al. (2008) it is a standardized residual, and if the variogram model is correct, the z_i should have mean and variance values close to 0 and 1.

Additionally, a comparison by principal components (PC) was made for the resulting spatial layers. All procedures were performed in the free statistical software R.

3 RESULTS AND DISCUSSION

An overview of the results of all interpolations is presented in fig.2. Boxes b and a show the results of the IDW interpolation. The effect of the values of the individual stations that generate concentric circles around them can be observed. This spatial pattern is not suitable and therefore this results were not considered correct. Boxes b through c show the results of the OK interpolation. The spatial pattern obtained is considered suitable and that is why these results were analyzed in greater detail in the following paragraph. Boxes b through d and e show the results of the FM-RK interpolation using TRMM at its original resolution (5 Km) as an external variable. It can be seen that the spatial pattern of the variable is not expected and that interpolation is therefore considered incorrect. Among boxes d, g and h it can be observed that an OK interpolation of the P values assigned to the TRMM cell centers generated an acceptable spatial distribution of the parameter. Therefore, this result and its cross-validation were analyzed in subsequent paragraphs. Boxes b and f show the results of the UK prediction from FM. These results are quite similar in terms of the spatial pattern obtained in box i. Therefore, in a subsequent section, the results shown among boxes b, h and i are analyzed and validated. The former correspond to the process of integration between FM and satellite products (TRMM) using RK. In this case the external variable is the TRMM layer obtained from the OK interpolation of the TRMM cell centers values. The box l of the figure shows the PC results comparison. RK of FM using the results of the OK interpolated TRMM cell centers as an external variable is located in the middle of all other methods, showing good results balance.

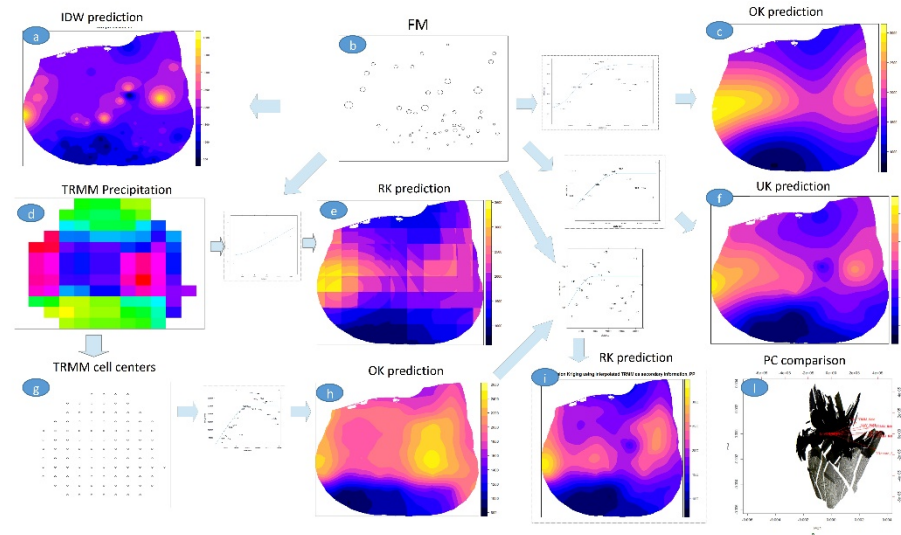


Figure 2: Overview of the results of all interpolations of annual 2001 P. At the bottom, a PC comparison of the interpolations results is shown.

3.1 OK of FM

The results of this interpolation are shown in fig. 3. The variable must have been logarithmically transformed to avoid large deviation from the nominal data to the lowest values. A gaussian semivariogram model was fitted. The result of the OK interpolation shows a smoothed spatial pattern with maximum P values in the western and eastern parts of the basin. Likewise, it is less affected by the values of the individual stations than the IDW results (shown in fig. 2). Nevertheless, since the spatial pattern of the layer shown in fig. 2 box i is more suitable in comparison to the results of this interpolation, the former were not evaluated by LOO-CV.

