

Collaborative Infrastructure for **Test-Driven** Scientific Model Validation

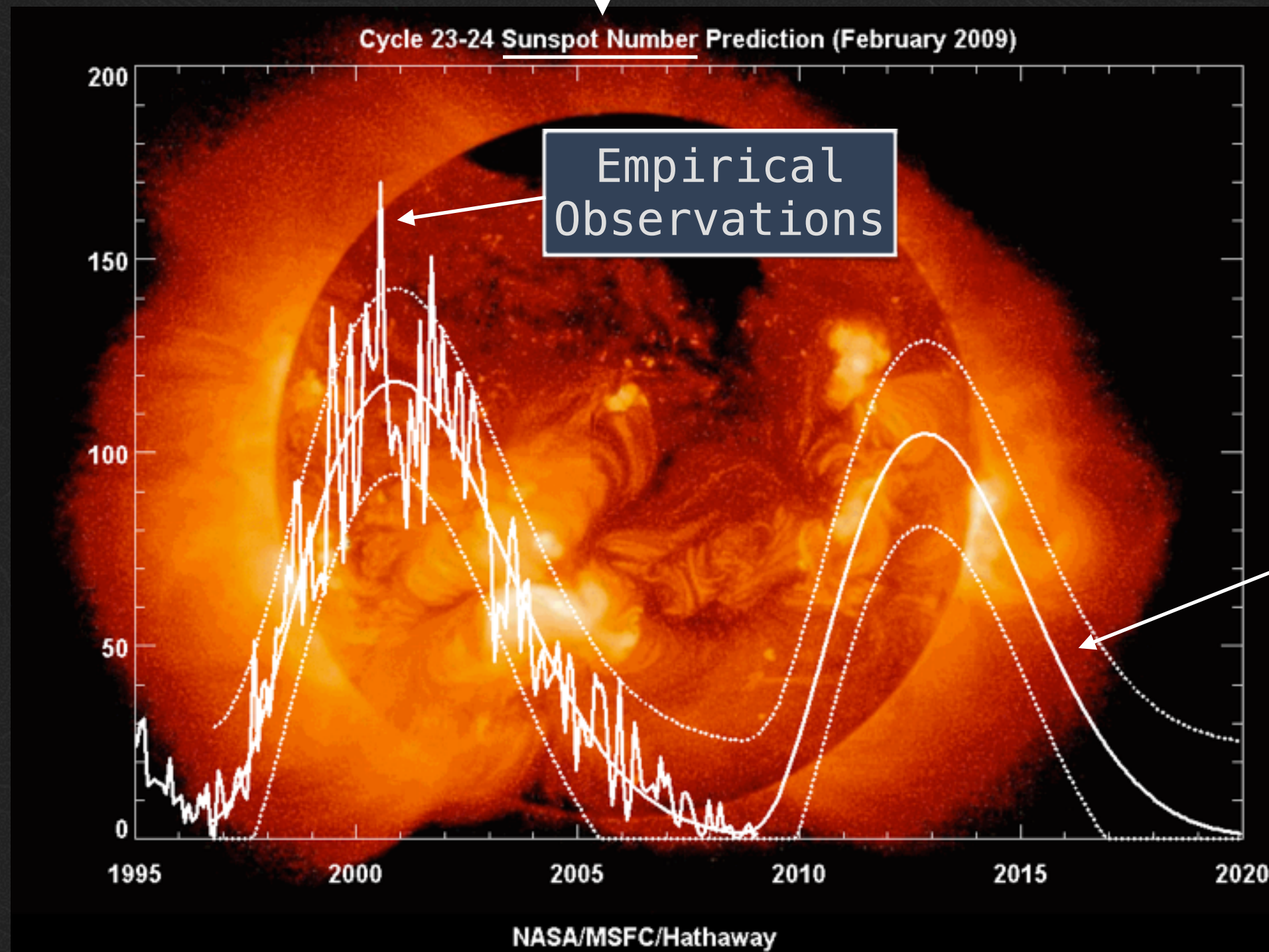
Cyrus Omar + Jonathan Aldrich

carnegie mellon university

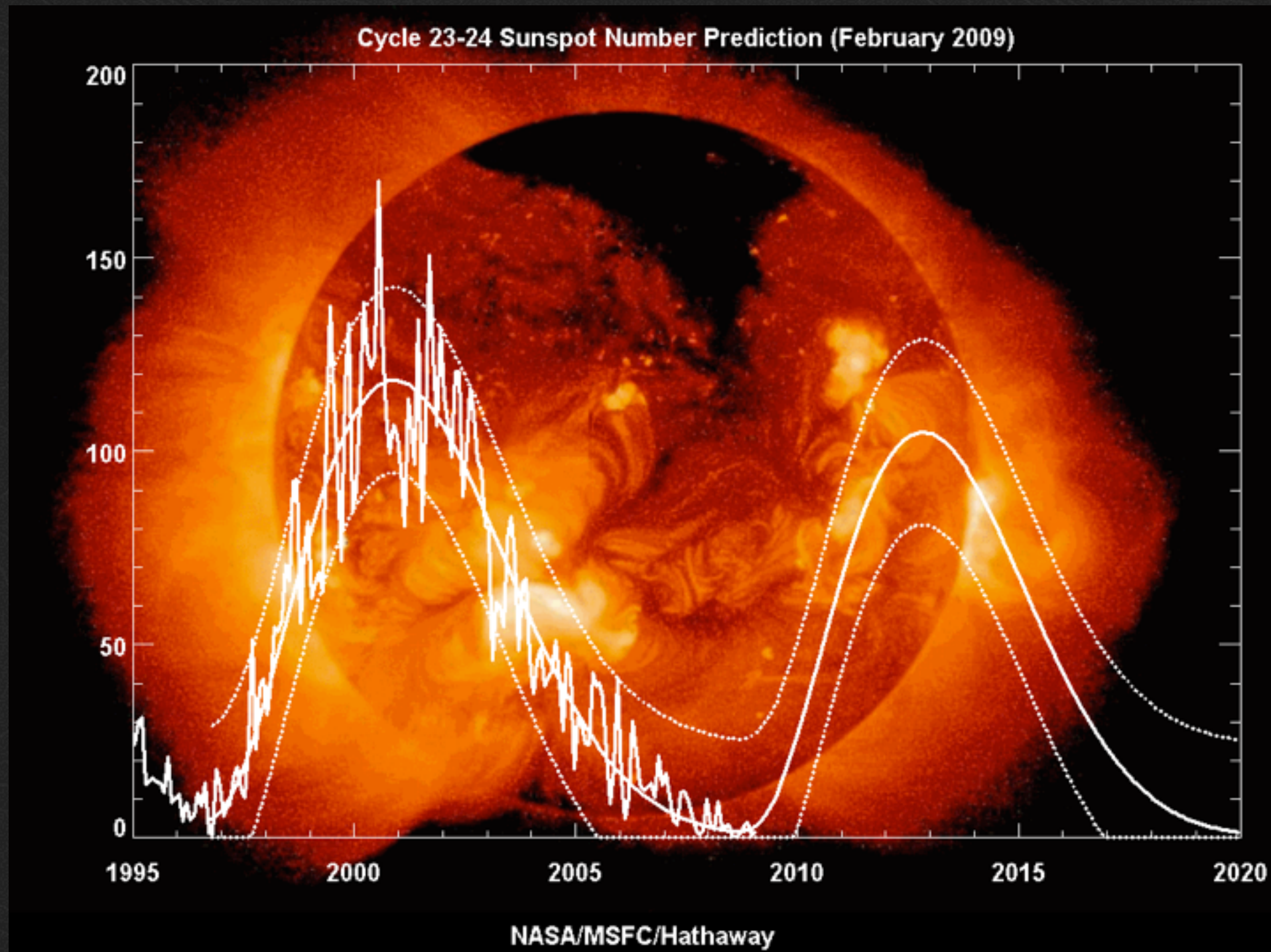
Richard Gerkin

arizona state university

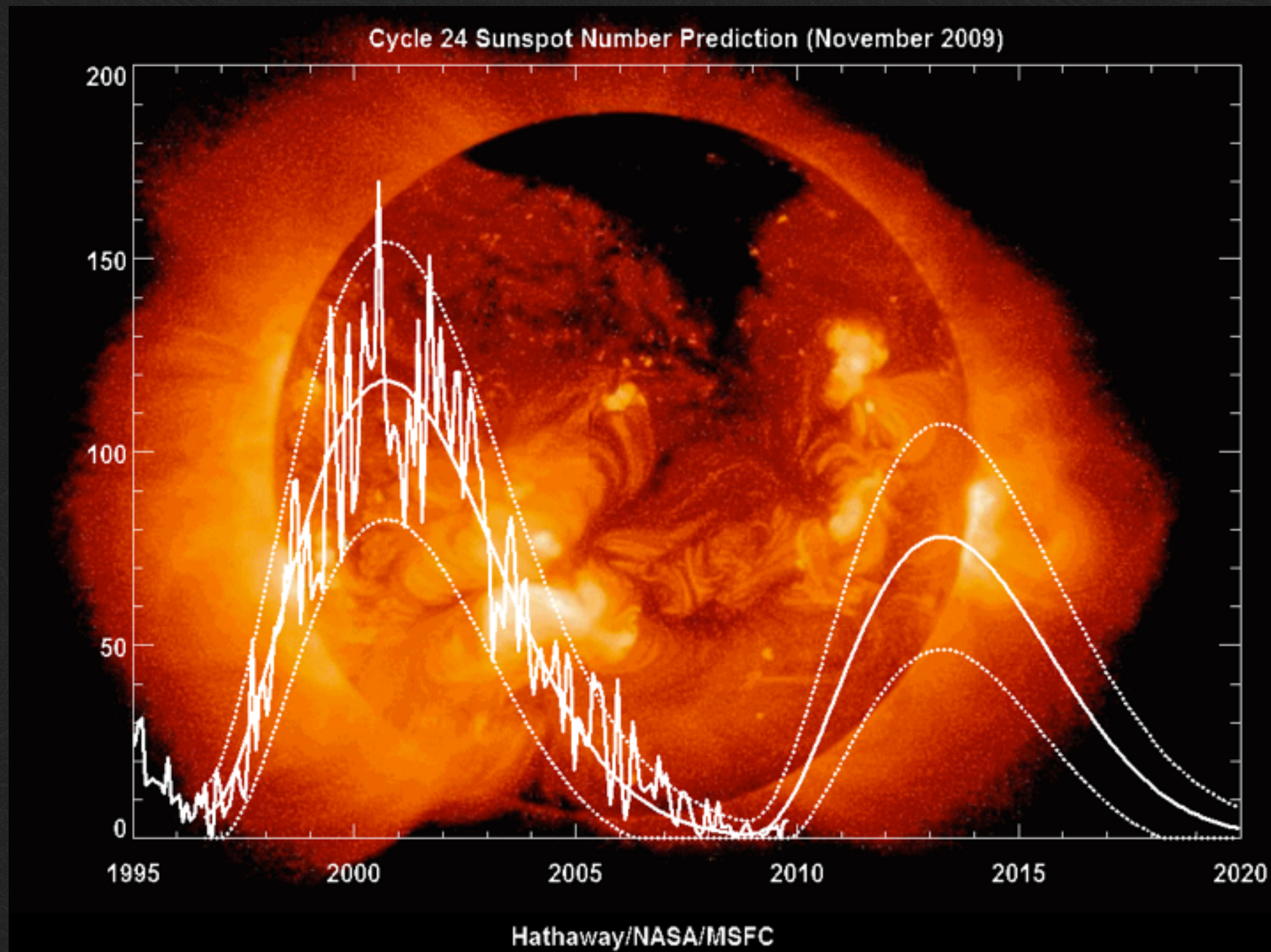
Observable
Quantity



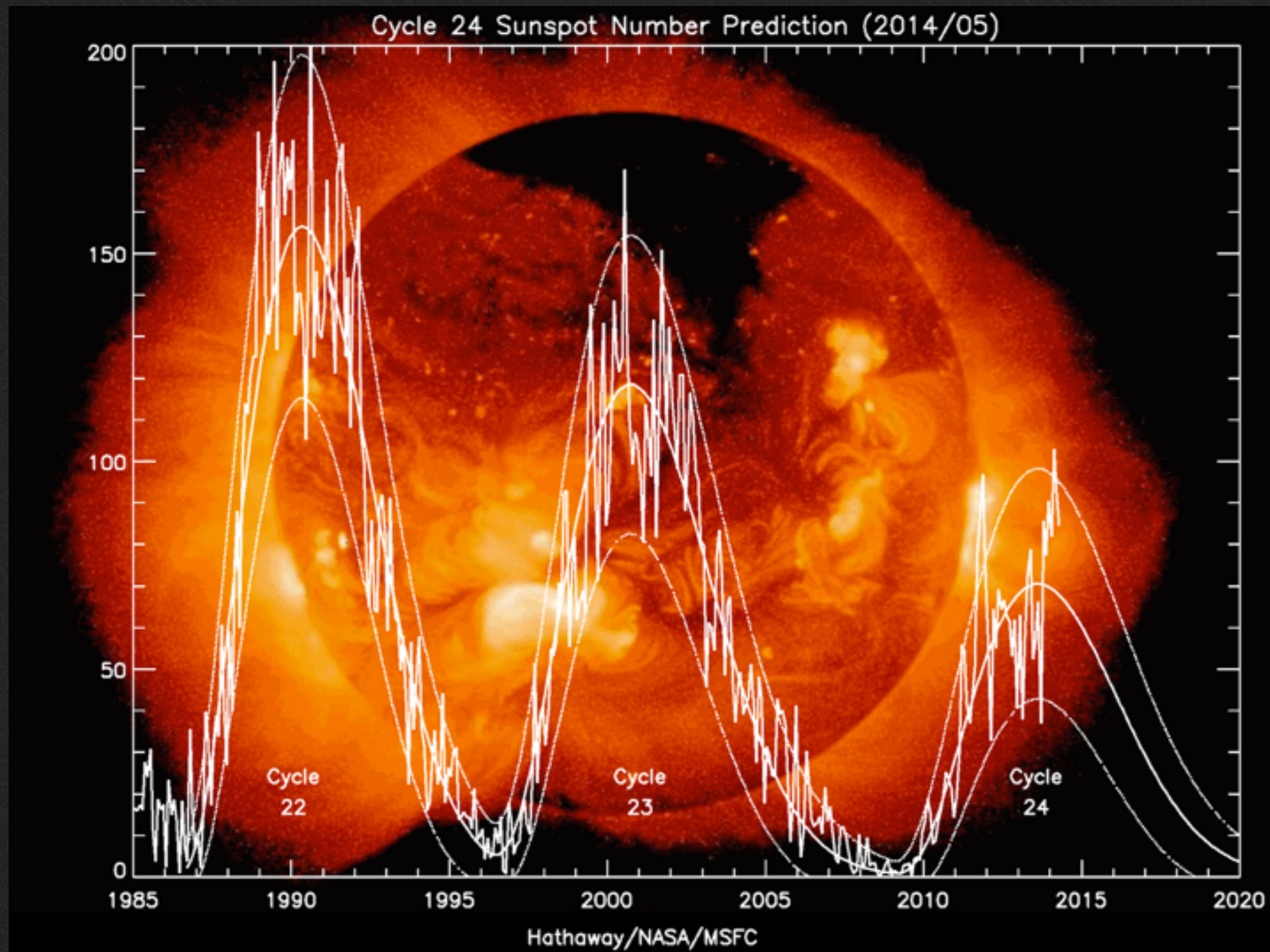
Science is Dynamic



Science is Dynamic



Science is Dynamic



A synthesis of solar cycle prediction techniques

