# Case Report

Jack Lichtenstein, Abbey List, Jingxuan Liu, Linda Tang, Mary Wang, and Justin Zhao

2021-10-12

## 1 Introduction

Turbulence is a fundamental concept in fluid mechanics. Understanding how fluids move has enormous importance to a vast range of problems, with practical applications in aeronautical engineering, environmental science, astronomy, and more. In contrast to laminar flow, turbulent flow is irregular, unpredictable, and energy-dissipating. Turbulence also enhances mixing and leads to non-uniform distribution of particles and cluster formation, which we are particularly interested in analyzing as a part of this study.

In general, the final state of particles subject to turbulent flow is determined by three main parameters: Reynolds (Re) number, Froude (Fr) number and Stokes (St) number. The Reynolds number is a measure of the intensity of turbulence, with a higher Reynolds number corresponding to a higher intensity of turbulence (J. den Toonder et al., 1997). Broadly speaking, a Reynolds number $> 4000$ generally represents a relatively chaotic and turbulent flow, a Reynolds number $< 2300$ generally represents a smooth laminar flow and any number in between typically represents transient flow (Schlichting et al., 2017). The Froude number is a measure of the impact of gravitational acceleration on fluid motion. For instance, a cumulonimbus cloud at a high level above the ground will have a smaller Froude number (compared to lower hanging stratus clouds) because it experiences a lower intensity of gravitational acceleration relative to other clouds (Chanson, 2009). Stokes number is a description of particle properties; a large Stokes number correlates with large particle size which tends to form relatively loose clusters (Ireland et al., 2016).

Current research literature on particle clustering in turbulent flow focuses primarily on the Direct Numerical Simulation (DNS) method to solve the Navier-Stokes equation, a process which is extremely time consuming and computationally expensive (Moin & Mahesh, 1998). Another limitation to this approach is that the DNS method cannot be practically applied in situations to simulate flows with large Reynolds numbers (*Reference*). Thus, the primary research **objective** of this study is to build a statistical model using the three parameters Re (Reynolds Number), Fr (Froude Number), and St (Stokes Number) that is capable of both prediction and inference for a particle cluster volume distribution. In particular, the response variable of this study is a particle cluster volume distribution. Voronoi Tessellation, a technique that looks at general features of individual clusters in the underlying turbulence, has been applied to generate a distribution of cluster volumes. Since this probability distribution of cluster volumes is harder for statistical learning methods to work with, we will summarize this distribution by its first four raw moments $E(X)$, $E(X^2)$, $E(X^3)$, and $E(X^4)$ and partition our focus into four new response variables. Theoretically, we are interested in the insights for each of these four moments - the mean of the distribution could be a good indicator of how flow behaves on average, the variance of the distribution could dictate how flow varies over time, the skew of the distribution could illustrate asymmetric properties of the flow, and the kurtosis of the distribution could indicate how particular cluster volumes deviate. We hope that our research model will enable a quick and efficient prediction of a particle cluster volume distribution and enhance our understanding of the relative influence of these parameters in turbulent flow.

## 2 Methods

### 2.1 EDA

Before we begin any analysis, let's further split our training data into a smaller train and test set. We also create cross-validation folds with $K = 5$. We do this in an effort to reduce the likelihood of overfitting to the full training data.

While Fr and Re consists of continuous values in real life, a brief examination of our training data reveals that both two variables contain only three levels. As such, we convert both variables to factors.

Below, we plot the relationship between the three predictor variables and each of the four response variables (Figure 1 contains graphs for Moment 1. For plots relating to Moments 2-4, please refer to Figures XX-XX in the Appendix). Note, the curves fit to the points are via local polynomial regression.

