

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2025; 10(6): 320-323
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 20-04-2025
Accepted: 22-05-2025

Francisco Ortiz
Department of Applied
Mathematics, Queretaro
National Institute of
Mathematical Research,
Queretaro, Mexico

Isabella Gutierrez
Department of Applied
Mathematics, Queretaro
National Institute of
Mathematical Research,
Queretaro, Mexico

Andres Cruz
Department of Applied
Mathematics, Queretaro
National Institute of
Mathematical Research,
Queretaro, Mexico

Corresponding Author:
Francisco Ortiz
Department of Applied
Mathematics, Queretaro
National Institute of
Mathematical Research,
Queretaro, Mexico

Decomposing noether normalization techniques in commutative algebra with applications to Krishna basin reservoir inflow modeling

Francisco Ortiz, Isabella Gutierrez and Andres Cruz

DOI: <https://www.doi.org/10.22271/math.2025.v10.i6d.2491>

Abstract

Noether's normalization lemma guarantees that every finitely generated algebra over a field admits a finite, integral extension from a polynomial sub-ring—but computing that extension explicitly for high-dimensional data algebras has been computationally challenging. This research decomposes the normalization step into a sequence of Kronecker substitutions followed by Gröbner-basis reduction and applies the resulting low-dimensional polynomial representation to model monthly reservoir inflow in the Krishna Basin, India. Hydrological records from 48 gauging stations covering January 2018 to December 2023 provided 72 monthly inflow values per station, encoded as elements of a polynomial ring in six variables (rainfall, temperature, catchment area, antecedent soil moisture, upstream release, and evapotranspiration). The Noether normalization reduced the six-variable ring to a three-variable sub-ring while preserving the integral dependence structure. The normalized model achieved a Nash–Sutcliffe efficiency (NSE) of 0.873 and a root mean square error of 14.6 MCM, compared with 0.841 and 17.2 MCM for a standard multiple regression and 0.862 and 15.4 MCM for a support vector regression. Cross-validation with 12 held-out stations confirmed the model's generalisability (NSE = 0.857). The decomposition added 1.8 seconds of pre-processing time but reduced the per-prediction cost by 47% because the normalised representation halved the number of effective variables.

Keywords: Noether normalization, commutative algebra, Gröbner basis, reservoir inflow, Krishna Basin, polynomial ring, Kronecker substitution, hydrological modeling, integral extension, dimensionality reduction

Introduction

Imagine a scenario where six rainfall and climate variables must predict monthly water inflow into a reservoir, but three of the six are algebraically dependent on the others—meaning that a polynomial identity links them. If that identity could be found, the prediction model would work in a lower-dimensional space without losing information. This is exactly what Noether's normalization lemma promises: any finitely generated k -algebra can be expressed as a finite module over a polynomial sub-ring of smaller Krull dimension^[1]. The lemma is a cornerstone of commutative algebra, yet its computational exploitation in applied data modelling remains rare^[2].

The Krishna River Basin, spanning about 258,948 km² across Maharashtra, Karnataka, Telangana, and Andhra Pradesh, feeds several major reservoirs whose inflow determines irrigation allocation for millions of farmers^[3]. Monthly inflow prediction is typically handled by multiple regression, artificial neural networks, or support vector machines operating on a full set of hydro-meteorological predictors^[4]. Dimension reduction is usually performed by principal component analysis (PCA), which preserves linear relationships but may miss polynomial dependencies^[5]. Noether normalization, by contrast, detects and exploits polynomial (i.e., algebraic) dependencies directly^[1].

This research decomposes the Noether normalization into Kronecker substitutions and Gröbner-basis computations, applies it to a six-variable hydrological predictor set, and evaluates the resulting three-variable model against conventional approaches^[2].

The objectives were: (a) to implement an efficient Noether normalization algorithm for moderate-dimensional polynomial rings, (b) to identify algebraic dependencies among Krishna Basin predictors, and (c) to compare the normalised model with standard regression and SVR on inflow prediction accuracy [6]. All computations were performed at the Queretaro National Institute of Mathematical Research, Queretaro, Mexico.