David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann

NASA Marshall Space Flight Center, Huntsville, Alabama

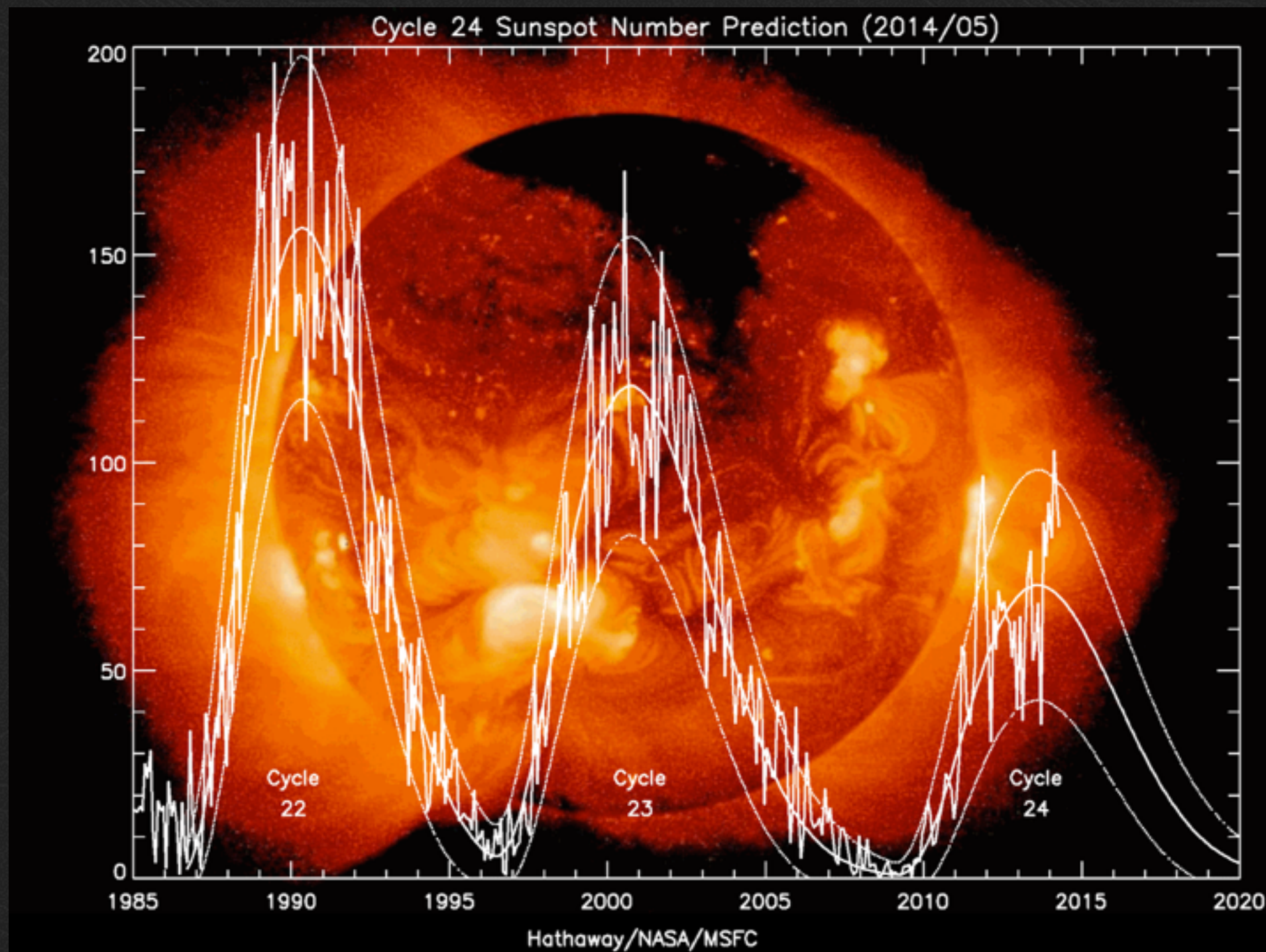
Empirical
Observations

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

Models

Goodness-of-Fit



The state of the art?

“Other [models] were examined and found promising but were not considered due to our inability to test the techniques”

PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.