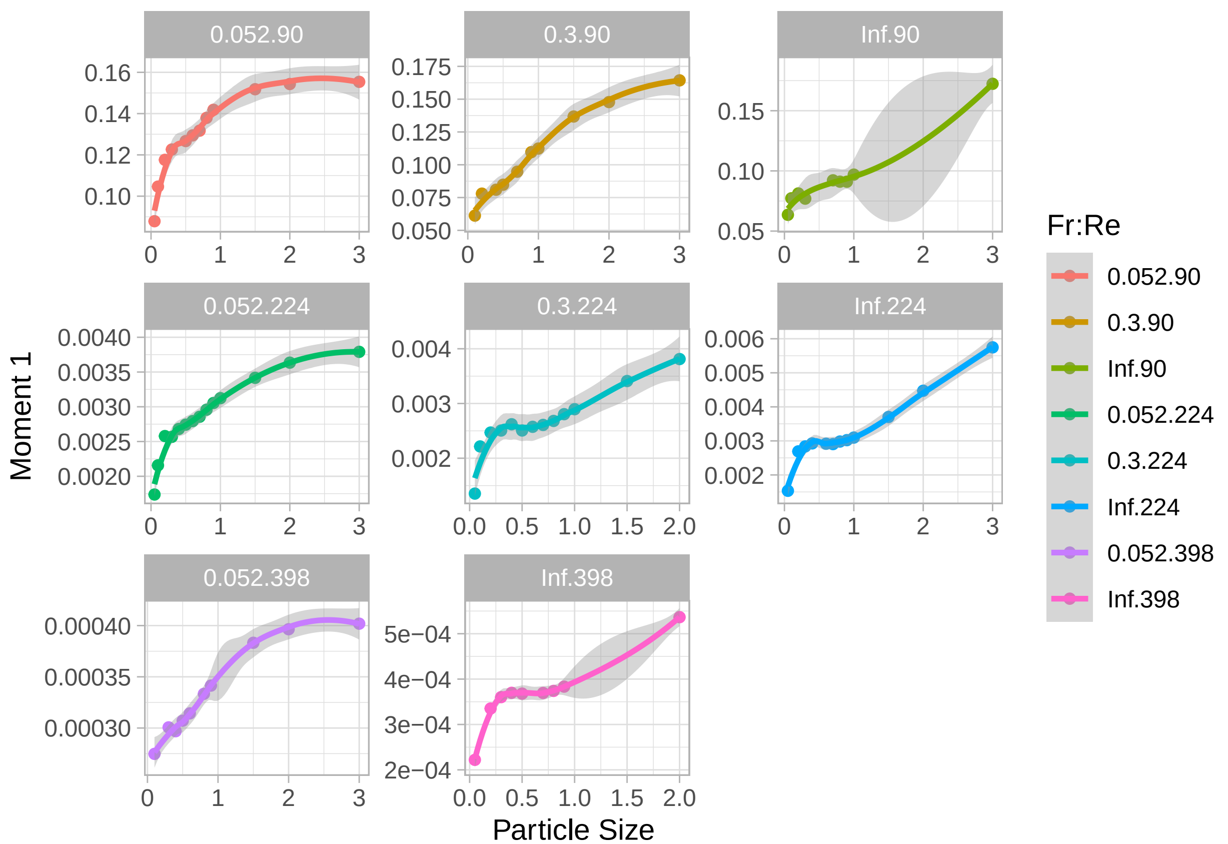
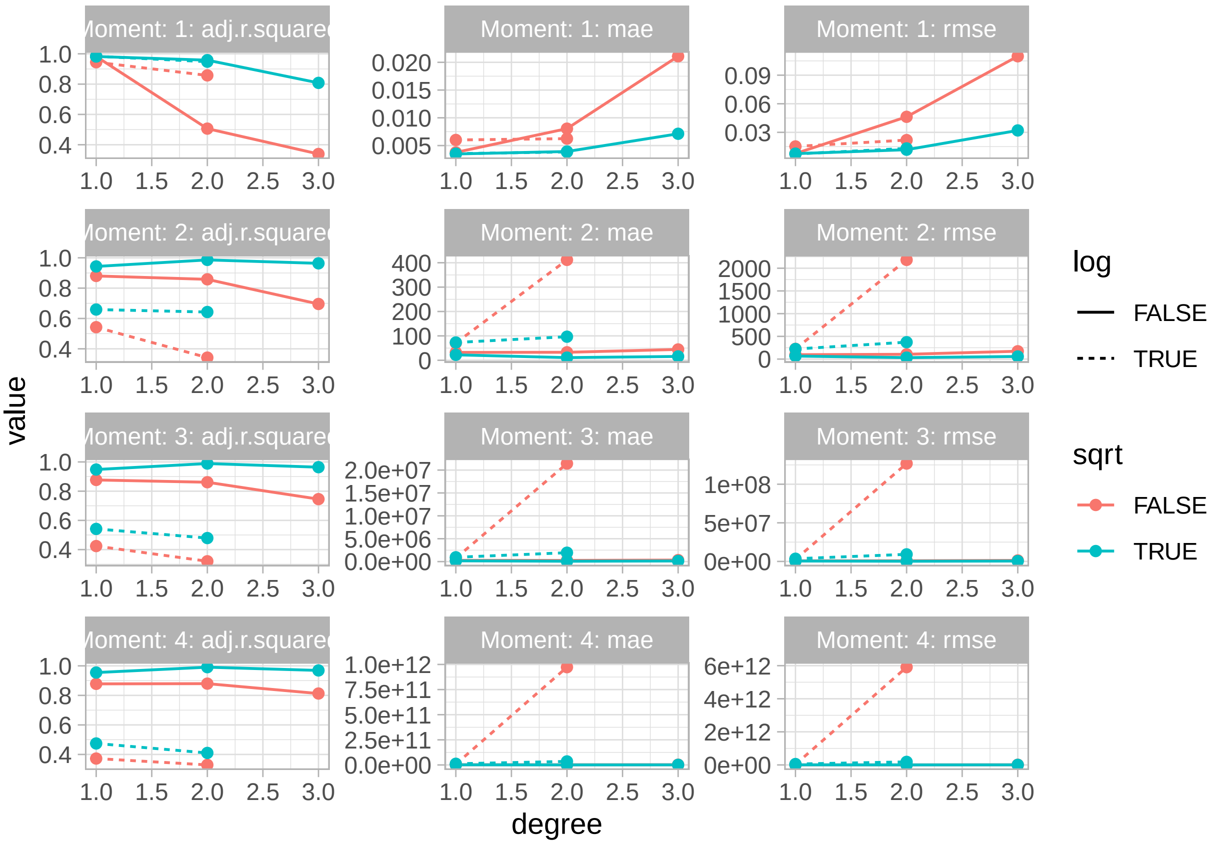



