PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **126**:711–732, 2014 August

Dissertation Summary

Techniques and Review of Absolute Flux Calibration from the Ultraviolet to the Mid-Infrared

RALPH C. BOHLIN,[1] KARL D. GORDON,[1,2] AND P.-E. TREMBLAY[1]

Received 2014 April 30; accepted 2014 June 05; published 2014 July 22

ABSTRACT. The measurement of precise absolute fluxes for stellar sources has been pursued with increased vigor since the discovery of dark energy and the realization that its detailed understanding requires accurate spectral energy distributions (SEDs) of redshifted Ia supernovae in the rest frame. The flux distributions of spectrophotometric standard stars were initially derived from the comparison of stars to laboratory sources of known flux but are now mostly based on calculated model atmospheres. For example, pure hydrogen white dwarf (WD) models provide the basis for the *HST* CALSPEC archive of flux standards. The basic equations for quantitative spectrophotometry and photometry are explained in detail. Several historical lab-based flux calibrations are reviewed; and the SEDs of stars in the major online astronomical databases are compared to the CALSPEC reference standard spectrophotometry. There is good evidence that relative fluxes from the visible to the near-IR wavelength of $\sim$2.5 μm are currently accurate to 1% for the primary reference standards, and new comparisons with lab flux standards show promise for improving that precision.

1. INTRODUCTION

Flux values in physical units for astronomical objects are required to make comparisons to physical models (Kent et al. 2009). Such comparisons are done regularly for observations at all astronomical distances; from solar system objects to individual stars to nebulae to entire galaxies. The need for a more precise flux calibration has recently been highlighted by the fact that the uncertainties in the flux distribution of stellar standards are the dominant systematic error in measuring relative fluxes of redshifted Type Ia (SN Ia) supernovae and, thus, in determining the nature of the dark energy that is driving the observed accelerating cosmic expansion (Sullivan et al. 2011). A more detailed understanding of dark energy requires a precise and accurate comparison between the fluxes of distant and nearby supernovae in the rest frame. Quantitative descriptions of dark energy in terms of the Einstein equations of general relativity are significantly improved when the relative flux with wavelength is known to an accuracy of 1% or better (Aldering et al. 2004). Proof of this primary justification for precise flux standards is the award of the 2011 Nobel Prize in Physics for the discovery of dark energy, which uses the SN Ia technique for mapping the history of cosmic expansion. The study of circumstellar dust rings in the mid-infrared (IR) (Su et al. 2006) provides a second example of the need for a more precise absolute flux calibration. In this case, model atmosphere calculations provide the basis for absolute mid-IR stellar fluxes; and more precise lab based absolute flux measurements of stellar standards are required to verify and improve the model atmospheres. The models are fit to the visible and near-infrared stellar SEDs, where there is little emission from the dust; and the dust signature is the difference between the measured mid-IR flux and the fitted model. Thus, the precision of the model flux distribution from the visible through the mid-IR is critical.

Our ability to measure stellar brightness has progressed from ancient times, when the eye was the only detector and the precision was about one magnitude, i.e., a factor of 2.512. Photographic film provided a somewhat more accurate brightness measures that improved again with bolometers and photomultiplier tubes in the 1950s and 1960s. Today, the state-of-the-art of modern two-dimensional electronic detectors makes possible a precision of order 1% in the determination of the physical energy distributions of stars. The precision of laboratory reference standards has also improved to better than 1% in absolute physical flux (Brown et al. 2006), but there are no recent results on the comparison of modern lab standards to stars. Currently, the best stellar flux standards are limited by the precision of model atmosphere calculations of pure hydrogen WD stars (see § 3.2.3.).

In principle, one nonvariable star with a known SED is sufficient to establish flux distributions for all stars in the sky. Ideally, a spectrophotometer located above any atmospheric absorption could measure the brightness of any star relative to the one standard candle, as long as dynamic range, linearity of response, and out-of-band stray light were not serious issues. Deustua et al. (2013) also review absolute flux calibrations but with a much broader brush that covers the entire electromagnetic spectrum from gamma waves to radio wavelengths. This work

[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218.
[2] Sterrenkundig Observatorium, Universiteit Gent, Gent, Belgium.

goes into greater detail on the restricted range of ultraviolet (UV) to mid-IR. Section 2 reviews the propagation of point-source stellar flux standards to absolute surface brightness of diffuse objects and details the mathematical basis of instrumental flux calibration. Section 3 reviews attempts to establish the fluxes of a few stars relative to laboratory flux standards, while § 4 presents the methodology for using model atmosphere calculations for standard star SEDs. In particular, § 4 explains the use of pure hydrogen WD models and *Hubble Space Telescope* (*HST*) spectrophotometry to establish a set of UV/optical/near-IR flux standards, which are available in the CALSPEC[3] archive. In § 5, several other archives of stellar flux standards are reviewed, and sample SEDs are compared to CALSPEC standards.

The CALSPEC database is the repository for the SEDs resulting from the work described in this review. Included in CALSPEC are the models for the three primary standard WDs, which are nonlocal thermodynamic equilibrium (NLTE) model atmosphere calculations as normalized to an absolute flux level defined by a reconciliation of physics-based visible and IR absolute measures. While SEDs from other sources appear in the CALSPEC database, the most precise and internally consistent set of fluxes are *HST* spectrophotometry mainly from the Space Telescope Imaging Spectrograph (STIS) and secondarily from the Near Infrared Camera and Objective Spectrograph (NICMOS) instruments that are calibrated with equally weighted observations of the three primary WDs. Estimates for IR fluxes from model atmosphere grids are often included with the STIS and NICMOS SEDs for wavelengths longward of the limits of those instruments.

2. CALIBRATION BASICS

2.1. Concept

In principle, the generation of a network of stellar flux standards is just a simple matter of measuring the background-subtracted net signal n in some units like electrons per second from the program star and N from a primary standard of known flux F, where the same instrument is used for both stars and where F is at the same spectral resolution as the measurements. The flux f of the program star is simply the ratio of the signals times the flux of the primary standard

$$f = \frac{Fn}{N} \tag{1}$$

or

$$f = Sn, \tag{2}$$

where $S = F/N$ is the instrumental sensitivity. This simple case is for a stable instrumental configuration with a linear response. Stability means that repeated observations produce the same response, while linearity implies that the count rate is directly proportional to the physical flux F; i.e., the ratio of flux to count rate will be the same ratio of F/N over the dynamic range of the system. There is no restriction on the entrance slit or extraction aperture as long as the same choice is made for both stars and the extracted count rate is repeatable for both stars. The measured count rate can be in a certain aperture radius for point source photometry or of a fixed height on the detector of an imaging spectrophotometer. The sensitivity S is really more properly an inverse sensitivity, because a more sensitive instrument will have a higher count rate for a source of the same flux.

One complication arises when the spectral resolutions of the flux standard differs from the unknown star. A common example is a standard star with a tabulated medium-resolution SED. In the case of a spectrometer with a resolution that is lower than the tabulated resolution of the standard, the calibration S as a function of wavelength is defined as the convolution of the known SED, F, with the instrumental line-spread function, LSF, divided by the count rate spectrum, N, of the standard convolved with the LSF of the standard star spectrum, which brings the numerator and denominator spectra of S to the same resolution and enables a pixel-by-pixel division of F by N after resampling to the same wavelength scale. This procedure may fail for the case where a low-resolution standard star SED must be bootstrapped to a calibration of a much higher resolution spectrometer where the sensitivity of the high-resolution data changes significantly over the resolution element of the standard SED. For example, a single echelle order may have a variation in sensitivity by a factor of 10 or more over a wavelength range covered by only one or a few resolution elements of the flux standard. The calibration of high-resolution spectroscopy is greatly simplified, whenever a high-fidelity, high-resolution model atmosphere calculation is available. For broadband photometry, the average flux of the standard over the bandpass must be calculated as detailed in the following, which is straightforward, if the spectral resolution is much better than the band width.

Other complications include nonlinearity and changing instrumental response with time or temperature. For CCD detectors, charge transfer efficiency (CTE) during readout is not perfect and degrades with time, especially in the presence of ionizing radiation which creates charge traps. A source near the readout amplifier will be brighter than when the same star is placed farther away, which creates a nonlinearity that must be corrected in the flux calibration process. A common nonlinearity for pulse counting detectors is coincidence of pulses at the higher count rates. Degradation of optical surfaces and detector quantum efficiency over the long term must be tracked and accounted in the correction of observed response to a reference epoch. For example, Figure 1 shows the mean degradation for one ultraviolet (UV) mode of STIS since its installation

[3] http://www.stsci.edu/hst/observatory/crds/calspec.html.

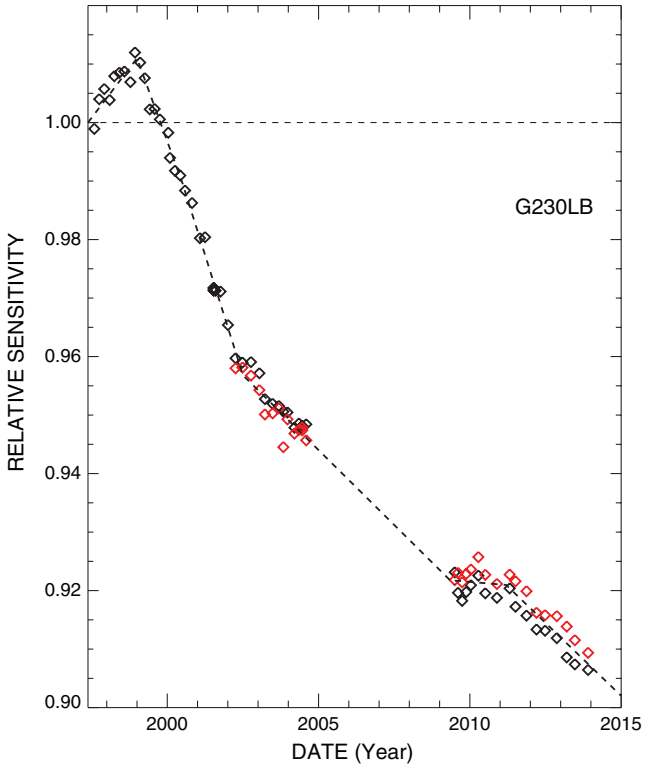

FIG. 1.—Changing sensitivity with time for the G230LB CCD mode of STIS on *HST*. *Diamonds*: average response to the monitoring standard AGK+81°266 over the 2000–3000 Å range; *black diamonds*, center of the CCD detector; *red diamonds*, E1 position near the readout amp. *Dashed lines*: piecewise linear fit to the changing sensitivity, as normalized to unity at the beginning of the mission. All data points are corrected for CTE losses. There is a data gap of almost five years before STIS was repaired by a servicing mission in 2009. Presumably, the small sensitivity increase at the beginning of the mission is caused by evaporation of contaminants from the optical surfaces.

in the *HST* in 1997, while Anderson & Bedin (2010) provide a good example of a CTE correction. Because CTE and optical degradation both manifest as slow losses in sensitivity, separate calibration programs are required to isolate and quantify the two different effects.

In the next section, N_e in electrons per second represents the instrumental response corrected for all nonlinearities, while corrections for temporal instabilities can be applied at any stage to the sensitivity, the count rate, or just the final fluxes.

2.2. Definitions & Equations for Spectrophotometry and Photometry

2.2.1. Photometry

Point source.—The *HST* method of flux calibration for filter photometry does not involve color corrections or nominal wavelengths and is always defined in terms of the photon weighted mean flux over the bandpass in wavelength units, where our flux is in energy units (e.g., erg cm^{-2} s^{-1} Å^{-1})

$$\langle F_\lambda \rangle = \frac{\int F_\lambda \lambda R d\lambda}{\int \lambda R d\lambda} = S_\lambda N_e \tag{3}$$

or in frequency units (e.g., erg cm^{-2} s^{-1} Hz^{-1})

$$\langle F_\nu \rangle = \frac{\int F_\nu \nu^{-1} R d\nu}{\int \nu^{-1} R d\nu} = S_\nu N_e \tag{4}$$

(Koornneef et al. 1986, Rieke et al. 2008). R is the system fractional throughput, i.e., the total system quantum efficiency; S_λ and S_ν are the instrumental sensitivities as a function of wavelength and frequency, respectively; and the integrals are computed over the full bandpass of the filter. The integration is done in photon units ($F_\lambda \lambda$), because UV/optical/near infrared (NIR) detectors are generally photon-detection devices, rather than total-energy-sensing bolometers, although Bessell & Murphy (2012) demonstrate that counting photons is equivalent to integrating the energy. Some authors (e.g., Cohen et al. 2003) define our product λR as their response function of the system. Also, see Stritzinger et al. (2005).

The instrumental count rate N_e can be either measured in an infinite aperture or calculated as

$$N_e = A \int \frac{F_\nu}{h\nu} R d\nu = \frac{A}{hc} \int F_\lambda \lambda R d\lambda, \tag{5}$$

where A is the collecting area of the primary mirror, h is Planck's constant, and c is the speed of light. The predicted throughput R can be adjusted as needed to make the predicted and measured count rates equal for observations of a stellar flux standard. N_e represents the instrumental response after making any required corrections for nonlinearities and temporal changes. For crowded field photometry, N_e is often measured in small radius apertures and corrected for the fractional enclosed energy. For example for the Advanced Camera for Surveys (ACS) on *HST*, Bohlin (2012) tabulates the fractional enclosed energy for isolated bright stellar images.

Source independent instrumental sensitivities S are defined by dividing the mean flux by the detected electrons s^{-1}, N_e, in an infinite-radius aperture. If N_e from equation (5) is substituted in equations (3)–(4),

$$S_\lambda = \frac{hc}{A \int \lambda R d\lambda} \tag{6}$$

$$S_\nu = \frac{h}{A \int \nu^{-1} R d\nu}. \tag{7}$$

For example, the *HST* standard flux units are normally per unit wavelength; and the constant S_λ appears in the headers of *HST* photometric images with the keyword name *photflam*. For NICMOS and Wide Field Camera 3 (WFC3), S_ν with

the keyword name *photfnu* is also included in the headers. Other instrumental archives, e.g., *Spitzer*, store calibrated images in units of surface brightness.

The *HST* calibration constants are normally derived from the source-independent equations (6)–(7) after any required adjustments are made to the R estimated from the product of laboratory component QE measurements. These adjustments are derived by making the measured $N_e(\mathrm{obs})$ in an infinite aperture match the predicted $N_e(\mathrm{pred})$ calculated from equation (5). In practice, a radius of something like the 5.5″ for ACS is defined as "infinite" (Sirianni et al. 2005, Bohlin 2012); and the primary pure hydrogen WDs G191B2B, GD71, and GD153 are the preferred standards used for F in equation (5). In theory, models of the instrumental PSF could help define the encircled energy for an infinite aperture. In the case of *HST*, considerable effort has been expended on the Tiny Tim point-spread function (PSF) modeling software. However, the Tiny Tim user manual states: "At short wavelengths, it may not be possible to compute a PSF larger than 7″. Generally, the models are not very good past a radius of ∼2″, due to the effects of scatter and high-frequency aberrations" (Krist and Hook 2004).