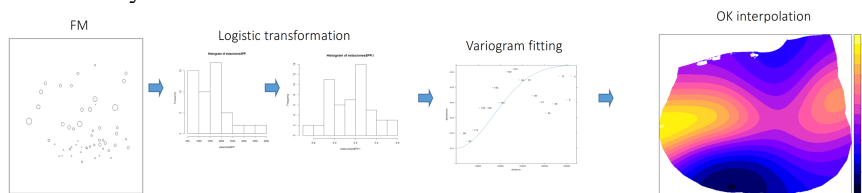


Figure 3: Results of the OK interpolation of the P values measured in stations of the Bajo Grijalva basin for 2001.

3.2 Analysis and validation of OK of cell centers from TRMM

The results of the OK interpolation of the TRMM cell center values are shown in fig. 4. In this case it was not necessary to perform a logarithmic transformation and it was possible to fit a spherical semi-variogram model. The spatial pattern obtained is very different from that of the original TRMM image, in addition to presenting the advantage of having a much higher spatial resolution (from 4 km to 300 m). The results of the validation of this interpolation are shown in the lower part of fig.4. The z_i shown in the bubble diagram are centralized to 0 and fluctuate between ± 1.6 which indicates a good fit of the

semivariogram model. r between predicted and residual values was 0.1 indicating a correct prediction (considering that the ideal is 0). Estimated variances are greater in locations where there are no measurements. r between predicted and residual values is very close to zero, which also gives validity to the model. Another result that indicates a high degree of validity is the r between the predicted and estimated values of 0.95.

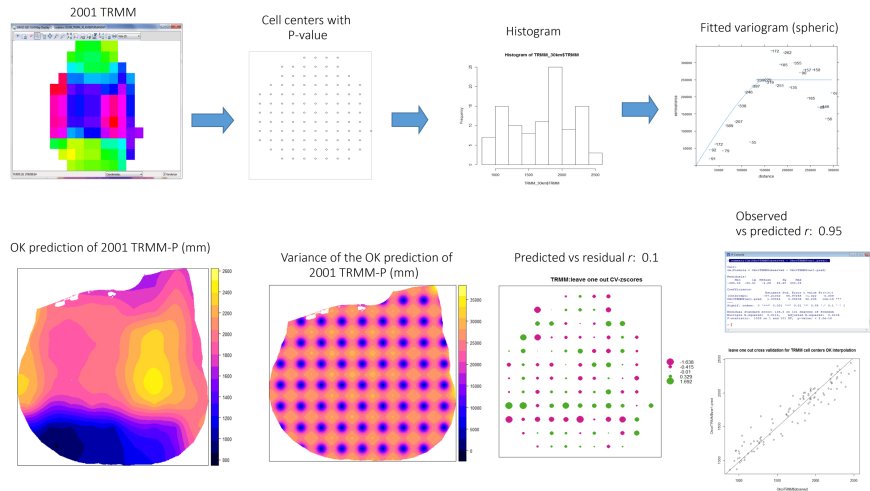


Figure 4: Results of the OK interpolation of the values of the TRMM cell centers. Results of LOO-CV are shown in the bottom.

3.3 Analysis and validation of RK interpolation of FM taking OK interpolation of TRMM cell centers as external variable

Figure 5 shows the results of the RK interpolation of the FM taking as an external variable the results of the OK interpolation of the TRMM cell centers. r between the auxiliary variable and the variable to be interpolated was 0.7 indicating the achievability of this integration scheme. This result is appropriate given that both sources of information allow quantifying the same parameter and can be taken as an indicator of the deviation between both sources of information for the analyzed case of study (approx. 30%). The spatial distribution of the variance is shown in the figure. This has a spatial pattern not determined by the locations where measurements are made as in OK (see fig. 4). The fitted semivariogram model was spherical. r between predicted and residual values was 0.1, indicating a correct prediction using this semivariogram model. The z_i are centralized to 0 and fluctuate between ± 1.69 indicating a good fit of the semivariogram model. Finally, r of the cross-validation exercise between the predicted and the measured values is 0.61, which indicates an adequate performance of the predictive model (Hengl et al. 2007).