Figure 1: Moment 1 values as a function of particle size at different levels of interaction between Fr and Re.

In general, these seem to be very strong relationships. We also notice that the interaction between `Fr` and `Re` seems to explain a lot of the variance in the response. That is, the relationship between the third variable `St` and the response depends a lot on the specific interaction between `Fr` and `Re`. We will very likely need to include this interaction in any model we build for these data.

Additionally, we notice that the relationship between `St` and the response may benefit from taking the square root of `St`.

Below, we perform cross validation of a number of candidate models for each moment. Specifically, for each moment, we train a model to predict the moment with the general formula `~ poly(St, degree)*interaction` where `interaction` is the factor interaction between `Fr` and `Re`. We vary the `degree` parameter from 1 to 3. Additionally, we may choose to take the square root of `St` or take the log of the response.




| moment | degree | sqrt | log | name | value |
|---|---|---|---|---|---|
| 1 | 1 | TRUE | FALSE | mae | 0.00348 |
| 1 | 1 | TRUE | TRUE | rmse | 0.00730 |
| 1 | 1 | TRUE | TRUE | adj.r.squared | 0.98384 |
| 2 | 2 | TRUE | FALSE | rmse | 32.42433 |
| 2 | 2 | TRUE | FALSE | mae | 10.84030 |
| 2 | 2 | TRUE | FALSE | adj.r.squared | 0.98613 |
| 3 | 2 | TRUE | FALSE | rmse | 240,845.80324 |
| 3 | 2 | TRUE | FALSE | mae | 76,364.57174 |
| 3 | 2 | TRUE | FALSE | adj.r.squared | 0.98910 |
| 4 | 2 | TRUE | FALSE | rmse | 1,837,866,060.54046 |
| 4 | 2 | TRUE | FALSE | mae | 558,091,625.13870 |
| 4 | 2 | TRUE | FALSE | adj.r.squared | 0.99101 |

