

A PDF is not a CDF

**Non-Parametric Information-Theoretic
Measures of One-Dimensional
Distribution Functions from
Continuous Time Series**

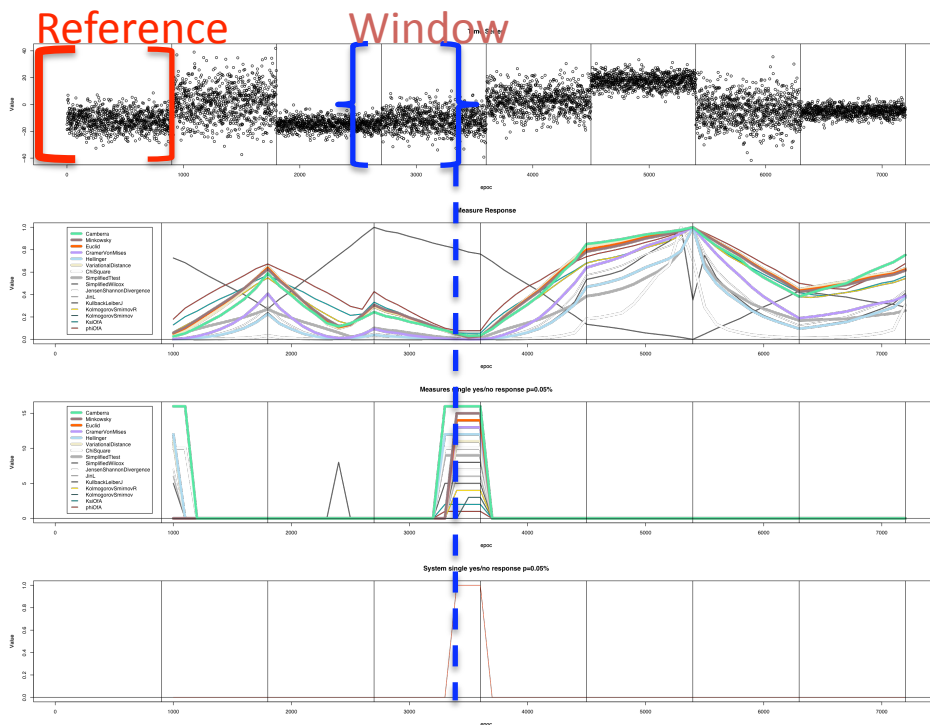
Paolo D'Alberto and Ali Dasdan

Yahoo! Inc.



A PDF is not a CDF: We extended distance measures so that to use CDFs (23 distance functions)

Our Approach:



Series: stochastic instance

Reference window: from series to CDF (F_r)

Moving Window : from series to CDF (F_w)

Measures:

Apply all measures of $\mathbb{D}(x = F_r, y = F_w)$

Confidence for all measures (i.e., $F_r = F_w$):

Given a confidence or p-value (5%) each measure flags when $F_r = F_w$ (i.e., $R \sim W$)

Disagreement and final vote ($F_r = F_w$):

We declare $F_r = F_w$, when no more than 1 out of 10 measures **disagrees** (disagreement 0.1)

Confidence for all measures: (spoiler) it is simplified by using CDF
(come by and I will show you why/how/when)