The reconciliation of laboratory component throughputs versus the truth of standard stars is achieved by adjusting the normalization of the filter throughput or even by changing the quantum efficiency (QE) as function of wavelength for the detector or filter when sufficient information exists (e.g., de Marchi et al. 2004; Bohlin 2012). Thus, information about individual component throughputs, such as the telescope or detector QE, may be inferred when reconciling sensitivities for several filters with different central wavelengths.

To complement these estimates of mean flux for stars imaged in a particular filter, an associated wavelength is often useful. In addition to the nominal wavelength λ_o of Reach et al. (2005), other common definitions are the *mean* and *effective* wavelengths

$$\lambda_{\mathrm{mean}} = \frac{\int \lambda R d\lambda}{\int R d\lambda}, \tag{8}$$

$$\lambda_{\mathrm{eff}} = \frac{\int F_\lambda \lambda^2 R d\lambda}{\int F_\lambda \lambda R d\lambda}. \tag{9}$$

Perhaps, most useful is the source independent *pivot-wavelength* λ_p and associated *pivot-frequency* ν_p, where $\lambda_p \nu_p = c$ and $\langle F_\lambda \rangle \lambda_p = \langle F_\nu \rangle \nu_p$:

$$\lambda_p = \sqrt{\frac{c\langle F_\nu \rangle}{\langle F_\lambda \rangle}} = \sqrt{\frac{\int \lambda R d\lambda}{\int \lambda^{-1} R d\lambda}}. \tag{10}$$

These various measures of the associated wavelengths for a filter are given by Koornneef et al. (1986).

Having calculated the source independent *pivot-wavelength* λ_p, equation (10) provides a convenient formula for calculating S_λ from S_ν values.

$$S_\lambda = \frac{c S_\nu}{\lambda_p^2} \tag{11}$$

Diffuse source.—For the surface brightness of diffuse sources, i.e., the specific intensity or radiance I, there are analogous equations for the instrumental calibrations, where Ω_{pix} is the size of a pixel in steradians or arcsec^2. If N_I in electrons per second represents the linearized instrumental response for the count rate per pixel in a region of diffuse surface brightness, then

$$\langle I_\lambda \rangle = \frac{\int I_\lambda \lambda R d\lambda}{\int \lambda R d\lambda} = C_\lambda N_I \tag{12}$$

or in frequency units

$$\langle I_\nu \rangle = \frac{\int I_\nu \nu^{-1} R d\nu}{\int \nu^{-1} R d\nu} = C_\nu N_I. \tag{13}$$

The calibration constants for the specific intensity calibration are related to those for point sources by

$$C = \frac{S}{\Omega_{\mathrm{pix}}}, \tag{14}$$

where C and S have either the λ or the ν subscript. The reason that S for a point source and an infinite aperture is required can be visualized by the following gedenken experiment: consider a field of point-source stars with flux F at the same spacing as the pixel grid of the detector, which is observationally indistinguishable from a field of uniform surface brightness. The total count rate N over all pixels for one isolated star is $N = F/S$, while the very same count rate N, but in each pixel, is recorded for the infinite field of point sources, because the contribution to the count rate at any distance for the isolated star is contributed equally by the star located at that same distance in the dense field. Thus, only a division by the solid angle per pixel is required to convert the infinite aperture point source calibration to a diffuse source calibration.

While diffuse source calibration is simple in principle, several practical considerations limit the precision. Because the wings of the PSF may cover a significant portion of the instrumental field-of-view, direct measurement of the signal in an infinite aperture suffers from inaccuracies in the flat-field correction and signal-to-noise (S/N) issues in the far wings, even for a bright isolated star. Alternatively, modeling of the PSF is also fraught with difficulties in calculating the last few percent of the energy in the far wings, which arise from diffraction, atmospheric blurring, ghosting, scattered light, etc. Sources such as galaxies with a diameter comparable to the PSF size will require a different calibration coefficient (eq. [14]) than for

an infinite diffuse source such as the night sky or zodiacal background. Such small objects require an effort to model the geometry of the source.

2.2.2. Spectrophotometry

Point source.—For the flux calibration of spectral data, the equations are analogous to those for photometry, but the count rates, sensitivities, fluxes, and throughput R are one-dimensional arrays ordered by increasing wavelength (or frequency). The implicit assumption is that the instrumental sensitivity varies slowly enough across any pixel so that the sensitivity of that pixel is well represented by its average sensitivity. Thus, the integrals go away in equations (6)–(7), i.e.,

$$F_\lambda = S_\lambda N_e = \frac{hc}{A\lambda R d_\lambda} N_e \tag{15}$$

and

$$F_\nu = S_\nu N_e = \frac{h\nu}{A R d_\nu} N_e, \tag{16}$$

where d_λ and d_ν are the spectral dispersions in angstroms per pixel or hertz per pixel, respectively. The product AR is known as the effective area of the instrument. S can be derived directly from an observation of a known flux standard F divided by the observed count rate spectrum N_e, where F is binned to the resolution of the observed N_e, as discussed in § 2.1. In practice, N_e for an infinitely high extraction width is $N(h)/\epsilon(h)$, where the correction $\epsilon(h)$ for $N(h)$ from a spectral extraction h pixels high is derived from high-signal spectral images.

Diffuse source.—A specific intensity calibration for spectra is analogous to equation (14), except that the uniform source of surface brightness must be limited by a slit of width W in arcseconds in the dispersion direction:

$$C = \frac{S}{mW}, \tag{17}$$

where m is the plate scale (arcseconds per pixel) in the direction perpendicular to the dispersion.

3. COMPARISON OF STARS TO LABORATORY FLUX STANDARDS

Direct comparisons between stars and laboratory flux standards, such as standard lamps, is complicated for ground-based telescopes by the need to place the lamp at sufficient distance from the telescope to simulate a point source and by the differential atmospheric absorption between the standard lamp and the star light. The use of standard, calibrated detectors as the basis for absolute fluxes is hampered by the faintness of the starlight and by the need to establish the throughput of a telescope plus detector system for a collimated, uniform beam that simulates stellar illumination. Making these sorts of comparisons in the Earth's atmosphere is further complicated by the need to know the atmospheric transmission as a function of wavelength and time. The remainder of this section summarizes attempts to establish stellar fluxes above the atmosphere with instrumentation that has been calibrated with respect to lab flux standards. Whenever possible, these laboratory pedigreed results are compared with the *HST*/STIS CALSPEC, WD-based fluxes, which are described in § 4.

3.1. Ultraviolet below 3300 Å

At far-UV wavelengths, the atmospheric opacity is so large that any attempts to measure stellar fluxes must be done from above the atmosphere. Even at balloon altitudes, the absorption by oxygen is overwhelming below 2000 Å. Thus, sounding rockets and satellites are the vehicles used in the far-UV. Although observing times are only a few minutes, a sounding rocket has the advantage that the preflight calibration can be confirmed postflight, as long as the contamination control for reentry and landing is adequate. Early results assumed that the visible fluorescence of fresh sodium salicylate is constant with wavelength in the UV (e.g., Opal 1968; Stecher 1970). At Lyα, the absolute flux was referenced to sealed nitrous oxide (NO) ionization cells. Later, more sophisticated measures of absolute flux became available, as discussed below.

This section is mainly a historical account of early attempts to compare UV starlight to physics based, laboratory flux standards. The ordering is roughly by time, where Figures 2–6 demonstrate improving agreement of the measured UV fluxes with the CALSPEC system from the ∼10% level in the 1970s to ∼3% for the Solar–Stellar Irradiance Comparison Experiment (SOLSTICE) fluxes of Snow et al. (2013). Both the available lab standards and the instrumental techniques have improved over the last five decades, which suggests that a precision of 1% is currently possible.

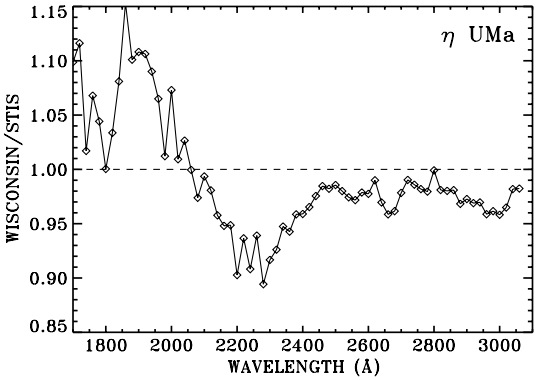

FIG. 2.—Ratio of *OAO-2* fluxes based on a synchrotron calibration to the baseline STIS fluxes for η UMa.

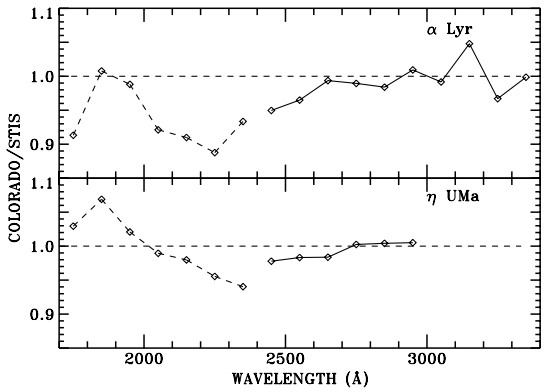

FIG. 3.—Ratio in 100 Å bins of Strongylis & Bohlin fluxes to the baseline STIS fluxes. For the rocket fluxes below 2400 Å (*dashed line*), the reference standard is an NBS standard photodiode detector, while the solid line connects fluxes that are referenced to an NBS standard light source. STIS fluxes for η UMa are not available longward of 3000 Å.

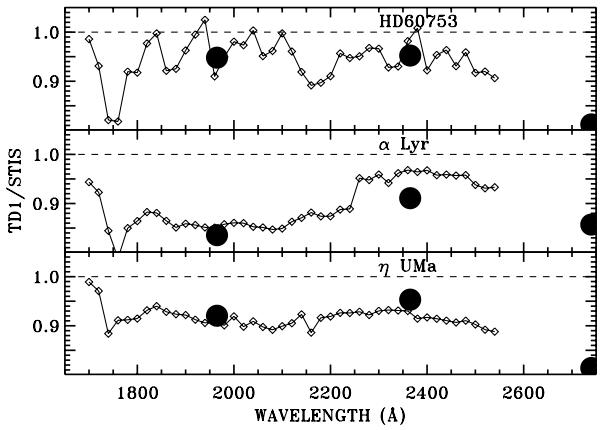

FIG. 5.—Ratio of *TD-1* fluxes to the baseline STIS fluxes with a bin size at the *TD-1* resolution of ~35 Å. *Small diamonds*: early results of Jamar et al. (1976); *large filled circles*: online catalog of Thompson et al. (1978).

3.1.1. Rockets

At the University of Wisconsin, an early sounding rocket program (Bless et al. 1976) provided stellar standards for the flux calibration of the *Orbiting Astronomical Observatory* (*OAO-2*). In the 1370–2920 Å wavelength range, the known spectral energy distribution of 240 MeV electrons in the Wisconsin Synchrotron storage ring established the sensitivity of the seven photometers that comprised the rocket payload. The absolute flux of the synchrotron beam was determined by basic physics and the number of circulating electrons. As the beam degraded from an initial strength of about 50 electrons, the electron count was determined by the incremental losses in signal strength. Other factors entering the calculation were the transmission of the MgF_2 window on the storage ring port and the geometry of the illumination. A complete recalibration was done after the flight and recovery of the payload. Three stars, α Vir, η UMa, and α Leo, were observed; and their measured fluxes were used to update the calibration of the spectrometer on *OAO-2*. These absolute physical fluxes that are based on the physics of synchrotron light emission have estimated uncertainties of 10% shortward of 2000 Å and 5% longward of 2000 Å. Over the wavelength range in common, Figure 2 compares the resulting *OAO-2* flux distribution for η UMa with the modern *HST*/STIS baseline WD system of absolute fluxes from the CALSPEC database.

At the University of Colorado, another rocket program based its measured fluxes on standards available from the National Institute of Standards and Technology (NIST, but was National Bureau of Standards [NBS] at the time). Bohlin et al. (1974)

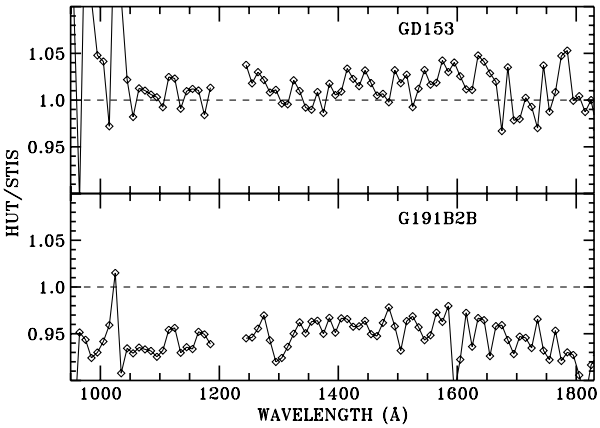

FIG. 4.—Ratio with a bin size of ~10 Å of *HUT* fluxes to the pure hydrogen WD models for two of the *HST* primary flux standards. The gap is where the *HUT* spectra are contaminated by geocoronal Lyα, while the 1302 Å OI aiglow has been clipped from the observations. The large scatter below 1050 Å is caused by a slight mismatch of the strong Lyman absorption lines between the observations and the models.

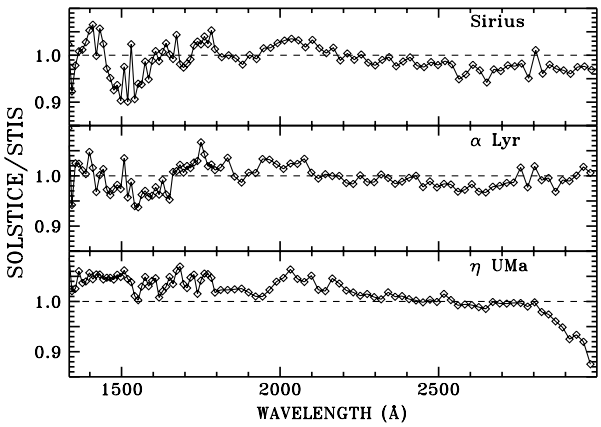

FIG. 6.—Ratio of SOLSTICE fluxes to the baseline STIS fluxes above 1700 Å and *IUE* at shorter wavelengths with a bin size at the SOLSTICE resolution. SOLSTICE fluxes are referenced to a synchrotron flux standard, while the STIS+*IUE* fluxes are on the WD *HST* system.

used an NBS pedigreed standard photodiode detector for the 1164–2385 Å range and a tungsten lamp for 2250–3400 Å to calibrate their spectrometer payload with its ~20 Å spectral resolution. Because of the expected faint stellar signals and the need to calibrate the spectrometer at these low levels, the NBS diode calibration was first transferred to a photomultiplier tube, which was used to scan the collimated, monochromatic input beam in the vacuum laboratory calibration chamber. Every transfer or correction of a NBS pedigree is associated with some added uncertainty, which degrades the precision of the measured stellar fluxes.

