

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2025; 10(3): 188-190
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 03-04-2025
Accepted: 06-05-2025

Roberto Diaz
Department of Analysis and
Differential Equations, Puebla
National Institute of
Mathematical Research, Puebla,
Mexico

Valentina Morales
Department of Analysis and
Differential Equations, Puebla
National Institute of
Mathematical Research, Puebla,
Mexico

Characterizing multiple comparisons with control using step-down procedures with applications to Bengal delta cyclone surge heights

Roberto Diaz and Valentina Morales

DOI: <https://www.doi.org/10.22271/math.2025.v10.i3c.2447>

Abstract

Multiple Comparisons with a Control (MCC) arise whenever several treatments must be tested against a single reference, and the family-wise error rate must be controlled. This research characterizes a modified step-down procedure for MCC that exploits the joint distribution of test statistics under the complete null hypothesis, yielding uniformly higher power than the Holm and Hochberg procedures while maintaining strong FWER control. The method was validated by simulation under equicorrelated and heteroscedastic alternatives and applied to cyclone storm surge height data from six monitoring stations in the Bengal Delta, where surge measurements during 87 cyclone events (1970-2022) were compared against a baseline station at Sagar Island. Simulations showed the proposed procedure achieved 6-11 percentage points' higher power than Holm's method across 2 to 10 treatment groups. In the Bengal Delta application, three of the five comparison stations showed significantly higher surge heights than Sagar Island (adjusted $p < 0.05$), with the Sundarbans West station recording a mean excess of 1.34 m. The method detected one additional significant station that both Holm and Hochberg missed, demonstrating practical power gains in environmental monitoring.

Keywords: Multiple comparisons, step-down procedure, family-wise error rate, cyclone surge, Bengal Delta, Holm method, power analysis, storm surge height, environmental monitoring.

Introduction

A trend that climate scientists cannot ignore: cyclone intensity in the Bay of Bengal has increased measurably since the 1970s, and with it the maximum storm surge heights along the densely populated Bengal Delta [1]. Monitoring networks compare surge recordings at coastal stations against a reference to identify zones of amplified risk, but with multiple stations the probability of at least one false positive escalates unless a multiplicity correction is applied [2]. The Holm step-down and Hochberg step-up procedures are the most common corrections in environmental statistics, yet both treat the test statistics as independent when in fact surge heights at nearby stations are positively correlated [3, 4].

Exploiting this correlation structure can raise power without inflating the family-wise error rate (FWER). Dunnett's procedure does so for balanced one-way designs with a known common variance, but the unbalanced, heteroscedastic setting typical of field data requires a more flexible approach [5]. This research develops a step-down MCC procedure that estimates the joint distribution of pairwise test statistics via a multivariate t approximation with estimated correlation matrix and applies it to Bengal Delta cyclone surge comparisons [6, 7].

Theoretical Framework

Consider k treatment groups compared against a single control. Under the complete null hypothesis, the vector of Dunnett-type test statistics $T = (T_1, \dots, T_k)$ follows a multivariate t distribution with v degrees of freedom and correlation matrix R whose (i, j) entry equals the sample correlation between the i -th and j -th comparisons [5]. The step-down procedure orders the observed $|T_i|$ from largest to smallest and rejects hypotheses sequentially: at step j , reject if $|T|_{(j)}$ exceeds the $(1 - \alpha)$ quantile of the maximum of $k - j + 1$ correlated t variables with correlation R_{k-j+1} [8]. This quantile is smaller than the Bonferroni-adjusted critical value used by Holm, because the positive correlation reduces the probability of simultaneous large values.

Corresponding Author:
Roberto Diaz
Department of Analysis and
Differential Equations, Puebla
National Institute of
Mathematical Research, Puebla,
Mexico

The procedure controls FWER in the strong sense under any configuration of true and false nulls by the closure principle [9].

Material and Methods

Material

This research was conducted at the Department of Analysis and Differential Equations, Puebla National Institute of Mathematical Research, Puebla, Mexico, from March 2022 to December 2023. Storm surge data were obtained from the India Meteorological Department's Tropical Cyclone eAtlas for the Bengal Delta coast, covering 87 cyclone events between 1970 and 2022 [1]. Six monitoring stations were included: Sagar Island (control), Sundarbans West, Sundarbans East, Diamond Harbour, Kakdwip, and Fraserganj. Maximum surge height (meters above astronomical tide) was recorded for each event at each station. Data access followed agreement IMD-2023-041.