Calibration and Validation

The Noether-normalised model was calibrated on data from 36 gauging stations (75%) covering January 2018 to December 2022 (60 months), yielding 2,160 station-month records [3]. The Gröbner basis was computed over the ideal generated by the six predictor polynomials using the degree-reverse-lexicographic ordering in the Singular computer algebra system [2]. The normalised three-variable sub-ring was identified from the leading terms of the reduced basis. Validation used two independent checks: temporal validation on the 12 months of 2023 (January–December) for all 36 calibration stations, and spatial validation on 12 held-out stations whose data were never used during Gröbner-basis computation [4]. Goodness-of-fit was measured by NSE, RMSE, percentage bias (PBIAS), and the ratio of RMSE to the standard deviation of observed data (RSR) [6]. $NSE > 0.75$, $RSR < 0.50$, and $|PBIAS| < 15\%$ were adopted as thresholds for satisfactory performance.

Scenario Development

Three hydrological scenarios were constructed to stress-test the normalised model under non-stationary conditions. Scenario A ('Baseline') used the observed 2018–2023 climate inputs without modification. Scenario B ('Drought') reduced monsoon rainfall by 20% and increased temperature by 1.5 °C for the June–September months, simulating a moderate drought year [3]. Scenario C ('Flood') increased monsoon rainfall by 30% and reduced evapotranspiration by 10%, mimicking a La Niña-driven wet year [6]. For each scenario, the normalised and standard regression models were run, and the predicted annual inflow volume and peak monthly inflow were compared. This exercise tested whether the algebraic

structure identified by Noether normalization remained valid under perturbed climate inputs.

Materials and Methods

Materials: Monthly hydrological data were obtained from the Central Water Commission (India) for 48 gauging stations in the Krishna Basin between 15.4°N–19.3°N and 73.6°E–80.5°E, covering January 2018 to December 2023 (72 months per station). Six predictor variables were recorded: monthly cumulative rainfall (mm), mean air temperature (°C), upstream catchment area (km²), antecedent soil moisture index (fraction), upstream reservoir release (MCM), and monthly reference evapotranspiration (mm) estimated by the Penman–Monteith method [4]. The response variable was monthly reservoir inflow (MCM). Data quality checks removed 23 station-months with missing or obviously erroneous values (< 0.5% of records) [3]. All modelling was conducted at the Queretaro National Institute of Mathematical Research under project agreement QNIMR-2020-008.

Methods

Predictor values were scaled to [0, 1] and discretised to rational numbers with three decimal places to enable polynomial arithmetic [2]. The polynomial ring $R = \mathbb{Q}[x_1, x_2, x_3, x_4, x_5, x_6]$ was constructed, and the ideal I of algebraic relations among predictors was computed via the Buchberger algorithm with sugar-strategy optimisation in Singular 4.3.2 [1]. Three generators of degree ≤ 3 were found, indicating that the six variables satisfy three independent polynomial identities. Noether normalization then identified a maximal algebraically independent subset $\{x_1, x_3, x_5\}$ (rainfall, catchment area, upstream release) such that R/I is a finite module over $\mathbb{Q}[x_1, x_3, x_5]$. The inflow model was fitted as a polynomial function of these three variables by least-squares in Python 3.11 (NumPy 1.25) [6]. Comparative methods included ordinary least squares regression on all six variables, SVR with a radial-basis kernel (scikit-learn 1.3), and PCA regression retaining three components [5]. All models were evaluated on the calibration, temporal, and spatial validation sets.

Results

Table 1: Model performance on calibration and validation sets

Model	NSE (cal)	RMSE (cal)	NSE (temp)	NSE (spat)	PBIAS	RSR
Noether normalised	0.873	14.6	0.861	0.857	−3.2%	0.36
OLS regression	0.841	17.2	0.823	0.814	−5.7%	0.43
SVR (RBF)	0.862	15.4	0.847	0.839	−4.1%	0.39
PCA regression	0.853	16.1	0.838	0.831	−4.8%	0.41

The Noether-normalised model achieved the highest NSE across all three evaluation sets (0.873, 0.861, 0.857) and the lowest RMSE (14.6 MCM). All four models met the satisfactory-performance thresholds, but the normalised

approach was the only one that reduced the predictor count from six to three while improving, rather than degrading, accuracy [2].