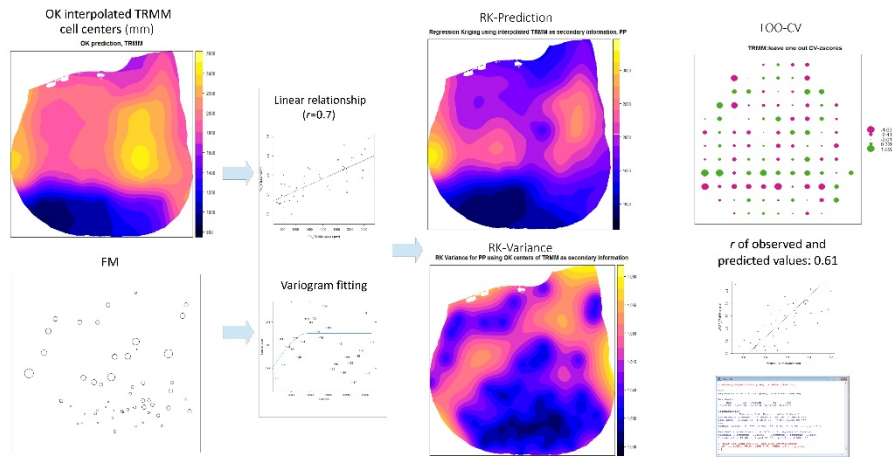


Figure 5: RK interpolation of P-FM taking as auxiliary variable the OK interpolated rainfall from TRMM cell centers. LOO-CV results are included.

3.4 General Discussion

The results obtained allow to confirm the possibility to implement a method of geostatistical integration between FM and TRMM satellite estimates. This scheme worked satisfactorily and substantially improved the information quality of the obtained P layers. Additionally, through schemes such as the one proposed, it is possible to perform downscaling processes of geographic information, given that for example in the case of OK interpolation of the TRMM cell centers, a final resolution of 300 m per pixel was obtained from one of 5 km. In this case the spatial pattern is greatly favored with statistical indicators that define low error for the generated layers. These discoveries must be confirmed by more research, but the results are promising.

As shown in Figure 2, the spatial pattern of the analyzed methods, particularly IDW, OK and RK is completely different. I argue that this pattern is a first element to select the method of obtaining the continuous spatial variable for research purposes as well to make decisions and formulate plans and programs. Among the unexpected results that were quite interesting is that an OK interpolation of TRMM cell center values allows for a substantial improvement in the spatial pattern and resolution of the resulting TRMM layer.

As Hengl et al. (2007) argue, I found that remote sensors data allow capturing spatial variability of the parameters to be interpolated and if they are integrated using geostatistical RK-type schemes with point FM that capture parameter accuracy, it is possible to improve the quality of the geographic layers of P (as done by Wang and Lin, 2015 and Lin and Wang, 2011). The integration method must be evaluated considering other climatic conditions, as well as its performance for other years, since this study only considers aggregates of rain for one year. The effects of the selected pixel size for the interpolation process, as well as the use of other auxiliary variables such as elevation and distance to the coast the use should also be investigated. In this research no linear relationship between P-values and altitude were found. However, for other conditions and geographic zones it is possible that elevation is a factor that partly defines the spatial variability of P and therefore its inclusion can improve the estimates within the proposed integration scheme. These activities are included in my future lines of research.

4 CONCLUSIONS

Geostatistical techniques allow increasing the effectiveness of the use of data from the field and satellites by improving their spatial representativity and enabling the generation of surface generation schemes (mapping) of the available information, including the use of additional information that provides value to the original information.

The main advantage of using these integration methods is that they increase the spatial representatives of the mapped variables and that is why their use should be encouraged, especially by researchers and organizations responsible for disseminating the basic information for decision-making about water resources and in general for the environment.

Acknowledgments

To CONACYT for supporting this research through the project 248719 of the 2014 National Problems Proposals Call and to Dr. Eugenio Gómez Reyes who invited me to participate and coordinated it at the UAM-Iztapalapa.

