We estimate the complexity of molecules directly from data that spacecrafts can obtain in the field.

Inferring molecular complexity from mass spectrometry data using machine learning
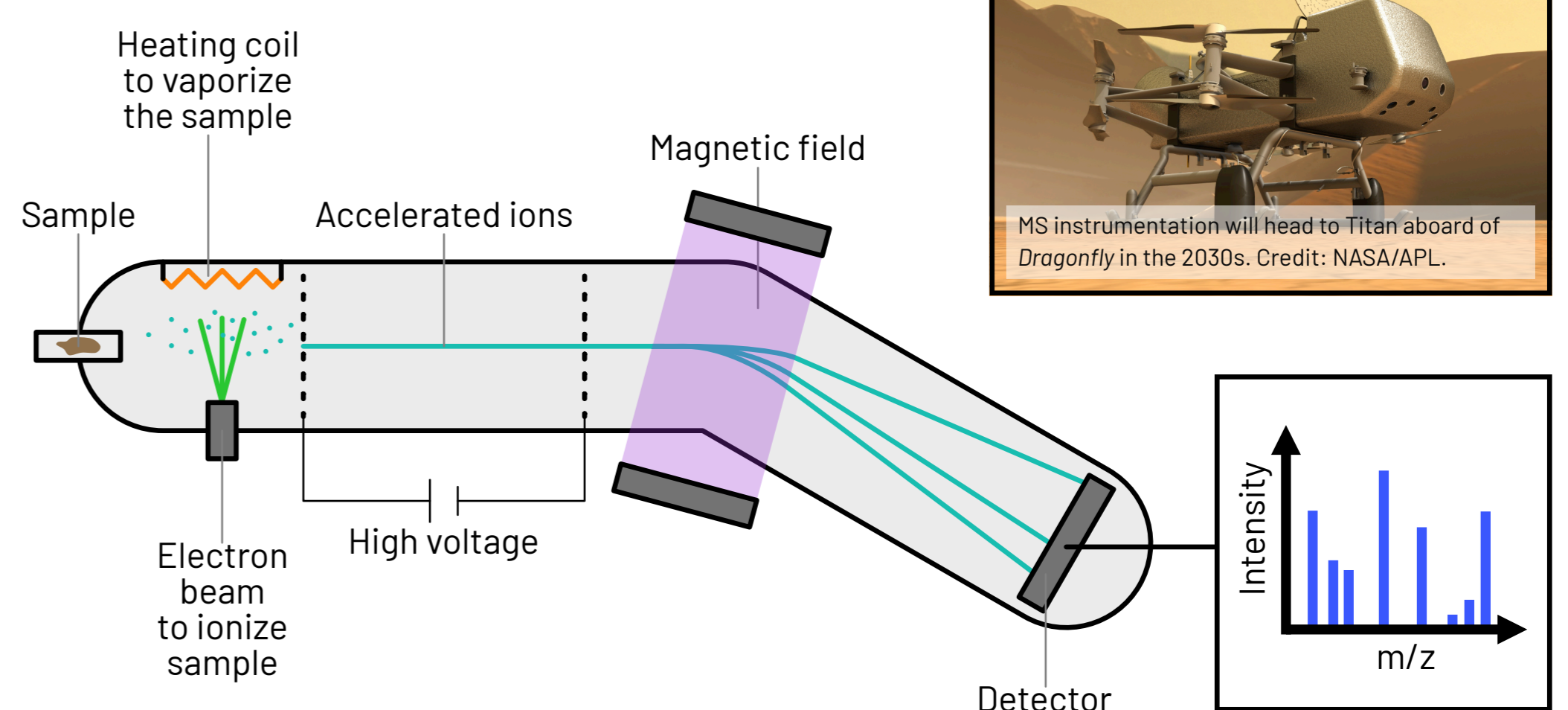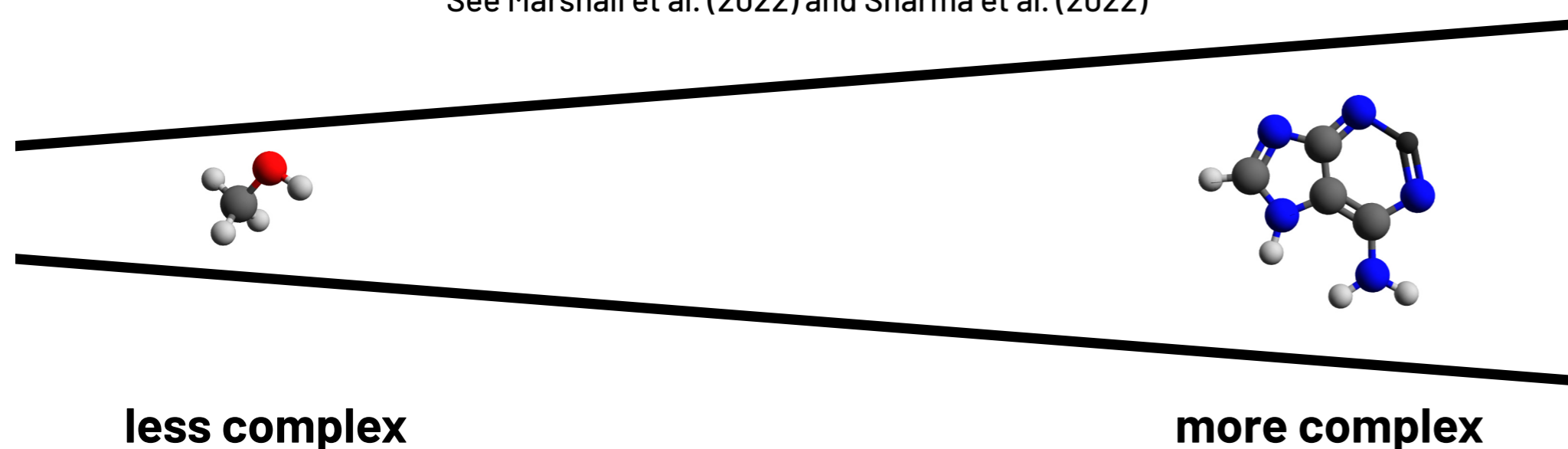
Timothy D. Gebhard*, Aaron C. Bell*, Jian Gong*, Jaden J.A. Hastings*, G. Matthew Fricke, Nathalie Cabrol, Scott Sandford, Michael Phillips, Kimberley Warren-Rhodes, Atılım Güneş Baydin