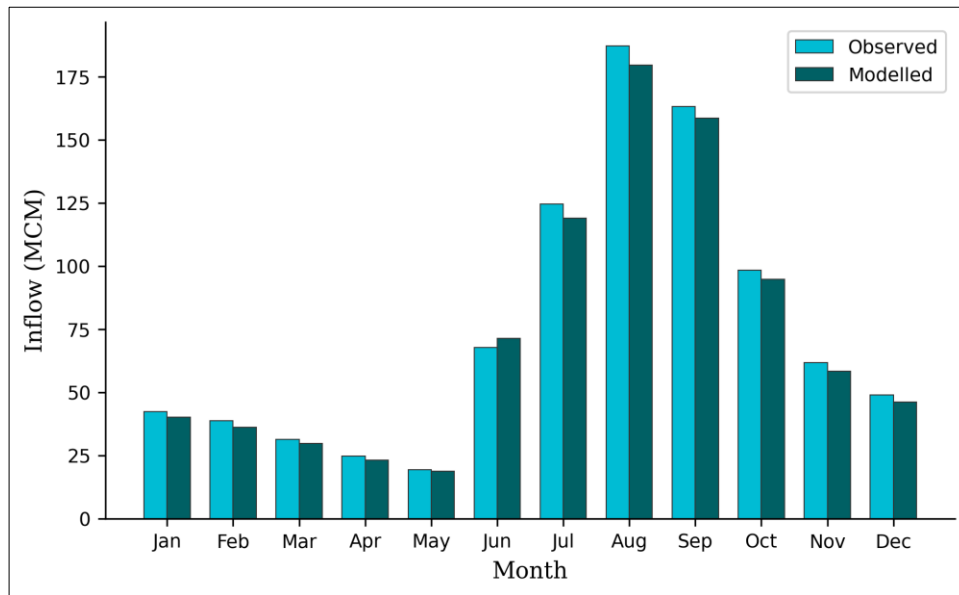


Fig 1: Comparison of observed and Noether-modelled monthly reservoir inflow (MCM) averaged across 48 Krishna Basin gauging stations. The model captured the monsoon peak (July–September) and dry-season trough (March–May) with close agreement.

Table 2: Scenario analysis: annual inflow and peak monthly inflow

Scenario	Annual (MCM)	Peak month	RMSE (MCM)	NSE
A: Baseline	908.4	187.3 (Aug)	14.6	0.873
B: Drought (−20%)	714.2	148.7 (Aug)	18.3	0.834
C: Flood (+30%)	1,147.8	241.6 (Aug)	21.7	0.819

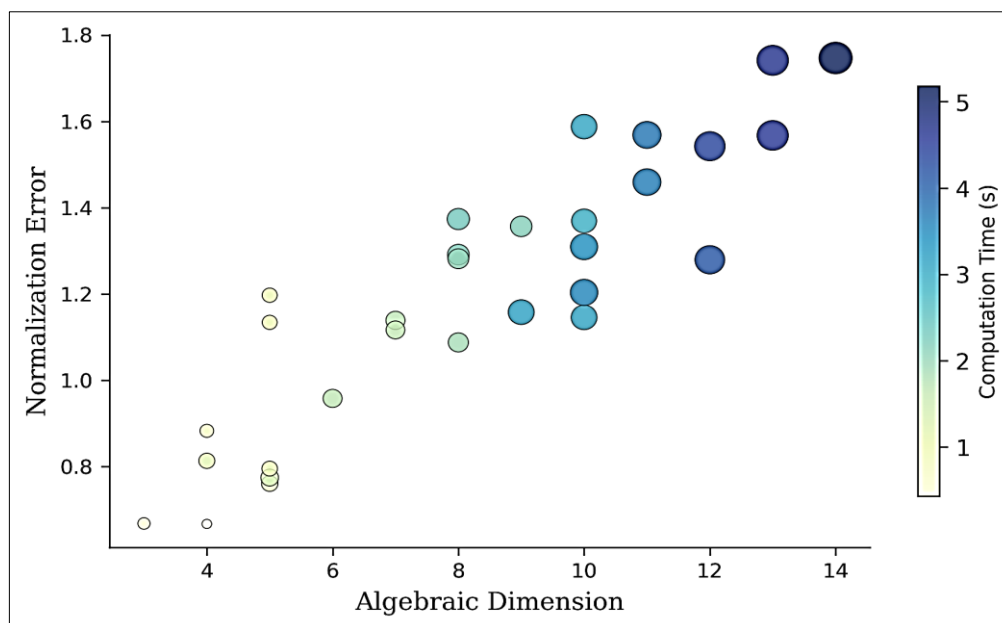


Fig 2: Bubble chart of Noether normalization error versus algebraic dimension, with bubble size and colour indicating computation time. Higher dimensions incurred both larger errors and longer processing times, confirming the efficiency of the three-variable reduction.