References

1. Agou, V.D., Varouchakis, E.A. & Hristopulos, D.T. *Environ Monit Assess* (191), 353 (2019). <https://doi.org/10.1007/s10661-019-7462-8>
2. Christian D. Kummerow, Sarah Ringerud, Jody Crook, David Randel, And Wesley Berg. *Journal of Atmospheric and Oceanic Technology* (28). (2011) <https://doi.org/10.1175/2010JTECHA1468.1>
3. Centro Nacional de Prevención de Desastres y Comisión Económica para América Latina y el Caribe, (Cepal y Cenapred). Tabasco: características e impacto socioeconómico de las inundaciones provocadas a finales de octubre y a comienzos de noviembre de 2007 por el frente frío número 4. Informe LC/MEX/L.864. Centro Nacional de Prevención de Desastres y Comisión Económica para América Latina y el Caribe, junio de 2008. <https://repositorio.cepal.org/handle/11362/25881>. Last accessed 01/09/2019.
4. Comisión Nacional del Agua, (Conagua). (2015). Estadísticas del Agua en México (ed. 2015). México: <https://agua.org.mx/biblioteca/estadisticas-del-agua-en-mexico-edicion-2015/>. Last accessed 01/09/2019.
5. Cressie, N., The origins of kriging. *Math. Geol.* (22), 239–252. <https://doi.org/10.1007/BF00889887>
6. Nerini, D., Zulkafli Z., Wang L.P., Onof C., Buytaert W., Lavadocasimiro W., et al. A comparative analysis of TRMM-rain gauge data merging techniques at the daily time scale for distributed rainfall-runoff modeling applications. *J. Hydrometeorol.*(16) pp.2153-2168. <https://doi.org/10.1175/JHM-D-14-0197.1>
7. Davis, J.C., *Statistics and Data Analysis in Geology*. John Wiley & Sons Inc, New York (1973)
8. Anagnostou E.N., V. Maggioni V., Nikolopoulos E.I., Meskele T., Hossain F., Papadopoulos A. Benchmarking high-resolution global satellite rainfall products to radar and rain-gauge rainfall estimates. *IEEE Trans. Geosci. Remote Sens* (48), pp.1667-1683,(2010). DOI: 10.1109/TGRS.2009.2034736
9. Holawe F., Dutter R. Geostatistical study of P series in Austria: time and space, *Journal of Hydrology* (219),1–2, pp.70-82. (1999). [https://doi.org/10.1016/S0022-1694\(99\)00046-3](https://doi.org/10.1016/S0022-1694(99)00046-3).
10. Skofronick-Jackson G., Kirschbaum D., Petersen W., Huffman G., Kidd C., Stocker E., Kakar R.. The Global P Measurement (GPM) mission’s scientific achievements and societal contributions: reviewing four years of advanced rain and snow observations. *Q J R Meteorol Soc.*(2018),144(Suppl. 1) pp. 27–48. ADVANCES IN REMOTE SENSING OF RAINFALL AND SNOWFALL. DOI: 10.1002/qj.3313. (2018).
11. Goovaerts, P., *Geostatistics for Natural Resource Evaluation*. Oxford University Press, New York. (1997).
12. Greene J. S., Morrysey M.L. Validation and uncertainty Analysis of Satellite Rainfall Algorithms. *Professional Geographer* (52)2 pp. 247-258. (2000).