The Colorado program used a standard tungsten ribbon light source at the longer wavelengths, but such a hot lamp drawing 38.722 amps must be operated in air to maintain proper convective cooling. A precision aperture in front of the tungsten ribbon defined an effective point source when the lamp was placed at a distance of 73 m from the flight spectrophotometer. However, this laboratory arrangement required a correction for the attenuation by the atmosphere (Strongylis & Bohlin 19791976), which caused the uncertainty to increase toward shorter wavelengths. The lamp calibration was used for the final calibration of the flight spectrometer down to 2400 Å, where the uncertainty was estimated to be $+18/-7\%$. A check of the relative sensitivity versus wavelength was provided by the molecular branching-ratio technique from observations of CO, NO, and N_2 spectra in the laboratory vacuum chamber.

Because of a failure in the short-wavelength channel on the Colorado rocket, results for only the 1700–3400 Å range were obtained for the target stars α Lyr, η UMa, and ζ Oph (Strongylis & Bohlin 1976). For two stars, Figure 3 compares those measured fluxes to the baseline CALSPEC database over their available common wavelength ranges in 100 Å bands.

3.1.2. Manned Space Flights

Manned space flights offer opportunities for longer integration times but have the complications of preventing contamination from human effluents and longer times from delivery to return of the flight instrument to the laboratory calibration chamber.

Apollo 17.—Henry et al. (1975) flew a UV spectrometer on the *Apollo 17* mission to the Moon, and the flux calibration was referenced to an NBS photodiode with an uncertainty estimate of 10% over its 1180–1680 Å wavelength range. Six stars were observed, including η UMa, which lies below the *OAO-2* fluxes by as much as 28% (Strongylis & Bohlin 1976).

Hopkins Ultraviolet Telescope.—The *Hopkins Ultraviolet Telescope*, *HUT*, (Kruk et al. 1997) detected light in the 830–1840 Å range. While the final calibration and archival fluxes were based on the modeled fluxes for the WD G191B2B, an independent lab calibration was also done with reference to NIST calibrated photodiodes and to the NIST synchrotron facility. Longwards of 912 Å, the WD and NIST lab calibrations agree to 3% rms with a worst-case difference of 7% near 1350 Å. On average, over the 912–1840 Å range, the lab synchroton calibration differs from the adopted WD calibration by less than 0.5%. At 830 Å, there is a 20% discrepancy between the two methods. Figure 4 illustrates ratios of the archived WD fluxes to the current model spectra for two of these primary *HST* standards.

3.1.3. Satellites

Spectrophotometry from space based satellites provides the best intercomparisons among a diversity of objects. Most space observatories relied on in-flight observations of standard stars and *not* on direct comparison to lab flux standards for their recommended flux calibration. This group includes the *International Ultraviolet Explorer* (*IUE*), the *Far Ultraviolet Spectroscopic Explorer* (FUSE), and the *Hubble Space Telescope* (*HST*). One observatory, *Copernicus* (*OAO-3*), obtained high resolution spectra from stars focussed on a narrow entrance slit. This *Copernicus* data set is not photometric, i.e., observations were not repeatable, because the part of the stellar PSF falling in the slit varied from one acquisition to the next.

In general, laboratory calibrations of space observatories are not done "end-to-end", because the whole instrument package is too big for most vacuum facilities. Furthermore, the time between any laboratory flux calibration and flight operations is long, allowing many opportunities for contamination of the optical surfaces by thin polymer films which can absorb significant fractions of the UV light beam. Often, estimates of total throughput are computed from the throughput, i.e., quantum efficiency (QE), of the individual components from primary mirror to detector with a corresponding accumulation of uncertainties. UV satellite observatories with catalogs of absolute fluxes that are referenced to lab flux standards are *TD-1*, *ANS*, and SOLSTICE.

Belgian/UK Ultraviolet Sky Survey Telescope.—The *Belgian/UK Ultraviolet Sky Survey Telescope* (S 2/68) in the European Space Research Organization (ESRO) *TD-1* astronomical satellite does have a flux calibration traceable to laboratory standards. Launched in 1972, the S 2/68 UV spectrometer package was independently flux calibrated about three months before launch by groups in the United Kingdom (UK) and Belgium against an absolute radiometric detector and a blackbody source, respectively (Humphries et al. 1976). The flight package was relatively small, with only a 27.5 cm diameter primary mirror; and the coverage is 1350–2550 Å with a resolution of 35 Å. There was also a photometer with a bandpass centered at 2740 Å. For the UK absolute calibration, the photomulipier tube used to measure the irradiance (flux) input to the *TD-1* flight instrument was calibrated against a thermopile with a pedigree traceable to the National Physics Laboratory in London. A thermopile with gold-black coatings to absorb the light can measure the total energy in a monochromatic beam but is far less sensitive than a photomultiplier detector tube. Thus, the crux of the absolute flux calibration consisted of a scheme to reduce the

intensity of the illuminating beam by a known geometric factor of 4×10^5 between the illumination of the thermopile and the detector.

For the independent calibration in Belgium, the photomultiplier used to scan the input beam to the flight instrument was referenced to a blackbody light source where the radiant flux could be calculated from the Stefan–Boltzmann relation for the operating temperature of ~500 K. The blackbody radiation at 6 μm illuminated an intermediate thermopile, which could also measure the energy in a monochromatic UV beam. Attenuation at the photomultipier was provided by a gold-coated MgF_2 filter with a measured transmission as a function of wavelength. This technique relies on the absorption coeficients of the gold-black-coated thermopile, which are 0.981 in the UV and 0.966 in the IR.

The adopted final calibration is the average of the independent UK and Belguim results, while the two separate calibrations differ from the mean by as much as 19%. A whole-sky catalog[4] of results is available online for 31,215 stars (Thompson et al. 1978) down to about $V = 10$ for unreddened B stars in four passbands at 1565, 1965, 2365, and 2740 Å. For a subset of the brighter stars, Jamar et al. (1976) present the full spectra with a finer sampling interval. For three stars, Figure 5 compares these *TD-1* fluxes to the baseline STIS dataset over their common wavelength range. Because these stars are too bright for the far-UV STIS MAMA detectors, STIS fluxes have short wavelength limits of 1680 Å.

Astronomical Netherlands Satellite.—The *Astronomical Netherlands Satellite* (*ANS*) collected UV photometry in five bands from 1550–3300 Å, but the catalog (Wesselius et al. 1982)[5] contains only 3573 stars, and the documentation of the laboratory flux calibration (Aalders et al. 1975) is not readily available.

Other UV space missions.—More recently, Snow et al. (2013) have published UV flux distributions for 18 bright A and B stars obtained by the SOLSTICE spectrometer on the *Solar Radiation and Climate Experiment* (*SORCE*) spacecraft, which also measured the solar flux. The absolute fluxes are tied to a synchrotron source, the NIST *Synchrotron Ultraviolet Radiation Facility* (SURF III) (Arp et al. 2000; McClintock et al. 2005), with coverage from 1150–3000 Å at 11–22 Å resolution. The estimated absolute accuracy is 3%. Figure 6 compares the SOLSTICE fluxes for the three stars in common with the STIS CALSPEC archive. However below 1700 Å, the STIS has been supplemented by *IUE* spectra matched in the overlap region, because the three stars are too bright for the STIS far-ultraviolet (FUV) multianode microchannel array (MAMA) detector. Differences between the synchrotron based SOLSTICE and the WD-based fluxes rarely exceed 2σ, i.e., 6%.

[4] http://webviz.u-strasbg.fr/viz-bin/VizieR?-source=II/59B.

[5] http://cdsarc.u-strasbg.fr/viz-bin/Cat?cat=ans&find=+.

3.2. Visible 3300–10000 Å

Compare the discussion in this section to the review by Deustua et al. (2013).

3.2.1. Historic

Early work by Oke & Schild (1970) and Hayes & Latham (1975) on measuring the absolute flux of the primary ground-based standard Vega has survived the test of time. From 3300 to 10800 Å, Oke & Schild used a NBS pedigreed tungsten ribbon lamp and two blackbody cavities operated at the melting point of copper to measure the flux, F(5556), of α Lyr with a value of 3.36×10^{-9} erg cm^{-2} s^{-1} Å^{-1} $\pm 2\%$ at 5556 Å. Hayes & Latham improved the atmospheric extinction corrections and utilized the earlier relative measurements of Hayes (1970) and absolute fluxes in the 6800–10800 Å range from Hayes et al. (1975) to derive $F(5556) = 3.45 \times 10^{-9}$ erg cm^{-2} s^{-1} Å^{-1} $\pm 2\%$. The Hayes fluxes are based entirely on copper melting-point blackbodies. Averaging with the corrected results of Oke & Schild produced a best estimate for F(5556) of 3.39×10^{-9} erg cm^{-2} s^{-1} Å^{-1} $\pm 2\%$. While Hayes & Latham present fluxes for the full 3300–10800 Å range, the value at 5556 Å is especially important for the WD flux system discussed in § 4.

Hayes (1985; hereafter H85) reviewed the available flux measurements for Vega and compiled a recommended SED from 3300–10500 Å, including a revised estimate of the monochromatic flux with a 25 Å bandpass at 5556 Å of 3.44×10^{-9} erg cm^{-2} s^{-1} Å^{-1} $\pm 1.5\%$ from an average of five independent measures. See Figure 6 of Bohlin et al. (2011; hereafter B11) for a graphical comparison of the five F(5556) flux values with the H85 SED for Vega. The ratio of this H85 SED to the STIS CALSPEC fluxes is illustrated in Figure 7. Any measured SED based on the H85 Vega fluxes, such as BD + 17°4708 (Fukugita 1996, Bohlin & Gilliland 2004b), differs from *HST*-based fluxes by more than 1% over much of the wavelength range as in Figure 7.

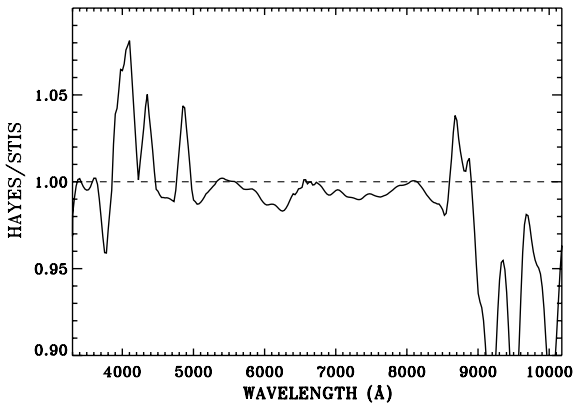

FIG. 7.—Smoothed ratio for Vega of the H85 to the baseline CALSPEC fluxes at the 25 Å bin spacing of H85.

Ten years after H85, Megessier (1995) reviewed the available absolute F(5556) determinations and eliminated the low Oke & Schild value because evidence suggested that result was based on a faulty tungsten ribbon lamp. Another set of results was scrapped because of discrepancies in the lab calibrations of their tungsten lamps. A weighted average of the three remaining measurements yielded 3.46×10^{-9} erg cm^{-2} s^{-1} Å^{-1} $\pm 0.7\%$ for F(5556). The rms scatter of 0.7% is lower than the H85 1.5%, because outliers were eliminated.

3.2.2. Solar Analogs

The solar analog method of establishing primary stellar flux standards is to assign the measured solar flux distribution to a star of the same G2V spectral type, where the assumption is that the solar SED is known to better precision than any star. One commonly adopted solar flux distribution is Thuillier et al. (2003, Th03), where observations from the ATLAS and EURECA missions with the space shuttle are referenced to the Heidelberg Observatory blackbody absolute-flux standard. Figure 8 compares existing CALSPEC flux distributions for solar analogs (Bohlin 2010; hereafter B10) with the Th03 SED. If the solar twins, 18 Sco (G2V) and HD101364 (alias HIP56948; Melendez & Ramirez [2007]) are ever observed with STIS in a photometric slit, those new data should be compared to Figure 8.

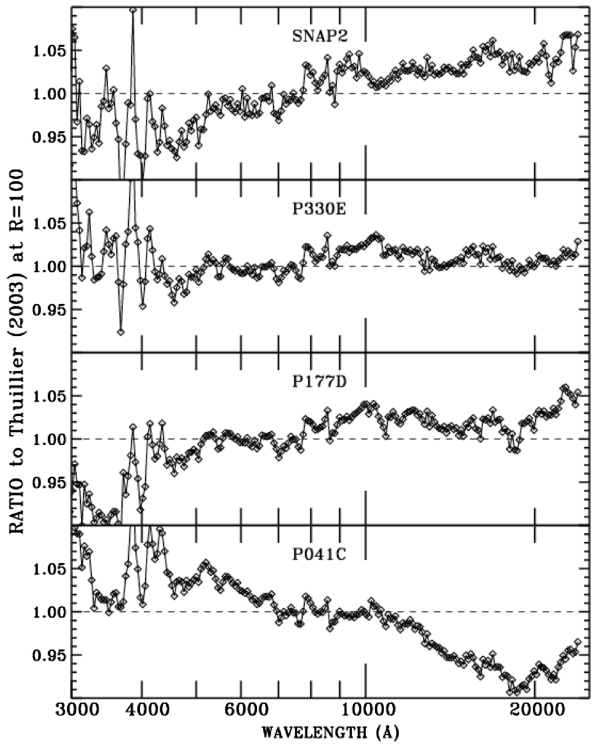

FIG. 8.—Ratio at $R = 100$ of CALSPEC fluxes for solar type stars to the Th03 solar SED, as normalized to unity at 7000–8000 Å.

While consistent to ~10% in Figure 7, these G stars do not agree with each other or with the solar flux to 1%. The dereddening for the small $E(B-V)$ interstellar extinction values of B10 is not done, but does not significantly flatten or improve the consistency of the ratios of Figure 7. Perhaps a 1% agreement with the Sun could be achieved with a true solar twin.

At present, the solar analog technique has a lower precision than the *HST*-adopted WD technique, even if an unreddened star is a perfect match to the solar line spectrum. The Th03 uncertainty at 2 μm is quoted as 1.3% in table III of Th03 but grows to "~2%" in the conclusions section, which also quotes a 2–3% uncertainty below 0.85 μm. The *HST* CALSPEC stellar SEDs have slightly smaller formal uncertainties over the UV/optical/NIR wavelength range. In the past, the brightness of the Sun facilitated accurate measurements, especially when bolometers were the primary detector. However, with modern sensitive detectors, comparison of laboratory flux standards with stars can be as precise as for the Sun. In space, the solar brightness is actually a disadvantage, because the large UV flux causes contamination of the optical components from polymerized hydrocarbons.