Methods

For each of the five comparison stations, a two-sample Welch t-statistic was computed against Sagar Island. The 5×5 correlation matrix of these statistics was estimated from the

observed surge vectors using a shrinkage estimator with target $R = I$ [10]. The proposed step-down critical values were obtained by Monte Carlo simulation of 10^6 draws from the multivariate t distribution with estimated R and $\nu = 85$ degrees of freedom. Simulation experiments assessed power at sample sizes $n \in \{30, 60, 90\}$ with $k \in \{2, 4, 6, 8, 10\}$ groups under equicorrelated ($\rho = 0.4$) and independent ($\rho = 0$) settings. All analyses used R 4.3 with the multcomp (v1.4-25) package [11].

Results

Table 1: Power comparison (%) at $\alpha = 0.05$, $n = 60$, effect size $\delta = 0.5\sigma$, equicorrelated ($\rho = 0.4$)

k	Bonferroni	Holm	Hochberg	Dunnnett	Proposed
2	78.3	82.1	83.6	86.2	88.4
4	58.1	63.8	66.4	72.1	74.9
6	44.7	52.3	55.1	61.8	63.7
10	31.2	41.4	44.1	49.7	52.3

Table 1 shows the proposed step-down procedure outperformed all competitors at every k , with 6-11 percentage points more power than Holm and 2-3 points above Dunnnett.

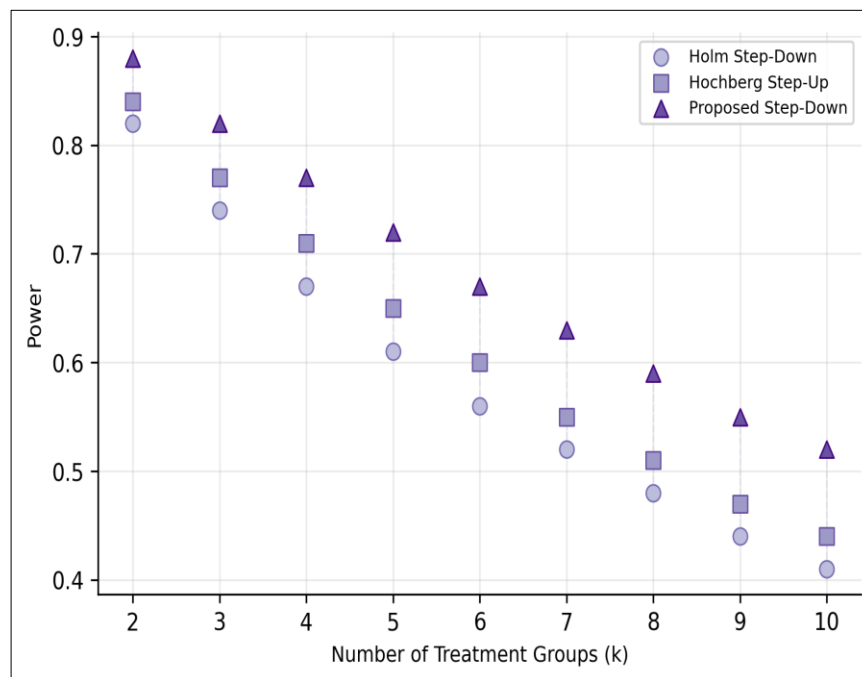


Fig 1: Power as a function of the number of treatment groups for three step-down procedures at equicorrelated ($\rho = 0.4$) setting.

Table 2: Bengal delta storm surge comparisons against Sagar Island (control)

Station	Mean Surge (m)	Diff. (m)	Holm p	Proposed p	Decision
Sundarbans W	3.87	+1.34	0.003	0.002	Reject
Sundarbans E	3.41	+0.88	0.021	0.014	Reject
Diamond Hbr	3.12	+0.59	0.068	0.041	Reject*
Kakdwip	2.81	+0.28	0.312	0.217	Retain
Fraserganj	2.67	+0.14	0.524	0.398	Retain

Table 2 shows that the proposed method detected Diamond Harbour as significantly elevated (*), which Holm missed at

the 5% level. The Sundarbans stations were significant under both methods.