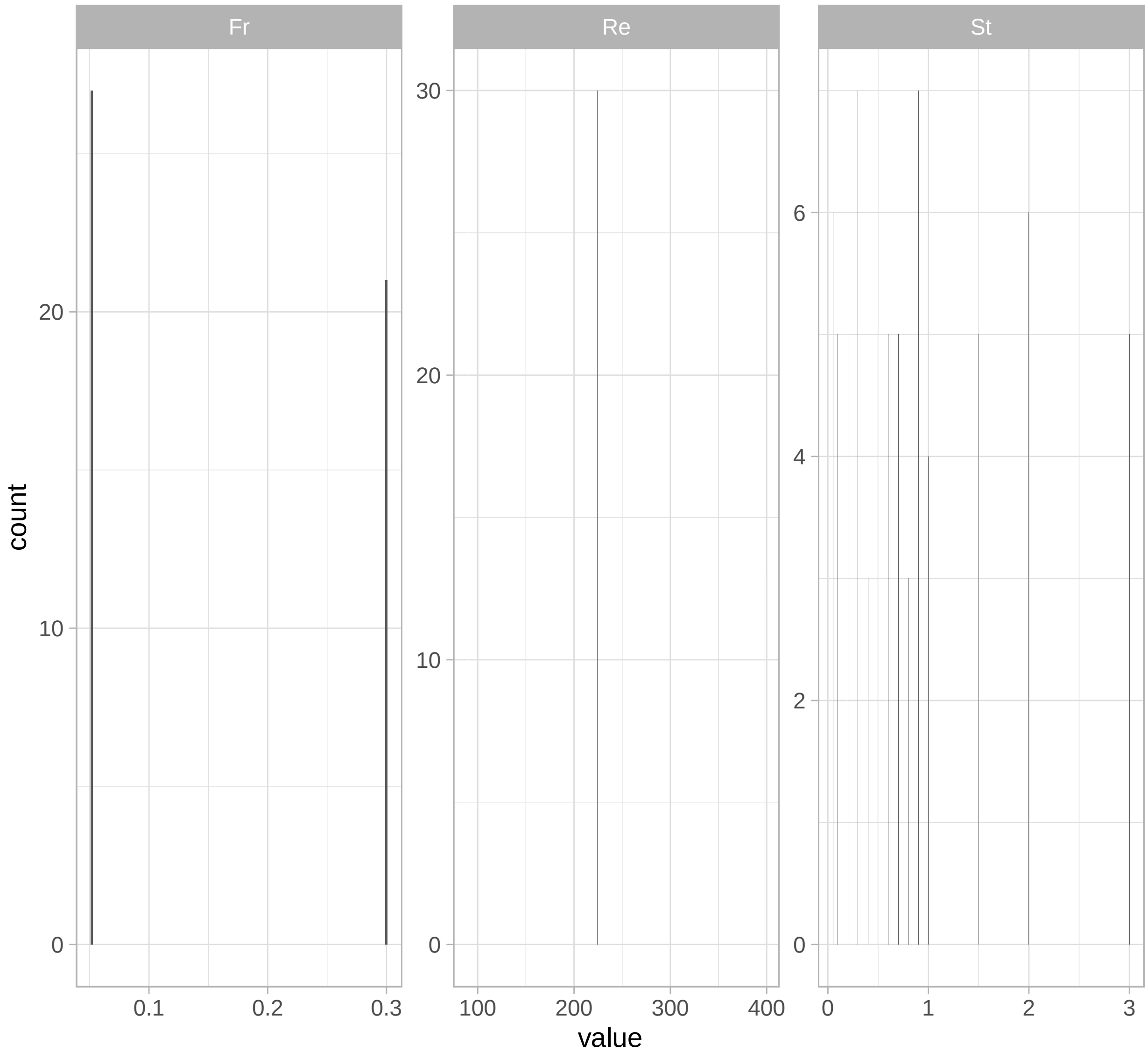
## Results

## Discussion

## References


- J. M. J. den Toonder, & Nieuwstadt, F. T. M. (1997, November 1). Reynolds number effects in a turbulent pipe flow for low to moderate re. AIP Publishing. Retrieved October 12, 2021, from https://aip.scitation.org/doi/pdf/10.1063/1.869451.
- Schlichting, H., Gersten, K., Krause, E., & Oertel, H. (2017). Boundary-layer theory. Springer.
- Chanson, Hubert (2009). "Development of the Bélanger Equation and Backwater Equation by Jean-Baptiste Bélanger (1828)" (PDF). Journal of Hydraulic Engineering. 135 (3): 159–63. doi:10.1061/(ASCE)0733-9429(2009)135:3(159).
- Ireland, P. J., Bragg, A. D., & Collins, L. R. (2016, May 11). The effect of Reynolds number on inertial particle dynamics in isotropic turbulence. part 2. simulations with gravitational effects: Journal of Fluid Mechanics. Cambridge Core. Retrieved October 12, 2021, from https://www.cambridge.org/core/journals/journal-of-fluid-mechanics/article/effect-of-reynolds-number-on-inertial-particle-dynamics-in-isotropic-turbulence-part-2-simulations-with-gravitational-effects/67C55CDC28B1B1C7868B7A402E279AF9.
- Moin, P., & Mahesh, K. (n.d.). Direct numerical simulation: A tool in turbulence research. Annual Reviews. Retrieved October 12, 2021, from https://www.annualreviews.org/doi/full/10.1146/annurev.fluid.30.1.539.