3.2.3. Modern

Fundamental standards of irradience, i.e., flux, are maintained by NIST and are traceable to a gold melting-point blackbody light source. Current NIST 2σ uncertainties in the absolute responsivity of standard detectors are 0.2% for Si photodiodes below 1 μm (Brown et al. 2006) and 0.5% for NIR photodiodes. Standard light sources with radiance temperatures as high as 3000 K are available with an accuracy of 0.55% at 5556 Å (Fraser et al. 2007).

One currently funded program to establish primary stellar flux standards relative to laboratory NIST irradiance standards is the Absolute Color Calibration Experiment for Standard Stars (ACCESS) rocket program, which will establish a few standards in the brightness range of Sirius to $V \sim 9.5$ (Kaiser et al. 2007, 2010a, 2010b, 2012, 2013). The wavelength coverage is 0.35–1.7 μm with an accuracy goal of 1% and a spectral resolving power of R = 500. Even though this wavelength range is accessible from the ground, observations from above the atmosphere eliminates that dominant source of uncertainty. ACCESS will be calibrated to both continuum and emission line fundamental radiance standards. Using emission from tunable lasers, the NIST Spectral Irradiance and Radiance Responsivity Calibrations Using Uniform Sources (SIRCUS) facility (Brown et al. 2006) will provide an end-to-end calibration transfer to ACCESS. These data also define the correction for out-of-band spectral stray light using a matrix correction algorithm (Zong et al. 2006, Smith et al. 2009). Stray light can cause serious errors when measuring stellar spectral distributions that differ from the spectral distribution of the calibrating light source. The use of a spectral light engine will calibrate ACCESS using a continuum spectral energy distribution similar to the spectral energy distribution of the various stellar targets (Brown et al. 2006; Smith et al. 2009).

Ground-based programs to measure stellar absolute fluxes seek to measure both the instrumental response function and the optical transmission function of the atmosphere (Stubbs & Tonry 2012; Tonry et al. 2012). NIST-calibrated photodiodes are utilized to measure the input to a telescope and establish the total throughput quantum efficiency (QE). Empirical atmospheric transmission is determined by water vapor as measured and matched to a MODerate resolution atmospheric TRANsmission (MODTRAN) atmospheric transmission model to determine the real-time atmospheric extinction. The ground-based program, NIST STARS, uses lidar backscatter to measure the atmospheric extinction at selected wavelengths and then fits a MODTRAN model to determine the atmospheric extinction (McGraw et al. 2010; Zimmer et al. 2010). In principle, these techniques applied at high temporal cadence can enable repeatable photometric stellar observation, even while atmospheric clarity is unstable. A goal of NIST STARS is to measure the absolute flux of stars across the sky with a 0.5% precision.

3.3. Infrared above 1 μm

Excellent reviews of the direct measurement of the physical fluxes of stars in the infrared have been published (Rieke et al. 1985; Price 2004; Rieke et al. 2008) and are summarized here.

3.3.1. Ground-based

The direct measurement of stars in the infrared is complicated by the atmospheric transmission that ranges from mostly transparent to totally opaque over the 1–40 μm wavelength range, where the water vapor and OH lines are especially problematic. Ground-based observations necessarily concentrate on measurements at wavelengths where the atmospheric opacity is low.

The earliest research measured absolute fluxes for a number of stars in near-infrared bands, roughly corresponding to z, J, H, and K, by referencing to blackbody sources at the telescopes (Walker 1969). This effort was extended to midinfrared wavelengths for a smaller sample of stars using Mars to transfer the stellar measurements to laboratory blackbodies (Low et al. 1973; Becklin et al. 1973; Rieke et al. 1985). Mainly because of the rapid variability of the atmospheric transmission, the precision of ground-based IR photometry is limited by the lack of accurate throughput measures of the instrumental bandpass that relates photons above the atmosphere to detected quantum events as a function of wavelength. Absolute spectrophotometry is always preferred to photometic average fluxes over a filter bandpass, but to our knowledge, no IR spectrophotometry beyond 1 μm has been published with a pedigree that is directly based on laboratory standards of absolute flux.

A significant body of work concentrated on precision measurements of Vega using the same technique (Selby et al. 1980, 1983; Blackwell et al. 1983; Mountain et al. 1985; Booth et al. 1989) with later works providing growing evidence that the infrared measurements of Vega deviate from extrapolations of optical fluxes into the IR. Detailed observations and modeling have shown that Vega is a rapidly rotating star, observed pole-on, and has a circumstellar disk that contributes to the stellar flux starting around 2 μm and extending to far-infrared wavelengths (Aumann et al. 1984; Su et al. 2005, 2013; Aufdenberg et al. 2006; Sibthorpe et al. 2010; Defrere et al. 2011; Absil et al. 2013; Bohlin 2014).

However, comparisons with modern results suggest that the early flux measurements are often correct within their quoted uncertainties. Figure 9 shows the results of Bohlin (2014) for the Vega photosphere and total flux of the dust plus photosphere in comparison to pioneering ground-based measures of total absolute flux and *Midcourse Space Experiment* (*MSX*) results (Price et al. 2004; see next section). For example, Booth et al. (1989) quote 3.86×10^{-11} erg cm^{-2} s^{-1} Å^{-1} $\pm 4\%$ at 2.250 μm with a bandpass of 10 Å for Vega. The corresponding CALSPEC monochromatic flux from the photosphere (i.e., alpha_lyr_stis_007.fits) is 3.676×10^{-11} erg cm^{-2} s^{-1} Å^{-1}, while the contribution from the dust in the K band is 1.26% (Absil et al. 2013) for a total of 3.72×10^{-11} erg cm^{-2} s^{-1} Å^{-1}. Thus, the Booth value is only 3.7% higher than the modern estimate and is within the Booth uncertainty of 4%. The Selby et al. (1983) measure of 3.92×10^{-11} erg cm^{-2} s^{-1} Å^{-1} at 2.20 μm is $\sim$3% lower than Bohlin (2014) but agrees within the quoted undertainty of 4%.

3.3.2. Airborne and Space-based

The challenges associated with observing through the atmosphere have motivated efforts to make direct measurements of stars in the infrared from airplanes and from space-based telescopes. Witteborn et al. (1999) used the Kuiper Airborne Observatory to provide direct measurements of a single star, α Boo,

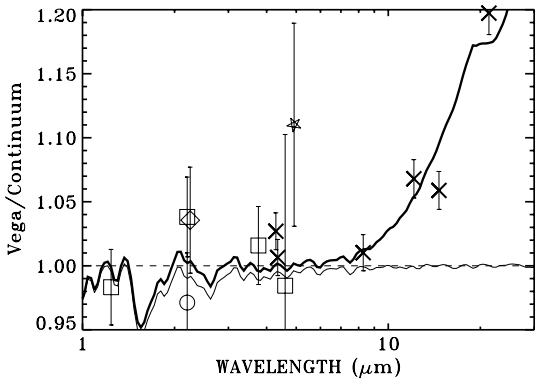

FIG. 9.—Comparison of historic IR absolute flux measures with the Vega SED of Bohlin (2014) at a resolution of $R = 20$ for the photospheric 9400 K model (*light line*) and the total flux including emission from the dust ring (*heavy line*). All values are divided by the theoretical photospheric continuum for clarity of display. The various symbols for the physics-based measurements are *circle* (Selby et al. 1983), *diamond* (Booth et al. 1989), *star* (Mountain et al. 1985), *squares* (Blackwell et al. 1983), and X (*MSX*).

which demonstrates the challenges of measuring the absolute flux of stars using observations through the atmosphere with the often complicated transfer to laboratory blackbodies.

A space-based experiment with onboard calibration sources significantly simplifies the direct measurement of stellar fluxes, which motivated the Spatial Infrared Imaging Telescope (SPIRIT) III on the *MSX* (Mill et al. 1994). The *MSX* SPIRIT III observations covered the 4.3 to 21 μm range in six bands and calibrated the observations of stars to five emissive reference spheres that were ejected at various times during the 8 months of operations. The properties of these reference spheres were measured in the laboratory, and their absolute infrared fluxes were determined as their temperatures increased due to illumination by the Sun and Earth shine. The *MSX* reference sphere calibration was checked by comparison to a set of predictions for the fluxes of eight reference stars from the CWW network (Cohen et al. 1992b; Cohen 2007). The *MSX*-measured mid-infrared fluxes have a quoted accuracy of 1.4% Price et al. (2004). Thus, the *MSX* mid-infrared measurements of the Galactic plane and other selected areas provide a network of stars that have been directly referenced to laboratory standards with fluxes calculated using basic physics.

3.4. Summary of Absolute Flux Zeropoint

Bohlin (2014) observed the primary IR standard, Sirius, with STIS and normalized an updated special Kurucz model[6] to the observed fluxes at 6800–7700 Å. The modeled IR extrapolation and the absolute Meggessier visible flux are reconciled with the *MSX* mid-IR fluxes (Price et al. 2004) at 8–21 μm. In order to minimize both the 5556 Å and *MSX* mid-IR residuals, the Megessier (1995) value with its 0.7% quoted uncertainty must be multiplied by 0.9945 to achieve a properly weighted average. The largest residual is for the *MSX* A band at 8.3 μm, where the residual is 1.8%, i.e., 1.3σ. The final result from Bohlin (2014) is $\mathrm{F}(5556) = 3.44 \times 10^{-9}$ erg cm^{-2} s^{-1} Å^{-1} with a formal uncertainty of 0.5%. While NLTE models establish the relative SEDs of the three primary *HST* standards as a function of wavelength, their overall absolute flux level is set by F(5556) for Vega, as discussed in the following.

4. USING MODELS TO ESTABLISH SEDS

4.1. History and Rationale

Despite valiant efforts to tie stars to lab flux standards in the 1970s, inconsistencies in the available standard stars still existed, as illustrated in some of the preceding figures. Thus, D. Finley and J. Holberg (Finley et al. 1984; Holberg et al. 1986; Finley et al. 1990) suggested the use of pure hydrogen WD model atmospheres for the UV flux calibration of the *IUE* satellite (Bohlin et al. 1990). Pure hydrogen WDs are preferred, because model atmosphere calculations are greatly simplified with only one element to consider. As the basis for all *HST* absolute fluxes, Bohlin, Colina, & Finley (1995) adopted the D. Koester model atmosphere SEDs calculated in local thermodynamic equilibrium (LTE) for G191B2B, GD153, GD71, and HZ43. Subsequently, HZ43 fell off this list of primary flux standards because of an M star companion that contaminates the STIS observations in the visible and IR (Bohlin et al. 2001). For the remaining three stars, interstellar reddening from the dust reduces the flux by $<0.6\%$ longward of 1150 Å, given the strict limits on $E(B-V)$ derived from the low hydrogen column densities and the Galactic average $N(\mathrm{HI})/E(B-V) = 4.8 \times 10^{21}$ of Bohlin et al. (1978). Of our three stars, G191B2B has the largest hydrogen column density $N(\mathrm{HI}) = 2.2 \times 10^{18}$ (Rauch et al. 2013, hereafter RWBK), which corresponds to $E(B-V) = 0.0005$. Dupuis et al. (1995) find $N(\mathrm{HI}) < 1 \times 10^{18}$ for GD153 and GD71. For G191B2B, the small reddening has been applied to the model according to the $R(V) = 3.1$ reddening curve of Cardelli et al. (1989) down to 2000 Å, where the LMC curve of Koornneef and Code (1981) for small grains is substituted, because the steeper far-UV slope fits the data better.

In order to establish the effective temperature and surface gravity (T_{eff} and $\log g$) for the pure hydrogen WDs, the calculated model lines are fit to the observed Balmer line profiles (e.g., Finley et al. 1997; hereafter FKB). FKB used Koester LTE model line profiles to fit the Balmer lines. However, NLTE calculations should be a better representation of the actual stellar physics. Consequently, Bohlin (2003) adopted the Hubeny Tlusty NLTE models (version 203) for pure hydrogen atmospheres (Hubeny & Lanz 1995). However, the T_{eff} and $\log g$ values originally used for the NLTE models were those derived from the LTE fits. Recently, this deficiency has been rectified by Gianninas et al. (2011, G11), who fit new Balmer line observations of the WDs with updated Tlusty NLTE models that include improved Stark broadening of the Balmer lines (Tremblay & Bergeron 2009).

4.2. The *HST* Absolute Fluxes

4.2.1. The WD Models

The absolute flux at 5556 Å (5557.5 Å in vacuum) of the NLTE models for the three primary WD standards is set by the STIS spectrophotometry of Vega (Bohlin & Gilliland 2004a; Bohlin 2007) relative to the WDs and by the absolute monochromatic flux at 5556 Å for Vega of $\mathrm{F}(5556) = 3.44 \times 10^{-9}$ erg cm^{-2} s^{-1} Å^{-1}, as discussed above. The small uncertainty of 0.5% in this F(5556) value affects the overall level and not the shape (i.e., "color") of the WD models used for *HST* flux calibrations. Despite suggestions that Vega is a variable star, H85 discusses the evidence and concludes that any variability "must be less than 0.01 mag". Engelke et al. (2010) present evidence for a 0.08 mag variation of Vega at visible

[6] http://kurucz.harvard.edu/stars/SIRIUS/.

wavelengths; but Bohlin (2014) demonstrated that this apparent variability seen in *Hipparcos* data is actually just a symptom of pulse counting saturation for this bright star.

Once the three WD flux distributions are fully defined by their NLTE models as normalized to the reconciled absolute visible/IR level, the STIS flux calibration proceeds as outlined in § 2. The fully corrected N_e count rate in electrons per second as a function of wavelength for each STIS observation of the three primaries is extracted from the STIS images using an extraction height of 11 pixels for the UV MAMA data and 7 pixels for the CCD images. The procedure for correction to infinite extraction height is outlined in Bohlin (1998). Each N_e spectrum is matched to the standard star flux, F_λ, with wavelengths adjusted to the instrumental rest frame and with the model smoothed to the STIS resolution. Because the STIS LSF is not precisely known and the model line profiles are not perfect, the absorption line regions are masked before the sensitivities $S_\lambda = F_\lambda/N_e$ per equation (15) are fit with splines. The *IDL* procedure *splinefit* is used for fitting the 1024 pixel sensitivities with 50–60 nodes per each of the five STIS low dispersion modes. The spline fits for all three WDs are averaged with equal weight for each star to get the STIS sensitivities that are used to calibrate other STIS observations and establish secondary flux standards. Because G191B2B is too bright for the two UV MAMA modes, those sensitivity functions are defined only by GD71 and GD153. All STIS observations of flux standards utilize the wide 52X2″ slit to avoid variable slit losses.