For *Anno* 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little with-
out Temple-Bar, and James Allestry in Duck-Lane,
Printers to the Royal Society.

scientific community ~ **software development team**
models ~ **software components**
papers ~ **component documentation**
implementations
interfaces
conformance tests

models ~ software components

model capabilities ~ interfaces

observable quantities ~ types

goodness-of-fit tests ~ unit tests

empirical observations ~ test parameters

summary tables ~ test result summaries

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

```
interface SunspotNumberPredictor
  predict : TimeSeries<SunspotNumber> -> TimeSeries<SunspotNumber>

OhlsMethod      : SunspotNumberPredictor = ...
FeynmansMethod  : SunspotNumberPredictor = ...
ThompsonsMethod : SunspotNumberPredictor = ...

SunspotTest : TimeSeries<SunspotNumber> -> Test<SunspotNumberPredictor, Metric> = ...

TestSuite([SunspotTest(cycle_data) for cycle_data in all_cycle_data])
.judge([OhlsMethod, FeynmansMethod, ThompsonsMethod]).view()
```

Model	◊ SunspotTest(19) ◊	SunspotTest(20) ◊	SunspotTest(21) ◊	SunspotTest(22) ▼
OhlsMethod	-55.4	19.1	21.8	4.4
FeynmansMethod	-42.8	9.6	26.9	3.6
ThompsonsMethod	-17.8	8.7	-26.5	-13.6


```

interface SunspotNumberPredictor
  predict : TimeSeries<SunspotNumber> -> TimeSeries<SunspotNumber>

OhlsMethod      : SunspotNumberPredictor = ...
FeynmansMethod  : SunspotNumberPredictor = ...
ThompsonsMethod : SunspotNumberPredictor = ...

SunspotTest : TimeSeries<SunspotNumber> -> Test<SunspotNumberPredictor, Metric> = ...

TestSuite([SunspotTest(cycle_data) for cycle_data in all_cycle_data])
.judge([OhlsMethod, FeynmansMethod, ThompsonsMethod]).view()

```

which programming language? a DSL?

Paper: SciUnit, a Python library

(widely adopted, good FFI support, simple syntax, IPython notebook)

collaborative workflow?