## Appendix



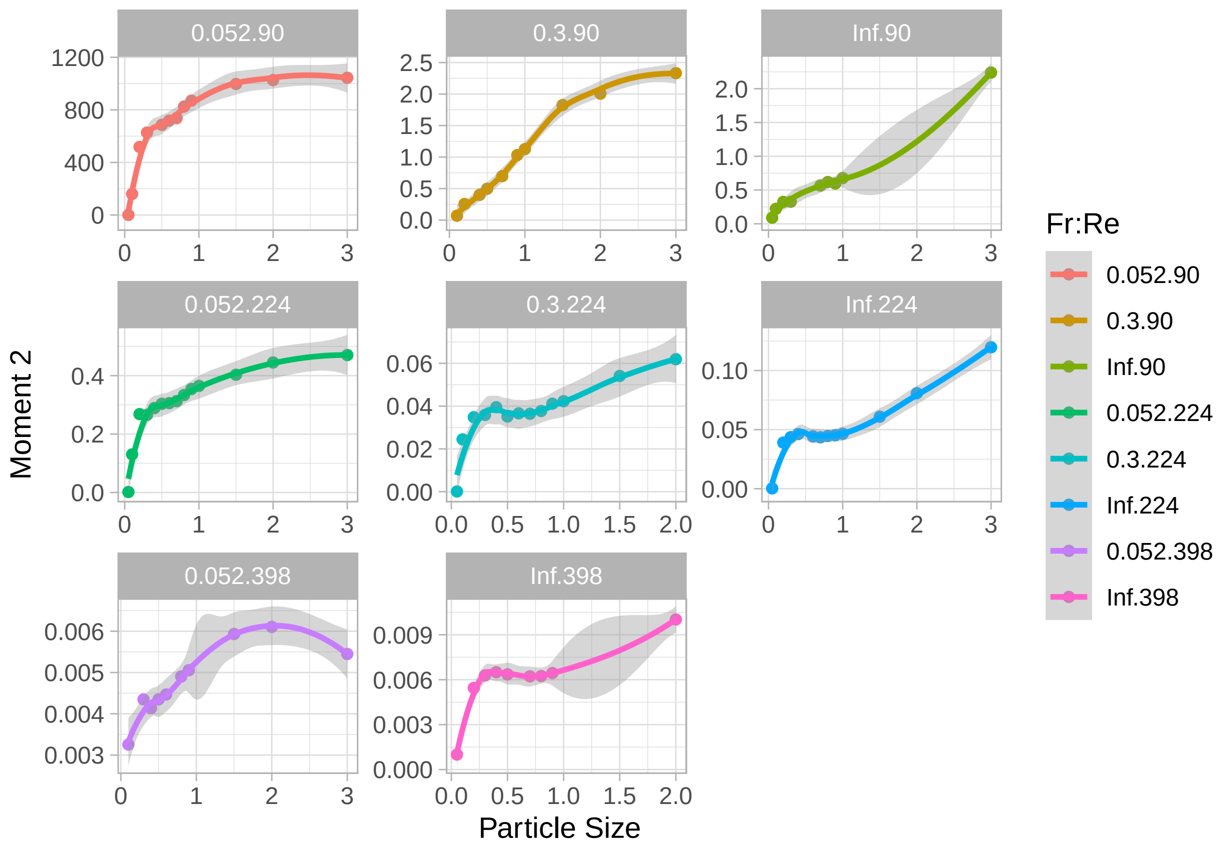



Figure 2: Moment 2 values as a function of particle size at different levels of interaction between Fr and Re.