Figure 10 illustrates the internal residuals after the old STIS low-dispersion fluxes of the WDs are divided by the Tlusty 203 pure-hydrogen NLTE models used to define the old flux calibration. The agreement at the subpercent level demonstrates that the calibrated fluxes agree with their reference SEDs to better than ∼0.5% over the 1150–10000 Å wavelength range, except for a few narrow bands at absorption lines. Thus, any updates to the modeled SEDs of these primary standard should retain the same subpercent level of internal agreement.

Figure 11 illustrates the difference between the original Tlusty 203 and the new Rauch NLTE models with $T_{\rm eff}$ and log g from G11 for GD153 and GD71. Figure 11 shows that the Rauch models from the registered virtual observatory service TheoSSA[7] (Werner et al. 2003) change by similar amounts over 1150–10000 Å for each star relative to the old Tlusty SEDs. The pure hydrogen model for G191B2B does *not* show the same ratio as the two cooler stars, but that model is inappropriate because of the trace metal lines observed in the UV. Instead, a special line blanketed NLTE model of RWBK is compared with the Hubeny SED for G191B2B in Figure 11. While RWBK found $T_{\rm eff} = 60000 \pm 2000$ K, a $T_{\rm eff} = 59000$ K model is within the uncertainty and is more consistent with the relative UV flux of the three stars. The parameters defining the three

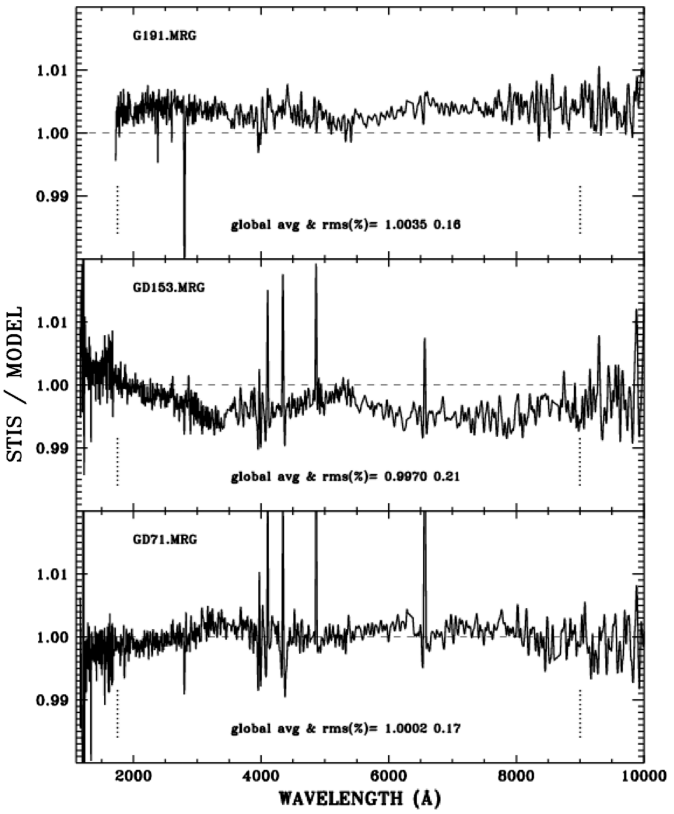

FIG. 10.—Ratio of old STIS flux distributions to the old Tlusty 203 NLTE models for pure hydrogen that were used to define the old fluxes. The average ratios and rms scatter between the vertical dotted lines at 1750 and 9000 Å are written on the plots. Narrow band differences are evident in the Balmer lines. G191B2B has a strong MgII interstellar or circumstellar absorption feature at 2800 Å. There are no data for G191B2B below 1700 Å, because the star is too bright for the STIS MAMA detectors.

primary flux standards appear in Table 1. After a complete re-calibration (see section 2) using the new set of model SEDs, Figure 12 illustrates the new internal STIS residuals, which are comparable to the old residuals in Figure 10. The net changes in STIS fluxes due to this update of the three prime standards are $\lesssim 1\%$. Updated fluxes that are based on the new Rauch model SEDs are identified in the CALSPEC database by delivery dates in 2013 or later.

4.2.2. The Uncertainties

The uncertainty in the overall level of the absolute fluxes is the 0.5% for the $\mathrm{F}(5556) = 3.44 \times 10^{-9}$ erg cm^{-2} s^{-1} Å^{-1} flux of Vega. The possible error in the WD SEDs could be larger, if Vega is variable or if STIS is nonlinear over the large dynamic range of 12–13 mag for the measured ratios of the WDs to Vega at 5556 Å. While the variability of Vega remains controversial, there is no charge transfer efficiency (CTE) loss in the STIS CCD observations for such a bright star according to the CTE correction formula of Goudfrooij et al. (2006).

[7] Theoretical Stellar Spectra Access, http://dc.g-vo.org/theossa.

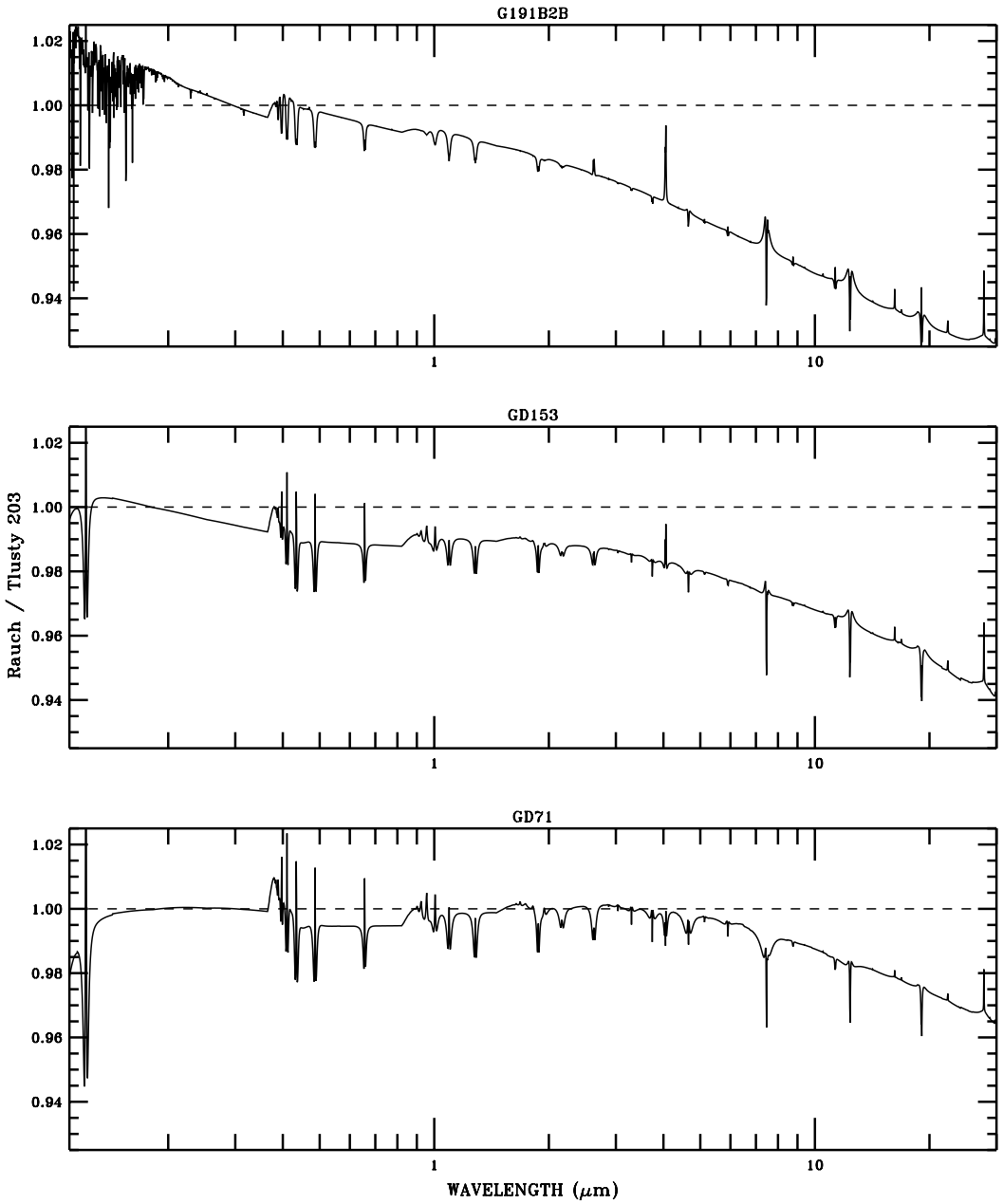

FIG. 11.—Ratio at $R = 500$ of the new Rauch model fluxes to the pure-hydrogen Tlusty 203 models that previously defined the three primary WD SEDs. All models are computed in NLTE with the parameters in Table 1. For GD153 and GD71, the new models are for pure hydrogen, while RWBK have computed a full metal line blanketed model of G191B2B that matches high-dispersion observations of the UV absorption lines. From Lyα to 1 μm over the STIS wavelength range, the three ratio plots are the same to ~2% in the continuum. Narrow band differences are evident in the hydrogen line profiles. The Rauch models include the new 0.6% gray reduction of the fluxes with respect to the old Tlusty normalization (Bohlin 2014).

More important is the uncertainty in the slopes of the adopted SEDs, i.e., the ratio of the model fluxes to their flux at 5556 Å. One measure of this uncertainty in the new set of primary standard models is the difference between sets of pure hydrogen NLTE models with the same $T_{\rm eff}$ and $\log g$. Our Tlusty NLTE models (version 204) include the Tremblay & Bergeron (2009) Stark profiles (see also Tremblay et al. 2011, G11), and incorporate similar physics to the Rauch models. The model ratios appear in Figure 13 for the G11 determinations of $T_{\rm eff}$ and $\log g$. In the STIS range, the Tlusty 204 models are systematically higher by up to 3% in the far-UV at Ly-α for GD71, while the IR fluxes are also higher, with a worst-case difference of nearly 5% at 30 μm for G191B2B. The pairs of models agree to ~1% from 0.2–5 μm.

TABLE 1
PRIMARY WD STARS

Star	V[a]	Sp. T.	FKB $T_{\rm eff}$	FKB $\log g$	G11 $T_{\rm eff}$	G11 $\log g$	Unc. $T_{\rm eff}$
G191B2B	11.781	DA.8	61193	7.492	60920[b]	7.55	993
GD153	13.346	DA1.2	38686	7.662	40320	7.93	626
GD71	13.032	DA1.5	32747	7.683	33590	7.93	483

[a] G191B2B–Landolt and Uomoto (2007), GD153–Landolt (1995, private comm.), GD71–Landolt (1992)
[b] The $T_{\rm eff}$ and gravity of the best fitting metal line-blanketed model are 59000 K and $\log g = 7.6$.

To understand the differences between the Rauch (TMAP) and Tlusty NLTE results, the actual computer codes should be compared in detail. However in the IR, the LTE versus NLTE differences (see B11) are significantly larger than the differences between the NLTE codes, which implies that IR fluxes at the few percent level are rather sensitive to the input microphysics and to the NLTE stratification of the upper layers of the model atmospheres.

Another smaller contributor to the uncertainty in the model slope is caused by the $T_{\rm eff}$ measurement errors from G11; but the uncertainties of 0.05 in $\log g$ cause, for example, at most a 0.2% continuum flux difference for G191B2B and are neglected. The dotted line in Figure 14 represents the rms differences from unity for the ratios of nominal temperature models to models that differ by the uncertainty in degrees Kelvin from the last column of Table 1. The dashed line in Figure 14 is the rms average difference from unity for the three ratios of Figure 13 and represents the uncertainty in the modeling procedure, while the heavy solid line is the combined estimate of total uncertainty in the WD flux system. Because the model SEDs are all normalized to unity at 5556 Å, the uncertainties are all relative to 5556 Å, where the uncertainty in flux due to slope errors

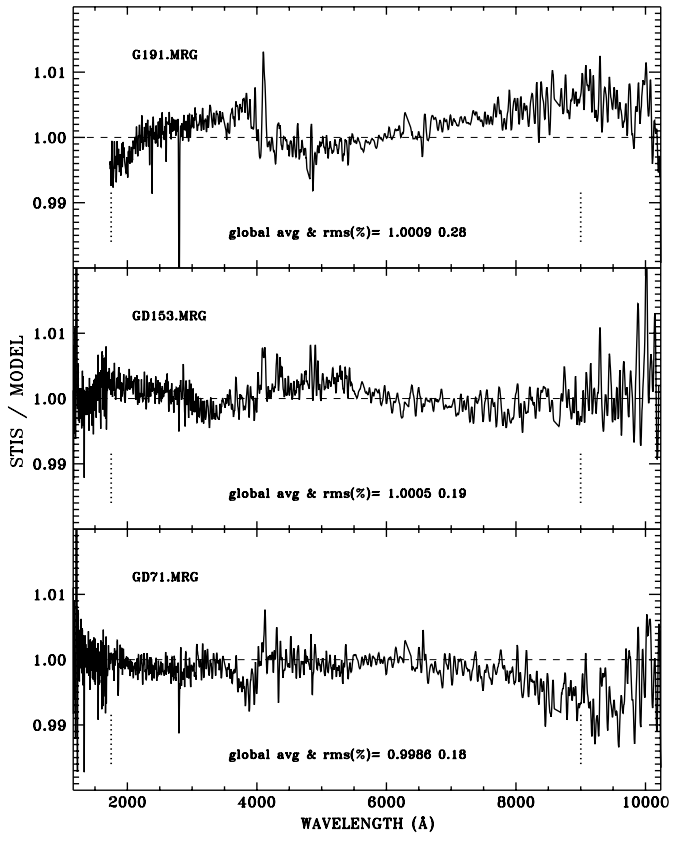

FIG. 12.—Ratio of new STIS spectrophotometry as in Fig. 10 to the new primary standards made from the Rauch models in Fig. 11. Notice the smaller residuals in the Balmer lines for GD153 and GD71 in comparison with Fig. 10.