Paper: SciDash, a wiki pointing to IPython notebooks hosted on Github

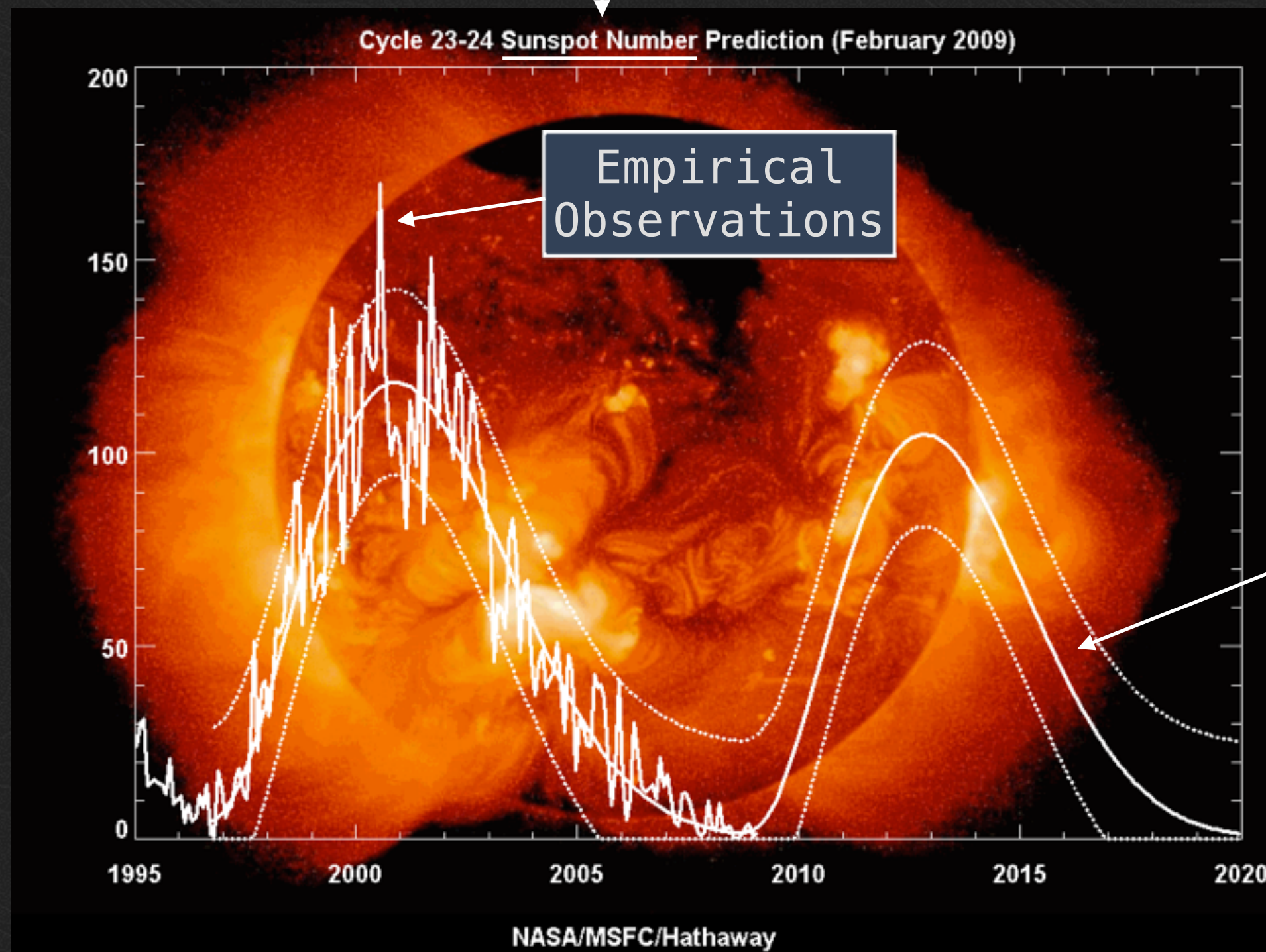
adoption?

Paper: people who organize modeling competitions (e.g. a neural modeling competition we are helping re-start)