13. Hengl, T., A practical guide to geostatistical mapping of environmental variables. European Communities, 2007. Publications Office, Luxembourg. (2007). <https://www.lu.lv/materiali/biblioteka/es/pilnieteksti/vid/A%20Practical%20Guide%20to%20Geostatistical%20Mapping%20of%20Environmental%20Variables.pdf>
14. Lin A. and Wang, X. L. An algorithm for blending multiple satellite precipitation estimates with in situ precipitation measurements in Canada. *Journal of Geophysical Research* (116)D21111. (2011). doi:10.1029/2011JD016359, 2011
15. Bowman K.P. Comparison of TRMM P Retrievals with Rain Gauge Data from Ocean Buoys. *Journal of Climate* (18)1, pp.178–190. (2005). <https://doi.org/10.1175/JCLI3259.1>
16. Sivasubramaniam K., Sharma A., Alfredsen K. Merging radar and gauge information within a dynamical model combination framework for P estimation in cold climates, *Environmental Modelling & Software* (119) pp. 99-110 (2019) <https://doi.org/10.1016/j.envsoft.2019.05.013>.
17. Kumar, V. Optimal contour mapping of groundwater levels using universal kriging—a case study. *Hydrol. Sci. J.* (52) pp. 1038–1050. (2007). <https://doi.org/10.1623/hysj.52.5.1038>
18. Kummerow, C., and Coauthors. The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.*(39) pp.1965–1982. (2000). [https://doi.org/10.1175/1520-0450\(2001\)040<1965:TSOTTR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2)
19. Laurent H., Jobard I., Toma A. Validation of satellite and ground based estimates of P over the Sahel. *Atmospheric Research* 47-48 pp. 651-670. (1998). [https://doi.org/10.1016/S0169-8095\(98\)00051-9](https://doi.org/10.1016/S0169-8095(98)00051-9)
20. New M., Todd M., Hulme M., Jones P. P. Measurements and Trends in the Twentieth Century. *Int. J. Climatol.*(21), pp. 1899–1922 (2001), DOI: 10.1002/joc.680
21. Matheron, G., 1969. Le krigeage universel. École nationale supérieure des mines de Paris, France.
22. Kebblouti M., Ouerdachi L., Boutaghane H., Spatial Interpolation of Annual P in Annaba-Algeria - Comparison and Evaluation of Methods, *Energy Procedia* (18) pp.468-475, (2012), <https://doi.org/10.1016/j.egypro.2012.05.058>.
23. Goovaerts, P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *Journal of Hydrology* (228)1–2, pp.113-129, (2000), [https://doi.org/10.1016/S0022-1694\(00\)00144-X](https://doi.org/10.1016/S0022-1694(00)00144-X).
24. Pielke, R. A., Stohlgren, T., Parton, W., Doesken, N., Money, J., Schell, L. Spatial representativeness of temperature Measurements from a single Site. *Bulletin of the American Meteorological Society* (81)4, pp.826-830. (2000). <https://www.jstor.org/stable/26215143>
25. Bivand R.S., Pebesma E.J., Gómez-Rubio V. *Applied Spatial Data Analysis with R*- Springer New York. (2013). DOI 10.1007/978-0-387-78171-6
26. Bell T. L. Comparing satellite rainfall estimates with rain gauge data: Optimal strategies suggested by a spectral model. *Journal of Geophysical Research* (108)D3-4121, (2003). doi:10.1029/2002JD002641
27. Hengl, T., Gerard B., Heuvelink M., Rossiter D.G. About regression-kriging: From equations to case studies, *Computers & Geosciences* (33)10, pp.1301-1315. (2007) <https://doi.org/10.1016/j.cageo.2007.05.001>.
28. West, R.C., Psuty, N.P. y Thom, B.G. (1985). Las tierras bajas de Tabasco en el sureste de México, México. México: Gobierno del Estado de Tabasco.
29. WMO (World Meteorological Organization). *Guide to Meteorological Instruments and Methods of Observation*. Geneva: WMO-No. 8. Seventh Edition. (2008). <https://www.weather.gov/media/epz/mesonet/CWOP-WMO8.pdf>
30. Zulkafli Z., Buytaert W., Onof C., Manz B., Tarnavsky E., Lavado W., et al. A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean-Amazon River Basins. *J. Hydrometeorol.*, (15) pp. 581-592. (2014). <https://doi.org/10.1175/JHM-D-13-094.1>
31. Wu Z., Zhang Y., Sun Z., Lin Q., He H. Improvement of a combination of TMPA (or IMERG) and ground-based P and application to a typical region of the East China Plain, *Science of The Total Environment* (640–641) pp. 1165-1175, (2018). <https://doi.org/10.1016/j.scitotenv.2018.05.272>.