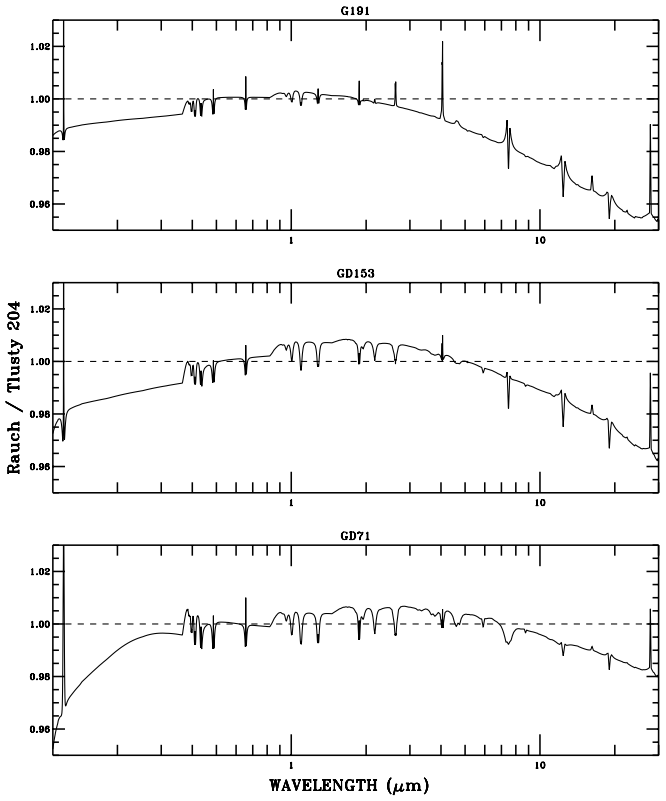

FIG. 13.—Ratios of pure hydrogen models for the results from the Tlusty 204 and Rauch NLTE codes, both using the Stark broadening profiles of Tremblay & Bergeron 2009. The resolution is $R = 500$ and both models for each star are normalized to the same value at 5566 Å. The deviation from unity for the three stars provides a measure of the uncertainty of using WD models to represent the SEDs of actual stars, and the worst agreement at the shorter wavelengths is for the coolest star GD71.

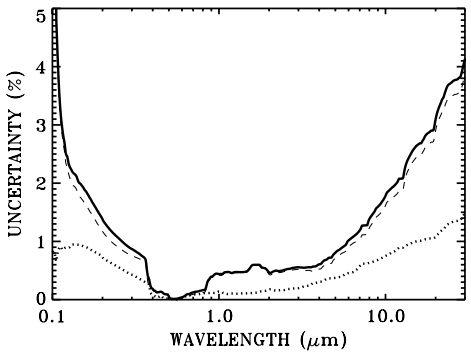

FIG. 14.—*Dotted line*: rms uncertainty in the WD flux scale from formal errors in $T_{\rm eff}$ from G11. *Dashed line*: Dominant rms uncertainty, as calculated from differences between pure-hydrogen Tlusty 204 and Rauch NLTE models with the same $T_{\rm eff}$ and $\log g$. *Heavy solid line*: Combination in quadrature of the above two uncertainties. These curves are all relative to the flux at 5556 Å, where the relative uncertainty is zero by definition.

is defined to be zero. When the primary stars are used for an instrumental calibration, the strong hydrogen lines are masked to avoid errors due to imprecisely known instrumental line-spread functions (LSF). Also, the modeling precision in the line cores and wings is less than the precision in the continuum. Thus, the effects of the lines are also removed in Figure 14.

While Figure 14 suffices for simple-minded estimates of the ensemble systematic uncertainty of the *HST* flux system, correlations of the errors over the wavelength range require the more comprehensive covariance matrix to fully characterize the uncertainties (Jeach 1985). This 573×573 matrix is available from the CALSPEC archive as the binary table WDcovar.fits in units of fractional error. The 573 wavelength bins represent a resolution of $R = 100$ from 1000 Å to 30 μm, while the square of the heavy solid line in Figure 14 is the diagonal of the covariance matrix after dividing the values from the figure by 100 to get the fractional error.

A set of WD models for the primary standards provides complete wavelength coverage for a system of absolute fluxes and for the zero points of the various standard magnitude systems; see for example Holberg & Bergeron (2006) and Pickles (2010). However, differences reach 2% at 8 μm in Figure 13 between the independent G191B2B models, which suggests that, pending further investigation, these WD SEDs are not appropriate *James Webb Space Telescope* (*JWST*) standards out to 30 μm.

In summary, the likely systematic errors in the slope of our WD-based flux system is the heavy solid line in Figure 14 or the covariance matrix WDcovar.fits. In order to get an estimate of the total systematic uncertainty at any wavelength, the uncertainty relative to 5556 Å must be combined with the uncertainty in F(5556) for Vega, which is 0.5%. Since the covariance is a measure of the uncertainty σ_{ij}^2, a value of $0.005^2 = 0.000025$ must be added to every matrix element to convert the covariance from values relative to 5556 Å to total σ^2. Our ensemble uncertainty estimates are applicable to instrumental flux measures when observations with equal weight for all three primary WDs are utilized for the flux calibration. For the best transfer of the WD flux system to an instrumental calibration, the observational data set for the three primary WDs should be robust enough to make statistical errors negligible, while spacing of the observations over the instrumental lifetime serves to reduce errors in accounting for any time dependent effects.

Of course, the total uncertainty in the observed flux for a sparsely observed program star is the uncertainty in the particular observational data set combined with the ensemble systematic uncertainty. For example, the signal-to-noise, the broadband repeatability, and any nonlinearities contribute to larger errors in the flux of any program star. See Figure 1 of Bohlin & Gilliland (2004b) for a discussion of statistical uncertainty for STIS.

The new models for the *HST* WD standard stars helps explain part of the problem with the WD fluxes in comparison with the *Spitzer* fluxes (B11). The worst discrepancy found by B11 was a 4σ difference of 12% for G191B2B with IRAC4 at 8 μm. The 4% lower flux at 8 μm for the new G191B2B SED in Figure 11, reduces the discrepancy to 8% with less than a 3σ significance.

4.3. IR Fluxes of Normal Stars

Because of the uncertainty in the IR fluxes of the WDs and because of the sparsity of bright pure hydrogen WDs, an alternate method is required for establishing a network of standard stars with known IR absolute fluxes. Longward of 1 μm, an extensive network of standard stars based on A-star models of Vega and Sirius has been established in a sequence of papers I–XIV (e.g., Cohen et al. 1992a, 2003). The essense of this technique is to find a model from a published grid,[8] e.g., Castelli & Kurucz (2003, hereafter CK04) that fits the observed shorter wavelength fluxes so that the longer wavelength part of the model establishes the absolute IR fluxes of the standard star. This technique was used by Bohlin & Cohen (2008) for A stars and by B10 for G stars with a fitting technique in four parameters, $T_{\rm eff}$, $\log g$, the metallicity [M/H], and the reddening $E(B-V)$. With the switch to Rauch NLTE models to define the primary WD SEDs, the measured *HST* fluxes change slightly; and new fits are required. Table 2 includes the revised best fit parameters for the Bohlin & Cohen and B10 stars that now have both STIS and NICMOS spectrophotometry. Additional new O, B, A, and G standards have only STIS spectrophotometry. The fitting technique is modified to use χ^2 minimization rather than the previous minimization of the rms residuals. These stars are prime candidate *JWST* flux standards; and their SEDs consisting of STIS fluxes plus the modeled IR extensions are in the CALSPEC database.

Despite changes in the SEDs of the three prime reference standards, improved observational data for the model fitting,

[8] http://wwwuser.oats.inaf.it/castelli/grids.html.

TABLE 2
SECONDARY STANDARDS WITH IR FLUXES DEFINED BY CK04 MODELS

Star[a]	R.A. J2000	Decl. J2000	Sp.T.	V	$T_{\rm eff}$	log g	[M/H]	$E(B-V)$	χ^2
ξ^2 Cet	02 28 09.54	+08 27 36.2	B9III	4.30	10360	3.90	−0.61	0.001	1.35
λ Lep	05 19 34.52	−13 10 36.4	B0.5V	4.29	27920	4.30	0.10	0.017	5.58
μ Col	05 45 59.89	−32 18 23.2	O9.5V	5.18	29840	4.00	0.12	0.003	7.00
10 Lac	22 39 15.68	+39 03 01.0	O9V	4.88	29880	3.80	0.03	0.069	7.67
HD014943	02 22 54.68	−51 05 31.7	A5V	5.91	8040	3.90	−0.10	0.037	3.05
HD37725	05 41 54.37	+29 17 50.9	A3V	8.35	8140	4.15	−0.32	0.024	1.93
HD116405	13 22 45.12	+44 42 53.9	A0V	8.34	10800	4.10	−0.37	0.002	0.76
BD+60 1753	17 24 52.27	+60 25 50.7	A1V	9.67	9440	4.00	−0.12	0.023	0.49
HD158485	17 26 04.84	+58 39 06.8	A4V	6.50	8700	4.30	−0.38	0.066	2.56
1732526	17 32 52.64	+71 04 43.2	A3V	12.53	8860	4.10	−0.20	0.061	2.64
1743045	17 43 04.48	+66 55 01.6	A5V	13.6	7350	3.50	−0.47	0.014	0.71
HD163466	17 52 25.37	+60 23 46.9	A2	6.86	7880	3.60	−0.56	0.030	4.77
1757132	17 57 13.24	+67 03 40.8	A3V	12.01	7860	4.10	0.34	0.071	0.88
1802271	18 02 27.17	+60 43 35.7	A3V	11.98	9070	4.10	−0.58	0.024	0.80
1805292	18 05 29.28	+64 27 52.0	A1V	12.28	8540	4.00	−0.13	0.033	0.50
1808347	18 08 34.70	+69 27 28.7	A3V	11.69	7900	3.90	−0.73	0.027	1.32
1812095	18 12 9.60	+63 29 42.2	A5V	12.01	7750	3.60	0.06	0.002	1.34
HD180609	19 12 47.20	+64 10 37.2	A0V	9.41	8560	4.00	−0.55	0.042	0.65
C26202	3 32 32.84	−27 51 48.6	F8IV	16.64	6200	4.40	−0.52	0.053	0.67
HD037962	05 40 51.97	−31 21 04.0	G2V	7.85	6000	5.00	0.00	0.059	1.87
HD038949	05 48 20.06	−24 27 49.9	G1V	8.0	6080	4.20	−0.12	0.016	1.11
HD106252	12 13 29.51	+10 02 29.9	G0	7.36	5940	4.70	0.00	0.016	1.40
P041C[b]	14 51 57.98	+71 43 17.4	G0V	12.16	6020	4.15	0.07	0.034	0.79
P177D	15 59 13.57	+47 36 41.9	G0V	13.36	5880	3.80	−0.11	0.052	0.98
SF1615+001A	16 18 14.23	+00 00 08.6	G0-5	16.75	5880	4.30	−0.73	0.118	0.45
SNAP-2	16 19 46.11	+55 34 17.8	G0-5	16.23	5760	4.90	−0.36	0.034	0.74
P330E	16 31 33.82	+30 08 46.5	G2V	12.92	5920	4.80	−0.13	0.051	1.59
HD159222	17 32 00.99	+34 16 16.1	G1V	6.56	5780	3.90	0.00	0.001	1.80
HD205905	21 39 10.15	−27 18 23.7	G2V	6.74	5920	4.10	0.00	0.025	1.91
HD209458	22 03 10.77	+18 53 03.5	G0V	7.63	6100	4.20	−0.04	0.003	0.53

[a] The following stars with data are omitted: HD27835 and HD60753 (double stars); HD165459 (dust ring); 1739431 (bad focus); 1740346 (dust ring); 1812524 (poor fit to model atmosphere).
[b] P041C has an M companion 0.57″ away (Gilliland & Rajan 2011)

and improved fitting techniques, the largest change in the modeled IR fluxes of Bohlin & Cohen or B10 is 3% at 30 μm for 1812095, which corresponds to a change in $T_{\rm eff}$ from 8250 K to 7750 K. Because the IR fluxes for G to A stars approach a Rayleigh–Jeans distribution, the fitted fluxes are robust and not a strong function of temperature. The reason for such large changes for 1812095 is that Bohlin & Cohen fitted only to the NICMOS spectrophotometry, which has a short wavelength cutoff of 0.8 μm, while the new fit is constrained by new STIS data down to 0.114 μm. For the G stars of B10, the original IR flux extrapolations differ typically by $<1\%$ from the revised parametizations in Table 2, because STIS data already existed for the B10 analysis.

Just as for the three primary WDs, the systematic errors of calculated model grids likely dominate the uncertainty in the extrapolated IR fluxes. For the G stars, B10 compared results for MARCS models (Gustafsson et al. 2008) with the CK04 grid and found models from both grids that agree with the STIS fluxes to $\sim 0.5\%$ at shorter wavelengths, but the MARCS grid does not extend past 20 μm, while the CK04 grid has rather coarse wavelength spacing with a total of just 1221 sample points and only one point between 10 and 40 μm. Recently, Mészáros et al. (2012) have computed an expanded set of the Kurucz ATLAS9 models with updated abundances. Unfortunately, the wavelength grid is even sparser with only 333 points. A third independent model grid, an update of the CK04 models with good wavelength resolution, and an extension of the MARCS grid beyond its current 20 μm limit are required to establish confidence in the IR flux distributions. Shortward of 20 μm, the best-fit CK04 models agree with the best-fit MARCS models to 1% for the G stars in broad continuum bands. Combining this 1% with a 1% systematic uncertainty in the *HST* WD flux scale at the anchor point of the fits at 1–2 μm from Figure 14 results in an estimate of a possible systematic 2% uncertainty of the broadband IR fluxes of individual G stars near 20 μm with respect to the V band. At 20 μm, this 2% uncertainty is less than the 3% uncertainty for WDs from Figure 14.

For the A stars, a second independent model grid is needed, because the MARCS models are limited to maximum

temperatures of 8000 K. Both the original Cohen SEDs and the newer Bohlin & Cohen (2008) SEDs are extrapolations of the measured fluxes into the IR using models that are based on computer code which is traceable to R. Kurucz. The best check on these IR fluxes of the A stars was from the two models provided by T. Lanz at 9400 K and 8020 K. While the agreement of the Lanz SED with CK04 is within 2% for the 9400 K model, the Lanz SED is brighter than CK04 by 4% at 10 μm for the 8020 K model, which raises the question of the accuracy of model SEDs, in general. A new high-resolution grid is currently being produced by Sz. Mészáros and should significantly improve the situation.

Cooler stars, such as K type, are often considered for extrapolating model SEDs to longer wavelengths. However, the extensive contribution of molecular line blanketing complicates the models, which suggests that hotter stars should be superior IR flux standards whenever the SED depends on model calculations.

4.4. Non-WD Models as Standard Star SEDs

In principle, a model atmosphere for any star can establish a primary standard SED, just as for the standard WDs G191B2B, GD153, and GD71 in § 3.2.1 above. However, the determination of precise $T_{\rm eff}$ and $\log g$ from Balmer line profiles is complicated by surface convection for G and F stars (e.g., Ludwig et al. 2009) and by perturbing metal atoms. Modeling stellar SEDs with significant metal line blanketing is a challenging enterprise when the goal is an SED with 1% relative precision over a broad wavelength region. However, several complete SEDs for selected stars are available on the R. Kurucz website,[9] e.g., the Sun, Vega, Sirius, and HD209458. As shown in Figure 15, the agreement of these Kurucz models with the CALSPEC fluxes is impressive. For the A stars in the top two panels, there is one glitch at the Balmer line convergence where the atomic physics is not quite perfected. Vega is a rapid rotator with a pole-to-equator surface temperature gradient (Aufdenberg et al. 2006), and the CALSPEC flux rises above the single $T_{\rm eff} = 9400$ K model below 0.32 μm, because the hotter polar region starts to dominate the UV output. At 0.17 μm, the true flux is ~15% above the 9400 K single temperature model. In the IR, Vega is also a poor standard, because a dust ring contributes significant flux starting near the K band at 2.2 μm, where Absil et al. (2013) find a 1.3% contribution from the ring.