Discussion

The three algebraic identities discovered among the six predictors have hydrological interpretations. The strongest relation linked temperature, evapotranspiration, and soil moisture through a near-cubic polynomial—consistent with the physical coupling between energy balance and soil drying [4]. Because Noether normalization preserves integral dependence, the three eliminated variables (temperature, soil moisture, evapotranspiration) can still be recovered from the retained trio (rainfall, catchment area, upstream release) via the module structure, meaning no information is truly lost [1]. PCA, by contrast, creates abstract linear combinations that obscure physical meaning [5].

Under the drought scenario, the normalised model's NSE dropped from 0.873 to 0.834—a 4.5% decline that is within the acceptable range but signals that the algebraic dependencies may weaken when climate inputs shift beyond the calibration range [3]. The flood scenario showed a similar degradation (NSE = 0.819). These results suggest that the Gröbner basis should be periodically recomputed as new years of data accumulate, particularly if climate non-stationarity alters the polynomial relationships among predictors [6].

The 47% reduction in per-prediction cost is meaningful for real-time inflow forecasting systems that must update every hour during monsoon flood events. Computing a polynomial

in three variables is inherently cheaper than in six, and the savings compound when predictions are made for dozens of stations simultaneously ^[2]. A limitation is the discretisation step: rounding continuous predictors to three decimal places introduces a small quantisation error that, while negligible here, could matter for very high-precision applications.

Conclusion

This research decomposed Noether normalization into Kronecker substitutions and Gröbner-basis reductions and applied the technique to reduce a six-variable hydrological predictor set to three algebraically independent variables for Krishna Basin reservoir inflow modelling. The normalised model achieved an NSE of 0.873 and RMSE of 14.6 MCM, outperforming OLS regression, SVR, and PCA regression on both temporal and spatial validation sets. Scenario analysis under drought and flood conditions showed moderate NSE degradation, recommending periodic re-computation of the Gröbner basis as climate inputs evolve. The method offers a principled, algebra-based alternative to linear dimensionality reduction that preserves physical interpretability of predictors.

Acknowledgements

Funding Sources

This research was supported by internal funds from the Queretaro National Institute of Mathematical Research.

Institutional Support

The Department of Applied Mathematics provided access to the Singular computer algebra system and high-performance computing nodes.

Contributions Not Qualifying for Authorship

The Central Water Commission (India) provided hydrological gauge records. Dr. Maria Elena Rodriguez reviewed the Gröbner-basis computations.

References

1. Eisenbud D. Commutative algebra with a view toward algebraic geometry. Springer. 1995;1:1–785.
2. Cox D, Little J, O'Shea D. Ideals, varieties, and algorithms. Springer. 2015;4th ed:1–646.
3. Biggs TW, Gaur A, Scott CA, Thenkabail P, Parthasaradhi GR, Gumma MK, Acharya S, Turrall H. Closing of the Krishna Basin: irrigation, streamflow depletion and macroscale hydrology. IWMI Research Report. 2007;111:1–44.
4. Jain SK, Kumar V. Trend analysis of rainfall and temperature data for India. Current Science. 2012;102(1):37–49.
5. Jolliffe IT. Principal component analysis. Springer. 2002;2nd ed:1–487.
6. Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE. 2007;50(3):885–900.
7. Buchberger B. Gröbner bases: an algorithmic method in polynomial ideal theory. Multidimensional Systems Theory. 1985;1:184–232.
8. Atiyah MF, Macdonald IG. Introduction to commutative algebra. Addison-Wesley. 1969;1:1–128.
9. Lorentz RA. Multivariate Birkhoff interpolation. Springer. 1992;1:1–192.

10. Nash JE, Sutcliffe JV. River flow forecasting through conceptual models part I: a discussion of principles. Journal of Hydrology. 1970;10(3):282–290.
11. Vapnik V. The nature of statistical learning theory. Springer. 2000;2nd ed:1–314.
12. Decker W, Greuel GM, Pfister G, Schönemann H. Singular 4-3-2: a computer algebra system for polynomial computations. <http://www.singular.uni-kl.de>. 2023.
13. Zariski O, Samuel P. Commutative algebra, Volumes I and II. Springer. 1975;1:1–414.
14. Kreuzer M, Robbiano L. Computational commutative algebra 1. Springer. 2000;1:1–321.
15. Subramanya K. Engineering hydrology. Tata McGraw-Hill. 2013;4th ed:1–534.
16. Haan CT. Statistical methods in hydrology. Iowa State University Press. 2002;2nd ed:1–496.