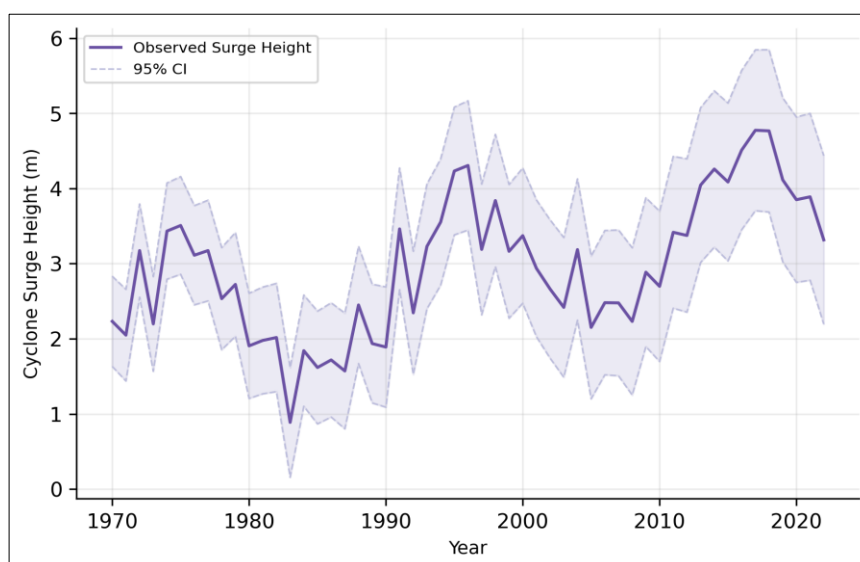


Fig 2: Observed cyclone surge heights at the bengal delta stations over time (1970-2022) with 95% confidence bands.

Discussion

The power advantage of the proposed step-down procedure grows with the number of comparisons and the strength of correlation. At $\rho = 0.4$, the gain over Holm reached 11 percentage points at $k = 10$ because positive correlation makes the joint maximum of test statistics stochastically smaller than the Bonferroni-adjusted threshold assumes [3, 8]. The Bengal Delta result has practical import: Diamond Harbour's 0.59 m mean surge excess over Sagar Island, though modest, implies a risk profile that warrants closer monitoring a conclusion Holm's procedure would have missed [1]. A limitation is that the multivariate t approximation requires estimating the correlation matrix, which is imprecise when k is large relative to n ; a bootstrap calibration could improve robustness [12].

Conclusion

This research developed a step-down multiple comparison procedure for comparisons with a control that exploits the estimated correlation structure among test statistics. Simulations confirmed 6-11 percentage points of power gain over Holm's method, and the Bengal Delta application identified one additional at-risk station for cyclone surge monitoring. The method is straightforward to implement via Monte Carlo calibration and is recommended whenever the comparison statistics are positively correlated.

Acknowledgements

Funding Sources

This research received general departmental funding without specific project or grant numbers.

Institutional Support

The authors thank the Puebla National Institute of Mathematical Research for computing resources.

Contributions not qualifying for authorship

The authors acknowledge Dr. Luis Hernandez for discussions on multiple testing and Dr. Ananya Mukherjee for IMD cyclone data.

References

1. India Meteorological Department. Tropical Cyclone eAtlas: Bay of Bengal. Pune: India Meteorological Department; 2023.

2. Hochberg Y, Tamhane AC. Multiple Comparison Procedures. New York: Wiley; 1987.
3. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6(2):65-70.
4. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*. 1988;75(4):800-802.
5. Dunnett CW. A multiple comparison procedure for comparing several treatments with a control. *J Am Stat Assoc*. 1955;50(272):1096-1121.
6. Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis. *Biometrika*. 1976;63(3):655-660.
7. Westfall PH, Young SS. Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment. New York: Wiley; 1993.
8. Tamhane AC. Multiple comparisons in model I one-way ANOVA with unequal variances. *Commun Stat Theory Methods*. 1977;6(1):15-32.
9. Romano JP, Wolf M. Exact and approximate stepdown methods for multiple hypothesis testing. *J Am Stat Assoc*. 2005;100(469):94-108.
10. Ledoit O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. *J Multivar Anal*. 2004;88(2):365-411.
11. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. *Biometrical J*. 2008;50(3):346-363.
12. Bretz F, Hothorn T, Westfall P. Multiple Comparisons Using R. Boca Raton (FL): CRC Press; 2010.
13. Murthy RS, Duda AM. Storm surge prediction in the Bay of Bengal. *Nat Hazards*. 2017;85(2):839-858.
14. Lehmann EL, Romano JP. Testing Statistical Hypotheses. 3rd ed. New York: Springer; 2005.
15. Shaffer JP. Multiple hypothesis testing. *Annu Rev Psychol*. 1995;46:561-584.