For the G stars in the bottom two panels of Figure 15, the comparison is not shown below 4500 Å, where the heavy line blanketing limits the fidelity of the models (cf., Fig. 10 of Bohlin [2007]). Otherwise, the CALSPEC data for HD209458 agrees with its model to ~2%, while the solar fluxes measured by Th03 deviate by as much as 6%. Thus, Figure 15 is evidence

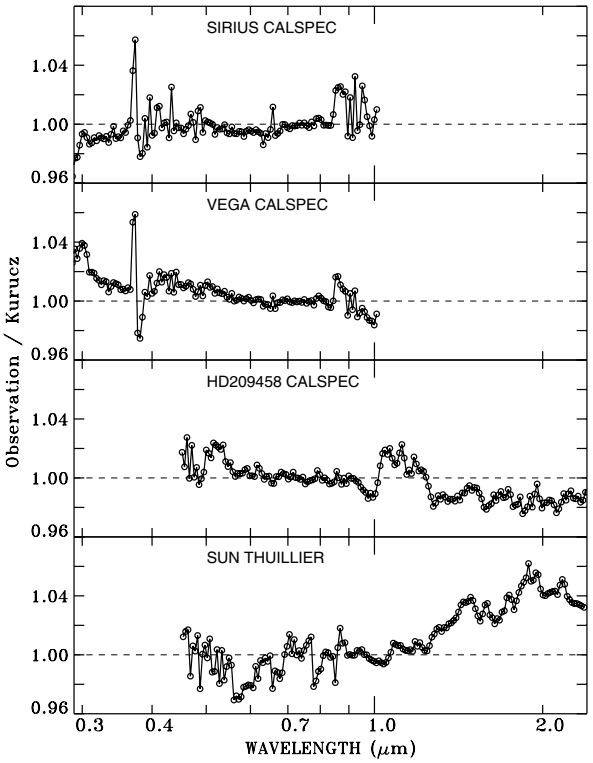

FIG. 15.—Ratio at a resolution of $R = 100$ of measured flux to the specially tailored Kurucz models for four stars. The models are normalized to the data at 0.7–0.8 μm.

that stellar absolute fluxes are known to a better accuracy than the solar absolute flux. Because Sirius is a slow rotator with no dust ring contamination, the Kurucz model is recommended as a primary IR SED, as suggested by Cohen et al. (1992a) and Engelke et al. (2010); see Bohlin (2014).

5. SPECTROPHOTOMETRY ARCHIVES

5.1. CALSPEC/*HST*

Finding charts and coordinates for many of the CALSPEC standard stars appear in Turnshek et al. (1990), while additional coordinates are tabulated by Bohlin et al.(1990) and on the CALSPEC website.

The CALSPEC archive contains the spectral flux distributions of stars with STIS or NICMOS observations, which are sometimes supplemented by older and less reliable data from the *HST* Faint Object Spectrograph (FOS) spectrograph, the *IUE* satellite, or the ground-based observations of Oke (1990). Complete model atmosphere flux distributions are also included for a few stars, while several stars have IR extensions based on models. The *FITS* headers specify the source of fluxes in the various wavelength intervals, while Bohlin et al. (2001) illustrate the relative precision of some of the various sources of SED measurements.

[9] http://kurucz.harvard.edu.

5.2. Pickles Library

Pickles (1998) presents a library[10] of stellar SEDs for 131 normal spectral types. The fluxes are tabulated at 5 Å intervals with a resolution of ~500 and cover 1150–10620 Å with an extension to 25000 Å for about half of the library. The SED for each type is constructed from a variety of sources and from observations of multiple stars of the same spectral type. Figure 16 compares the Pickles SED for A3V at $T_{\rm eff} = 8790$ K with two CALSPEC stars of the same type after dereddening and normalizing to unity at 0.7–0.8 μm, where the line blanketing is minimal. These two stars are the only CALSPEC A3V spectral types with both STIS and NICMOS measurements.

Below 1 μm, the three curves diverge more toward shorter wavelengths, but differences in the effective temperature, metallicity, and reddening cause the greatest changes in flux at far-UV wavelengths. The $T_{\rm eff} = 9070$ K for 1802271 in Table 2 suggests that its spectral type should be somewhat hotter than the Pickles A3 at 8790 K, which explains why the green curve is too high in Figure 16 in the UV. For example, a $T_{\rm eff} = 9070$ K continuum denominator for 1802271 would bring the green curve down by 13% at 0.2 μm, while only raising the level at 2.5 μm by 3%. For the red 1732526 curve, a dereddening corresponding to the $E(B-V) = 0.061$ from Table 2 is applied, but agreement with Pickles below 1 μm is within the uncertainty in the reddening. Above 1 μm, the NICMOS SEDs are within a few percent of each other, while the IR part of the Pickles SED is ~25% too low.

Figure 17 illustrates another comparison with Pickles at G0V and $T_{\rm eff} = 5800$ K for the two CALSPEC stars of the same type, viz., P041C (6020 K) and P177D (5880 K) in Table 2. Below ~0.3 μm, solar type stars have low flux levels, which differ significantly from star-to-star. P041C is hotter than the Pickles SED for G0V, i.e., a bit high at the shorter wavelengths and too low in the IR. P177D (green curve) is the better reference for the Pickles SED, which is lower than CALSPEC in the IR by ~15%.

5.3. IUE

The original flux calibration of the UV spectra from the *IUE* satellite was based on η UMa (Bohlin et al. 1980, Bohlin 1988). Later, Bohlin (1996) published a correction for *IUE* fluxes to the same WD scale that is the basis of the *HST* CALSPEC fluxes. However, the *IUE* project reprocessed the whole *IUE* archive (Nichols & Linsky 1996) with an optimal extraction technique (Kinney et al. 1991), which required a new flux calibration that was based entirely on the 60,000 K pure-hydrogen model for the WD G191B2B. This final archive of *IUE* spectra is available from the Mikulski Archive for Space Telescopes (MAST).[11]

[10] http://www.ifa.hawaii.edu/users/pickles/AJP/hilib.html.
[11] http://archive.stsci.edu/iue/.

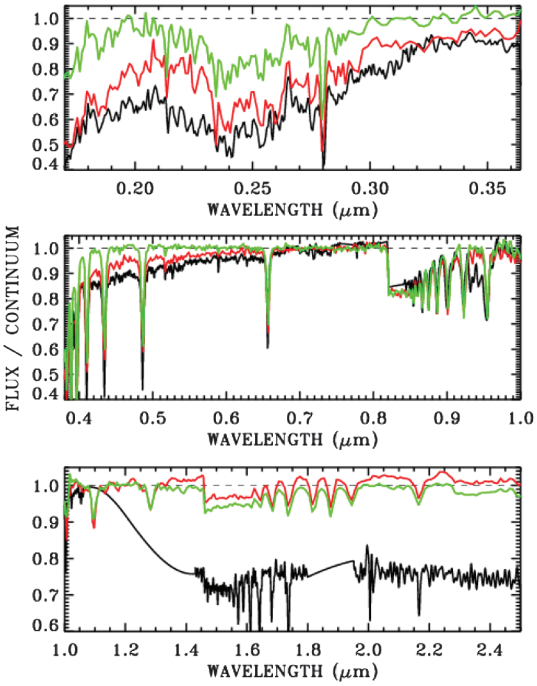

FIG. 16.—Comparison of three stellar SEDs to the same theoretical continuum level for $T_{\rm eff} = 8790$ K, where the line blanketing generally increases toward ultraviolet wavelengths. *Black line:* Pickles SED for an A3V star at $T_{\rm eff} = 8790$ K; A3V stars 1732526 (*red*) and 1802271 (*green*) have both STIS and NICMOS data.

Figure 18 compares the *IUE* final archive MAST SED for G191B2B to the current reference model from RWBK. The ratio shows little slope and is consistent with the expectation of Nichols & Linsky (1996). While a pure-hydrogen WD model for G191B2B provided the basis for the *IUE* flux calibration as a function of wavelength, the normalization of the model was to the UV fluxes of TD-1 (Fig. 5), instead of a normalization in the visible, e.g., the Megessier value of $F(5556) = 3.46 \times 10^{-9}$ for Vega. Nichols & Linsky state that *IUE* normalization is 6% lower than the HST flux scale, which is consistent with Figure 18.

5.4. FUSE

The *FUSE* astronomical satellite recorded far-UV spectra in the 905–1187 Å spectral region with a resolution of $R = 20,000$ (Moos et al. 2000). The dynamic range of FUSE was 10–11 magnitudes with a bright limit of $V \sim 11$ for an unreddened O star. The flux calibration is based on TLUSTY version 200 NLTE pure hydrogen models for six WDs, including the three primary *HST* flux standards (Dixon et al. 2007). The models are normalized to the V magnitude of each star. A comparison of one *FUSE* observation (P104120300000nvo4histfcal .fit) for G191G2B is compared to our Rauch model in Figure 19 after correcting for the radial velocity of 22 km s^{-1}. This *FUSE*

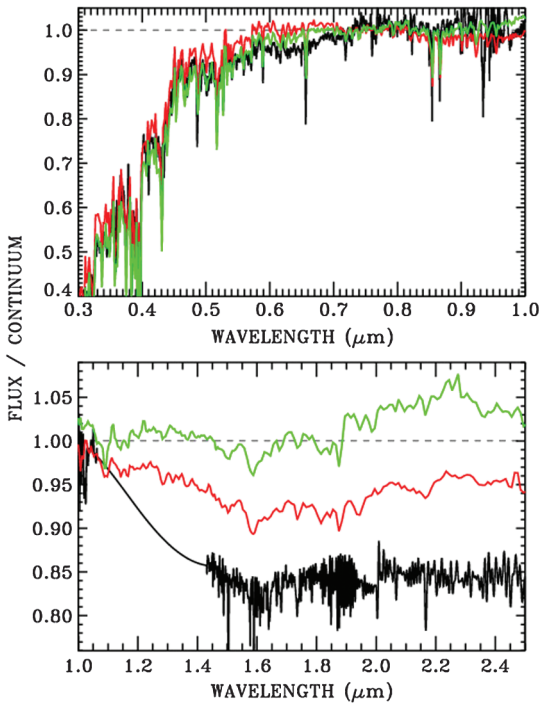

FIG. 17.—As in Fig. 16 for G0V stars. The Pickles SED for a G0V star at $T_{\rm eff} = 5800$ K is the black line, while the two CALSPEC G0V stars P041C (*red*) at $T_{\rm eff} = 6020$ K and P177D (*green*) at $T_{\rm eff} = 5880$ K have STIS data below 1 μm and NICMOS data that extends to 2.5 μm.

spectrum is one of 61 observations of G191B2B, from the MAST archive,[12] which contains over 6000 *FUSE* data sets.

5.5. SDSS

The Sloan Digital Sky Survey (SDSS) is a massive program to image a large fraction of the northern hemisphere (Ahn et al. 2012). In conjuction with the imaging survey, fiber spectra were obtained for about 2.6 million objects, including ~0.7 million stars in the range of g magnitude from 14 to 20 (Yanny et al. 2009). The flux calibration was originally based on the absolute flux of BD + 17°4708 (Fukugita 1996), whose flux was compared to STIS data by Bohlin & Gilliland (2004b). Betoule et al. (2013) discuss improvements to the SDSS fluxes. The spectra for the two stars in common with the CALSPEC archive are retrieved from the SDSS DR9 database[13] and are compared with the CALSPEC fluxes in Figure 20. The solid and dotted lines for the early SDSS data releases are consistent with each other and with the CALSPEC fluxes mostly within ~5%, while the newer DR9 dashed line for WD1657 + 343 is discrepant by up to ~25%. A revised flux calibration for the new DR9 instrumentation was not yet implemented.

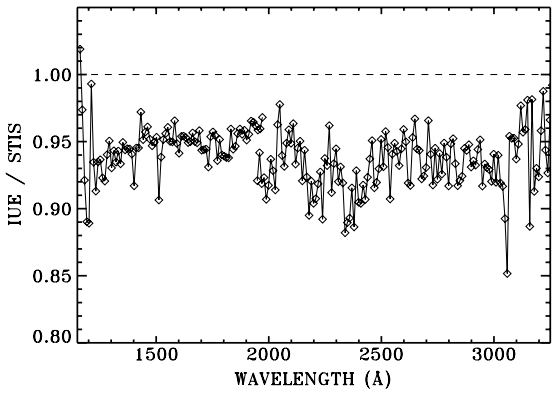

FIG. 18.—Comparison of *IUE* final archive fluxes and the reference standard model for G191B2B. The ratio is in 10 Å bins, and the *IUE* spectra are averages of 57 SWP and 48 LW individual observations of G191B2B with total exposure times of 8300 and 13500 s, respectively.

5.6. Infrared Space Observatory

A comparison of the *ISO* short wave spectrometer fluxes[14] for Sirius with the CALSPEC reference standard of Bohlin (2014) appears in Figure 21. The different curves correspond to different scan speeds with the differences reflecting different samplings of the residual nonlinearities of the detectors. If the two separate *ISO* scans could be averaged, the agreement would be excellent at the few percent level.

5.7. Spitzer Space Telescope

Figure 22 compares measured spectra from the CASSIS reductions[15] of *Spitzer Space Telescope* infrared spectrometer observations with the CALSPEC fluxes that are extrapolated from the shorter wavelengths using CK04 models. One star, HD163466, has multiple observations in the IRS archive, and the illustrated average of three of these independent observations is on an expanded y-scale. This average IRS data for HD163466 agrees with CALSPEC to <1% at the shorter wavelengths and to ~2% in the 25–30 μm range. Judging from the scatter of the other stars at the longer wavelengths, this ~2% difference for HD163466 probably reflects the statistical reproducability of the IRS data, and more IRS scans would need to be analyzed to measure any systematic deviation of IRS fluxes from the CALSPEC scale.

Even though the other seven stars have only single IRS scans with correspondingly poorer S/N, the following conclusions are suggested:

1. The O and the B stars lie considerably more above unity than the A and G stars. So, either the hot models are less precise than the cooler part of the CK04 grid, or there is an instrumental

[12] http://dr9.sdss3.org/advancedSearch.
[13] http://archive.stsci.edu/fuse/.
[14] http://isc.astro.cornell.edu/~sloan/library/2003/swsatlas.html.
[15] http://cassis.sirtf.com/atlas/.