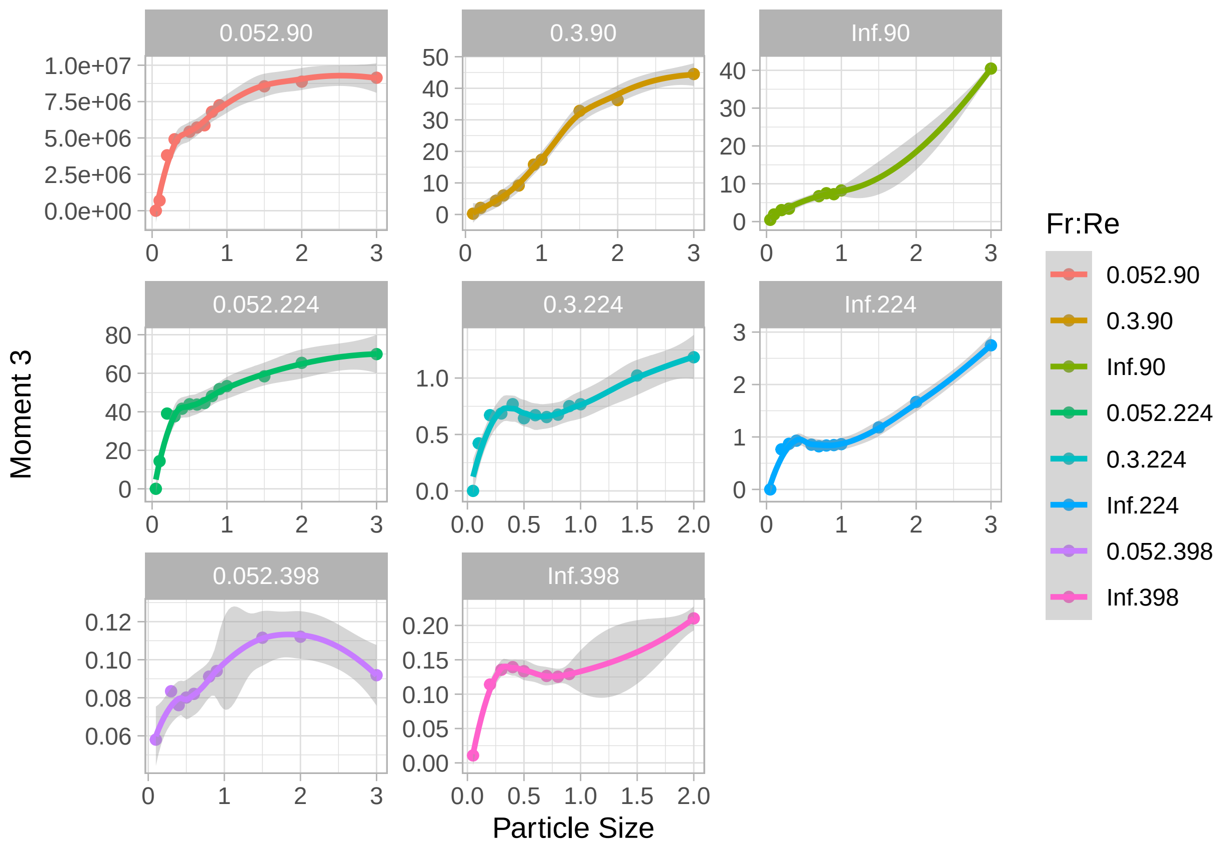



Figure 3: Moment 3 values as a function of particle size at different levels of interaction between Fr and Re.

