Validating the Transit Spectra: An Automated Flagging System

Kate McCarthy
Jet Propulsion Laboratory, California Institute of Technology

Validating Transit Spectra

Traditionally, we determine whether a spectrum is of “good” or “bad” quality by looking at it manually:

HAT-P-26 b
The Need for Automated Validation

• EXCALIBUR is expanding with more targets and instruments
• It is not feasible to manually validate data products for 1,000+ targets
• How can we automate transit spectra validation?
## Data Quality Flags

<table>
<thead>
<tr>
<th>green</th>
<th>yellow</th>
<th>red</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passed Classifier tests (perfection not guaranteed)</td>
<td>Proceed with caution</td>
<td>Use at your own risk</td>
</tr>
</tbody>
</table>
Example: Points in Transit
The number of points in the full and total transit.

Normalized Flux

Separation (Stellar Radius)
Example: Points in Transit
Example: Points in Transit

Points in Transit Across Targets (G141)

Points in Transit Across Targets (G430L)
Classifier: A Suite of Algorithms

- Points in Full and Total Transit
- Residual Standard Deviation Metric
- Scientific Utility
- Median Error
- Light Curve Symmetry
- Spectral Channels Rejected

planet flag
Planet-Specific Flags

“classifier.flags” state vector

- PLANET: e
  Overall Flag: green

Viewing State Vector:
- Run ID: 1
- Target: 55 Cnc
- Task: classifier
- Algorithm: flags
- State Vec: transit-HST-WFC3-IR-G141-SCAN

PLANT: e
Overall Flag: green

POINTS IN FULL AND TOTAL TRANSIT
Flag: green
Flag Description: Sufficient points between 1st and 4th contact points. Sufficient points between 2nd and 3rd contact points.
Number of points in total transit: 10
Number of points in full transit: 10

PLANET: e
Overall Flag: green

POINTS IN FULL AND TOTAL TRANSIT
Flag: green
Flag Description: Sufficient points between 1st and 4th contact points. Sufficient points between 2nd and 3rd contact points.
Number of points in total transit: 10
Number of points in full transit: 10
Pipeline-Wide Flag Summaries

“classifier.summarize_flags” state vector
Pipeline-Wide Flag Summaries

Points in Transit Across Targets (G141): Full Transit, Total Transit, Pre-Transit Points, Post-Transit Points

Light Curve Symmetry Across Targets (G141): Number of planets vs. Number of points

Residual Standard Deviation Metric (RSDM) (G141): Mean RSDM (Shot noise) vs. Number of planets

Spectral Channels Rejected Across Targets (G141): Spectral channels rejected (%) vs. Number of planets

Median Error Across Targets (G141): Median error (Noise model units) vs. Number of targets
Conclusions

• Classifier is a system of algorithms to flag data quality for each target
• Automates the initial spectra judgment process
• Can alert you of when to proceed with caution
• Offers insight into EXCALIBUR’s overall performance at each stage
• Provides validation infrastructure to support EXCALIBUR’s expansion