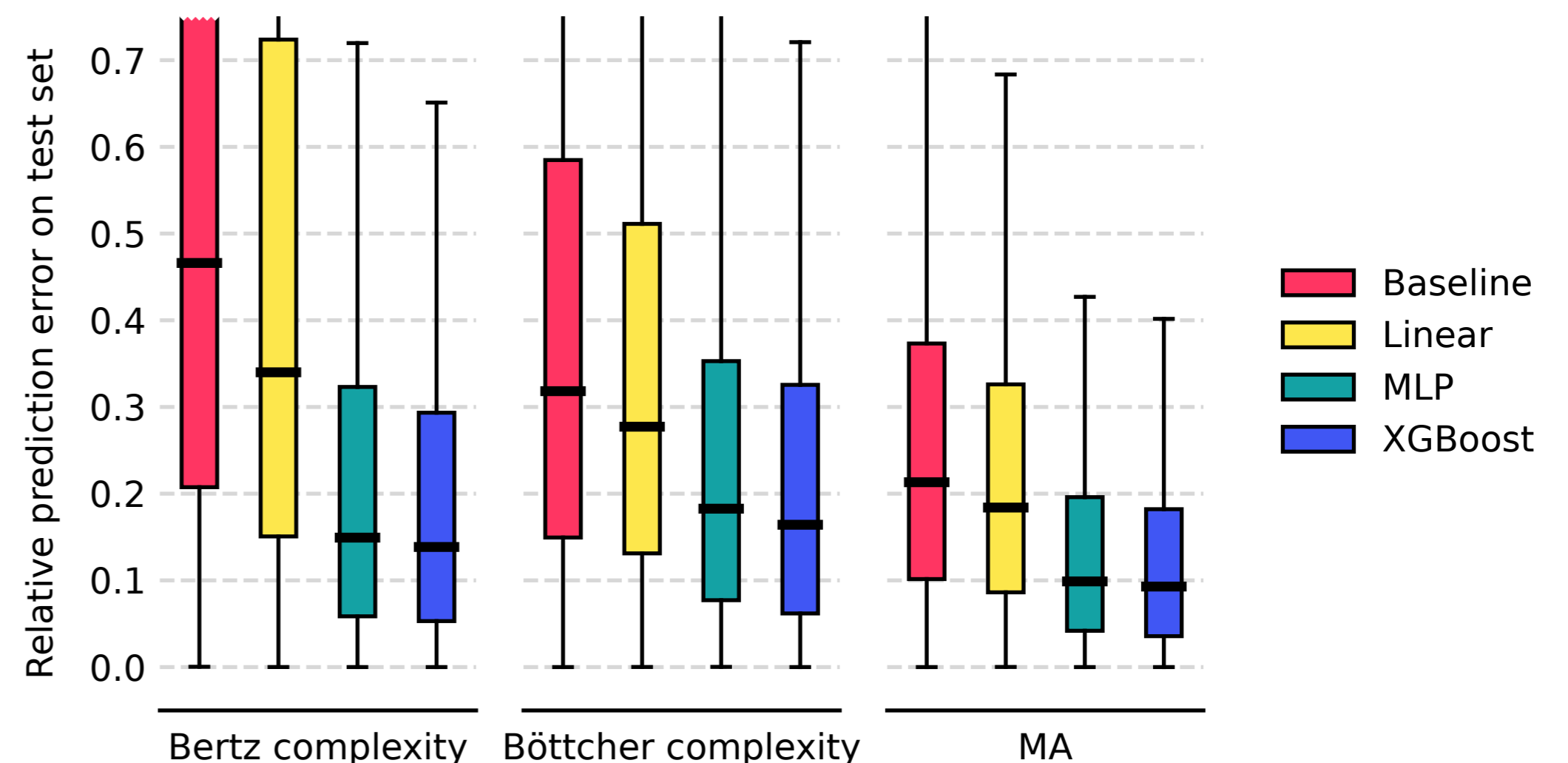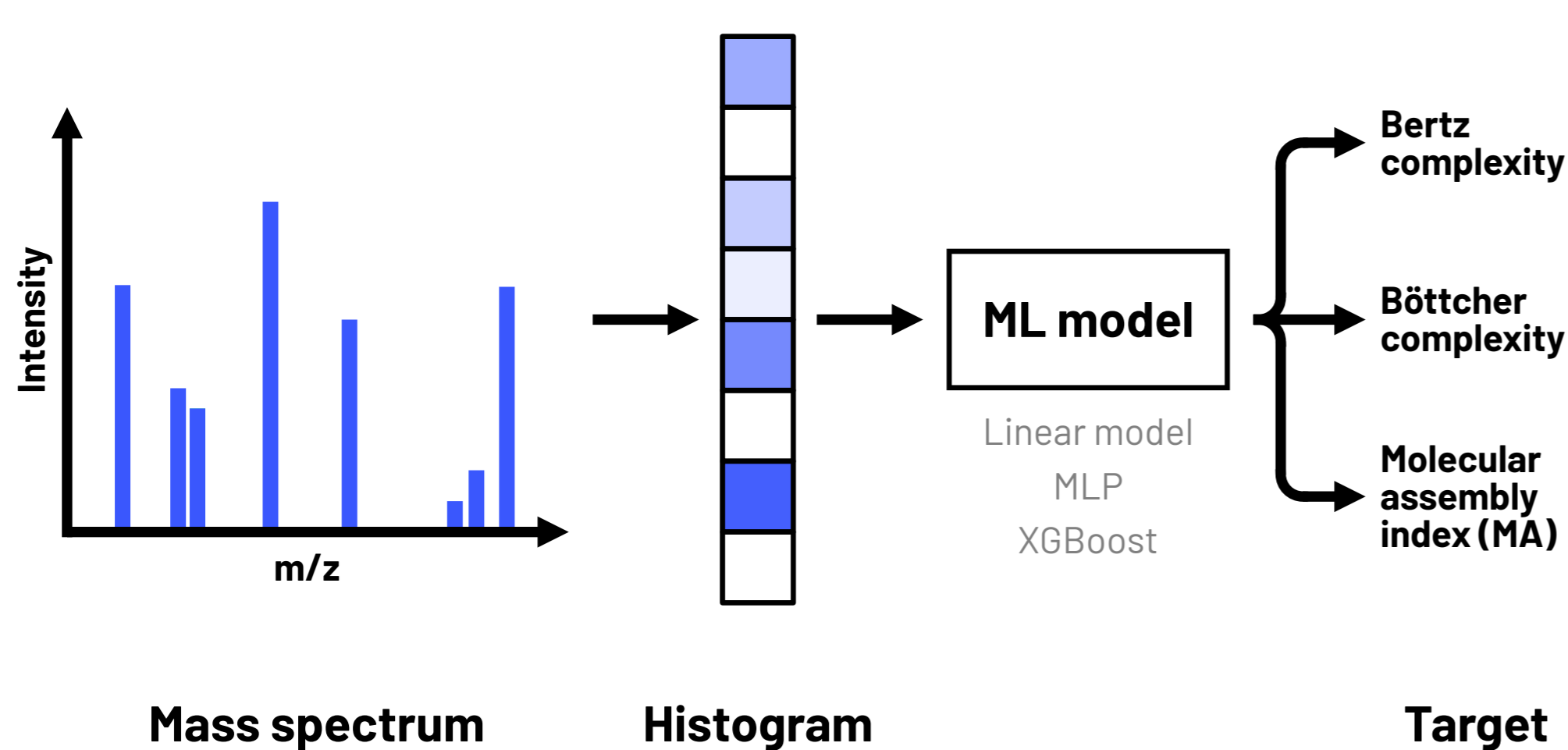
Key hypothesis:

»Only life is able to create complex molecules in larger quantities.«

See Marshall et al. (2022) and Sharma et al. (2022)

less complex   **more complex**

MS instrumentation will head to Titan aboard of *Dragonfly* in the 2030s. Credit: NASA/APL.

1   **Aim:** Connect molecular complexity metrics to experimental data.

2   **Mass spectrometry (MS):** How to identify molecules in the field.

Mass spectrum   Histogram   Target

ML model

Linear model
MLP
XGBoost

→ Bertz complexity
→ Böttcher complexity
→ Molecular assembly index (MA)

3   **ML workflow:** Pipeline from mass spectra to complexity.

4   **Results:** Distribution of (relative) prediction error on our test set.