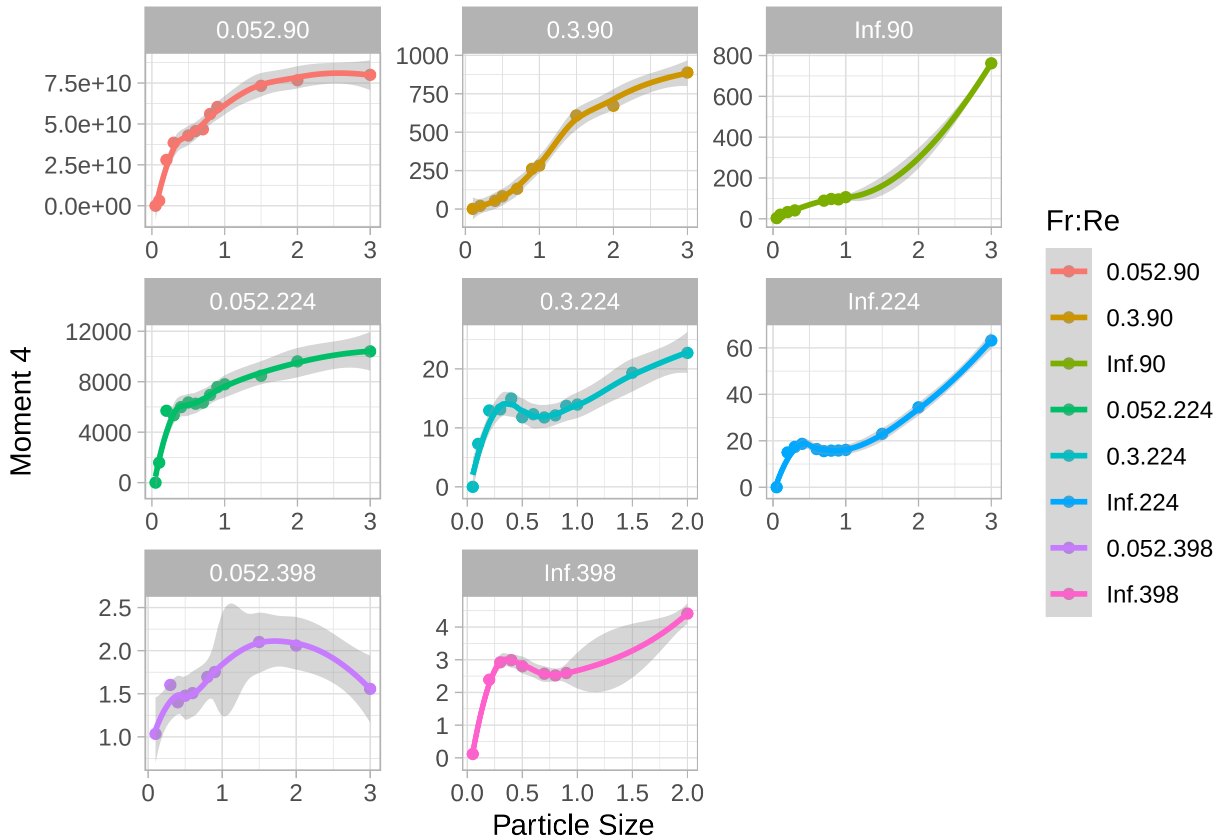



Figure 4: Moment 4 values as a function of particle size at different levels of interaction between Fr and Re.