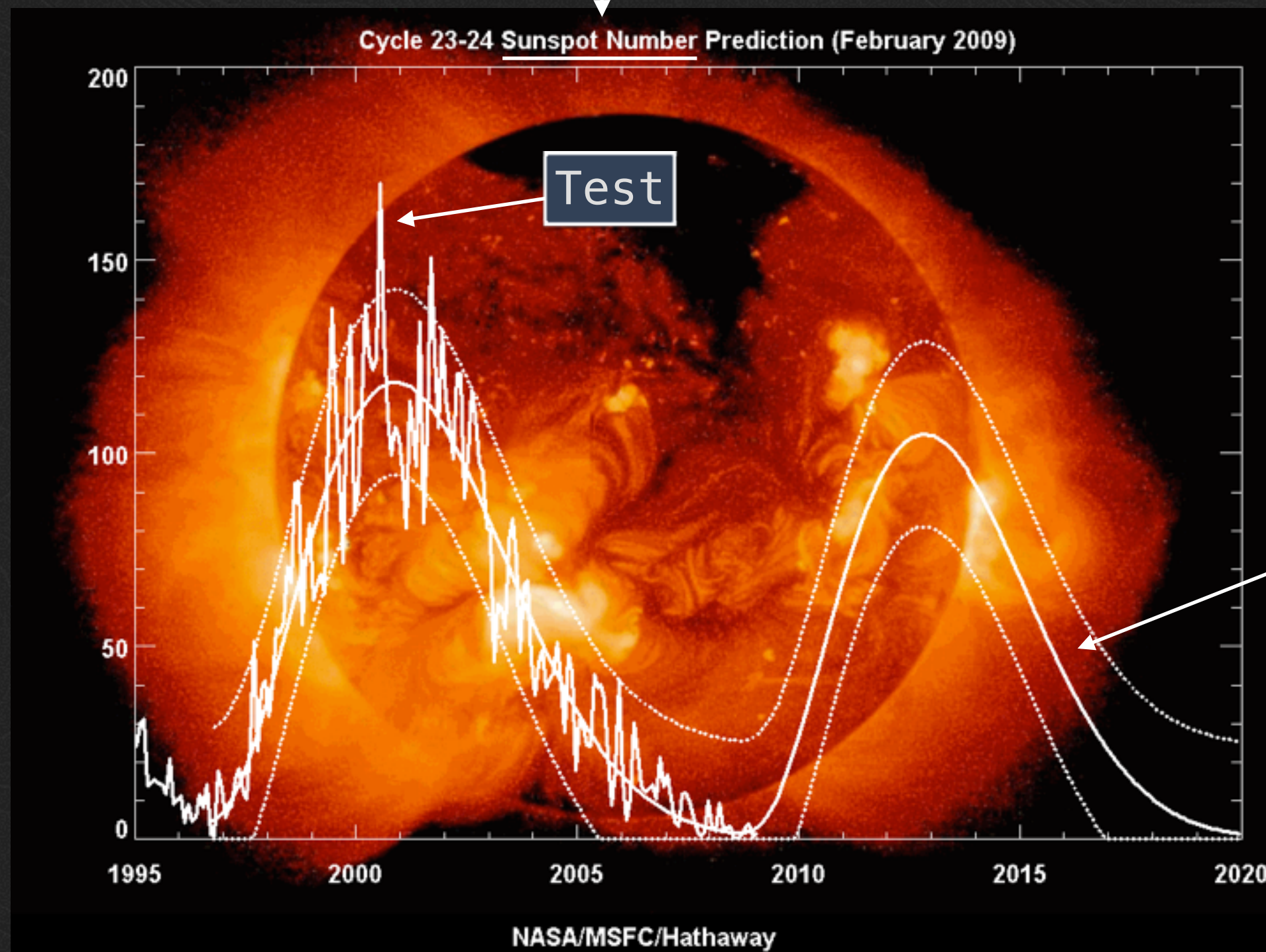
Ongoing work: automatic test and model generation from existing repositories

Contact Me: cyrus@cmu.edu or @neurocy

Observable
Quantity



Data Type



A synthesis of solar cycle prediction techniques

David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann

NASA Marshall Space Flight Center, Huntsville, Alabama

Unit Tests

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Mean cycle	-94.8	-9.1	-53.5	-48.6	59.8
Secular trend	-91.6	8.7	-36.2	-25.3	51.0
Gleissberg cycle	-80.4	18.5	-51.6	-51.1	55.0
Even-odd	-59.3	-	-22.3	-	44.8
Amplitude-period	-74.1	0.3	-61.2	-25.3	49.7
Maximum-minimum	-83.9	21.6	-22.9	-15.0	45.4
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

Models

Goodness-of-Fit