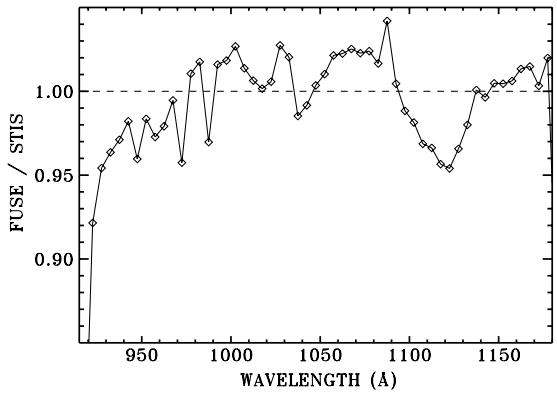

FIG. 19.—Comparison of FUSE archival fluxes with the reference standard model for G191B2B. The ratio is in 5 Å bins, and the total FUSE exposure time is 15051 s.

problem such as a shorter wavelength light leak or contamination from a different spectral order that is significant for the hottest stars. The elevated plateau in the 14–20 μm range is strong evidence for some sort of instrumental problem for the two hottest stars.

2. The scatter increases for the fainter stars, especially longward of 14 μm. Thus, the best photometric comparison is below 14 μm for the six cooler stars.

3. For the A and G stars, CALSPEC and the *Spitzer* IRS flux shortward of 14 μm agree to better than 3%, which suggest that neither estimate has systematic errors of more than ~3% in absolute flux.

6. FUTURE DIRECTIONS

On the theoretical front, disparities in the NLTE modeling is the largest limitation to the precision of the three primary

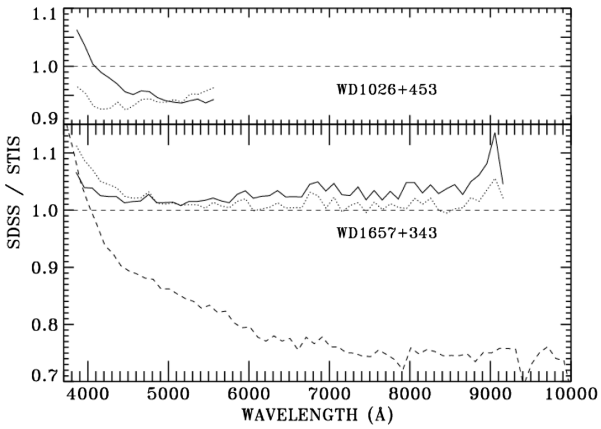

FIG. 20.—Comparison of SDSS archival fluxes with two of the faintest CALSPEC WDs in 100 Å bins. Unfortunately, there are no long-wavelength G750L STIS observations for WD1026 + 453, so that comparison ends at 5500 Å. The solid and dotted lines are for early SDSS data releases, while the dashed line for WD1657 + 343 is from the recent DR9 with expanded wavelength coverage.

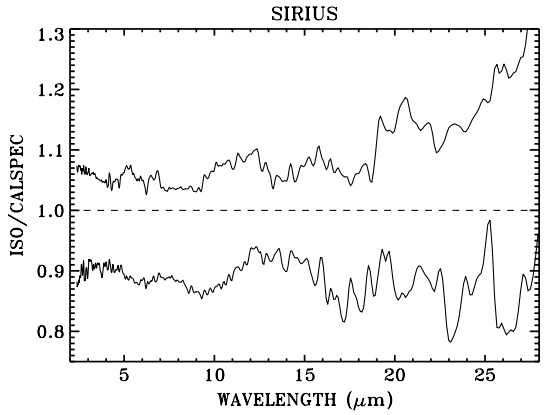

FIG. 21.—Comparison of ISO archival fluxes with the Sirius CALSPEC SED at a resolution of $R = 100$. The ratio above unity is for an ISO speed of 4, while the lower curve is the ISO SED at a speed of 1.

WD flux standards. Ideally, the preeminent developers of the modeling codes might confer and determine whether the discrepancies in SEDs calculated for the same $T_{\rm eff}$ and $\log g$ are caused by different input physics or numerical imprecisions. Secondarily, the uncertainties on the derivation of $T_{\rm eff}$ and $\log g$ are a floor to the precision of the models (*dotted line* in Fig. 13). Perhaps, better data or improved analysis techniques could reduce this contribution of the systematic error budget of the fundamental WD models.

Experimentally, the NIST laboratory reference standards that have subpercent precision should be transferred to the stars. Rocket or other space platforms provide the most straightforward path for implementation. However, the extensive programs to measure the temporal and wavelength variations of the

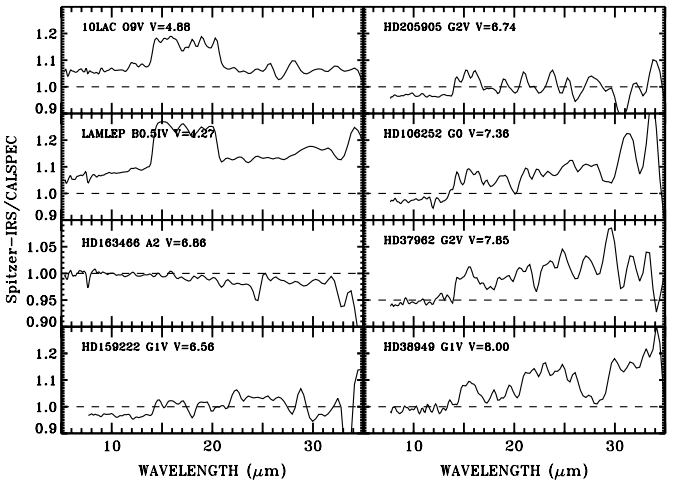

FIG. 22.—Comparison of *Spitzer* IRS archival fluxes with the CALSPEC SEDs at a resolution of $R = 50$. Notice the expanded y-scale for HD163466, where three IRS spectra are averaged to improve the S/N. The five G stars are arranged in decreasing brightness down the columns; and show increasing scatter toward longer wavelengths for fainter stars with lower signal strength.

atmospheric transmission in the visible and near-IR may well produce stellar standards with precise flux distributions at the top of the atmosphere. To verify the *MSX* results in the mid-IR and extend the *MSX* concept to other wavelengths, a new mission is needed with onboard, in situ absolute flux sources like the *MSX* ejected reference spheres. The success of these experimental programs would free the flux standards from the uncertainties of the theoretical model calculations.

In summary, the current set of CALSPEC stars are the best available stellar flux standards for the FUV to mid-IR, where the SEDs as a function of wavelength are based on NLTE models atmospheres for the primary standard WD stars of Table 1. Table 2 includes many of the secondary CALSPEC standards.

Thanks to Derck Massa for a patient explanation of the nuances of covariance matrices. S. Deustua, M. Kaiser, R. Kurucz, and the referee provided constructive comments on preliminary drafts. Support for this work was provided by NASA through the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. This research made use of the SIMBAD database, operated at CDS, Strasbourg, France.

REFERENCES

Aalders, J. W. G., van Duinen, R. J., Luinge, W., & Wildman, K. J. 1975, Space Sci. Instr., 1, 343
Absil, O., et al. 2013, A&A, 555, A 104
Ahn, C. P., et al. 2012, ApJS, 203, 21
Aldering, G., et al. 2004, PASP, in press (astro-ph/0405232)
Arp, U., Friedman, R., Furst, M., Maker, S., & Shaw, P. 2000, Metrologia, 37, 357
Anderson, J., & Bedin, L. 2010, PASP, 122, 1035
Aufdenberg, J. P., et al. 2006, ApJ, 645, 664
Aumann, H. H., et al. 1984, ApJ, 278, L 23
Becklin, E. E., Neugebauer, Hansen, & Kieffer 1973, AJ, 78, 1063
Bessell, M., & Murphy, S. 2012, PASP, 124, 140
Betoule, M., et al. 2013, A&A, 552, A124
Blackwell, D. E., Leggett, S. K., Petford, A. D., Mountain, C. M., & Selby, M. J. 1983, MNRAS, 205, 897
Bless, R. C., Code, A. D., & Fairchild, E. T. 1976, ApJ, 203, 410
Bohlin, R. C. 1988, in New Directions in Spectrophotometry, ed. A. G. D. Philip, D. S. Hayes, & S. J. Adelman (Schenectady, N.Y.: L. Davis Press), 121
Bohlin, R. C. 1996, AJ, 111, 1743
———. 1998, Instrument Science Report, STIS 1998-01 (Baltimore: STScI)
———. 2003, in 2002 HST Calibration Workshop, ed. S. Arribas, A. Koekemoer, & B. Whitmore (Baltimore: STScI), 115
———. 2007, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric, and Polarimetric Standardization, ed. C. Sterken (Ann Arbor: Sheridan Books), 315
———. 2010, AJ, 139,1515 (B10)
———. 2012, Instrument Science Report, ACS 2012-01 (Baltimore: STScI)
———. 2014, AJ, 147, 127
Bohlin, R. C., & Cohen, M. 2008, AJ, 136, 1171
Bohlin, R. C., Colina, L., & Finley, D. S. 1995, AJ, 110, 1316
Bohlin, R. C., Dickinson, M. E., & Calzetti, D. 2001, AJ, 122, 2118
Bohlin, R. C., Frimout, D., & Lillie, C. 1974, A&A, 30, 127
Bohlin, R. C., & Gilliland, R. L. 2004a, AJ, 127, 3508
———. 2004b, AJ, 128, 3053
Bohlin, R. C., Harris, A. W., Holm, A. V., & Gry, C. 1990, ApJS, 73, 413
Bohlin, R. C., Holm, A. V., Savage, B. D., Snijders, M. A. J., & Sparks, W. M. 1980, A&A, 85, 1
Bohlin, R. C., Savage, B. D., & Drake, J. F. 1978, ApJ, 224, 132
Bohlin, R. C., et al. 2011, AJ, 141, 173 (B11)
Booth, A. J., Selby, M. J., Blackwell, D. E., Petford, A. D., & Arribas, S. 1989, A&A, 218, 167
Brown, S. W., Eppeldauer, G. P., & Lykke, K. R. 2006, Appl. Opt., 45, 8218
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Castelli, F., & Kurucz, R. 2003, in Poster A20, IAU Symp. 210, Modeling of Stellar Atmospheres, ed. N. Piskunov, W. Weiss, & D. Gray (Cambridge: Cambridge Univ. Press) (CK04)
Cohen, M. 2007, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric, and Polarimetric Standardization, ed. C. Sterken (Ann Arbor: Sheridan Books), 333
Cohen, M., Walker, R., Barlow, M., & Deacon, J. 1992a, AJ, 104, 1650
Cohen, M., Walker, R. G., & Witteborn, F. C. 1992b, AJ, 104, 2030
Cohen, M., Wheaton, W. A., & Megeath, S. T. 2003, AJ, 126, 1090
Defrère, D., et al. 2011, A&A, 534, A 5
De Marchi, G., et al. 2004, Instrument Science Report, ACS 2004-08 (Baltimore: STScI)
Deustua, S., Kent, S., & Smith, J. A. 2013, in Planets, Stars and Stellar Systems, Vol. 2, ed. H. E. Bond (New York: Springer)
Dixon, W. V., et al. 2007, PASP, 119, 527
Dupuis, J., Vennes, S., Bowyer, S., Pradhan, A. K., & Thejll, P. 1995, ApJ, 455, 574
Engelke, C. W., Price, S. D., & Kraemer, K. E. 2010, AJ, 140, 1919
Finley, D. S., Basri, G., & Bowyer, S. 1984, in Future of Ultraviolet Astronomy Based on Six Years of IUE Research, ed. Y. Kondo, R. D. Chapman, & J. M. Meade, (NASA CP-2349; Washington: NASA), 277
———. 1990, ApJ, 359, 483
Finley, D. S., Koester, D., & Basri, G. 1997, ApJ, 488, 375 (FKB)
Fraser, G. T., Brown, S. W., Yoon, H. W., Johnson, B. C., & Lykke, K. R. 2007, Proc. SPIE, 6678, 66780
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
Gianninas, A., Bergeron, P., & Ruiz, M. T. 2011, ApJ, 743, 138 (G11)
Gilliland, R. L., & Rajan, A. 2011, Instrument Science Report, WFC3 2011-03 (Baltimore: STScI)
Goudfrooij, P., Bohlin, R., Maiz-Apellaniz, J., & Kimble, R. 2006, PASP, 118, 1455
Gustafsson, B., Edvardsson, B., Eriksson, K., Jorgensen, U. G., Nordlund, A., & Plez, B. 2008, A&A, 486, 951

Hayes, D. S. 1970, ApJ, 159, 165
———. 1985, in IAU Symp. 111, Calibration of Fundamental Stellar Quantities, ed. D. S. Hayes, L. E. Pasinetti, & A. G. Davis Philip (Dordrecht: Reidel), 225 (H85)
Hayes, D. S., & Latham, D. W. 1975, ApJ, 197, 593
Hayes, D. S., Latham, D. W., & Hayes, S. H. 1975, ApJ, 197, 587
Henry, R. C., Weinstein, A., Feldman, P. D., Fastie, W. G., & Moos, H. W. 1975, ApJ, 201, 613
Holberg, J. B., & Bergeron, P. 2006, AJ, 132, 1221
Holberg, J. B., Wesemael, F., & Basile, J. 1986, 306, 629
Hubeny, I., & Lanz, T. 1995, ApJ, 439, 875
Humphries, C. M., Jamar, C., Malaise, D., & Wroe, H. 1976, A&A, 49, 389
Jamar, C., Macau-Hercot, D., Monfils, A., Thompson, G. I., Houziaux, L., & Wilson, R. 1976, Ultraviolet Bright-Star Spectrophotometric Catalogue, ESA SR-27 (France: European Space Agency, Paris,)
Jeach, J. L. 1985, Statistical Analysis of Measurement Errors, an Exxon Monograph, (New York: Wiley)
Kaiser, M. E., Kruk, J., McCandliss, Sahnow, D., Dixon, W., Bohlin, R., & Deustua, S. 2007, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric, and Polarimetric Standardization, ed. C. Sterken (Ann Arbor: Sheridan Books), 361
Kaiser, M. E., et al. 2010a, Proc. SPIE, 7731, 77313 I
———. 2010b, STScI Calibration Workshop, ed. S. Deustua, & C. Oliveira (Baltimore: STScI), 108
———. 2012, SPIE, 8442, 844246
———. 2013, SPIE, 8660, 88600 Y
Kent, S., et al. 2009, ASTRO2010, 155, (astro-ph/0903.2799)
Kinney, A., Bohlin, R., & Neill, D. 1991, PASP, 103, 694
Koornneef, J., Bohlin, R., Buser, R., Horne, K., & Turnshek, D. 1986, Highlights of Astronomy, ed. J.-P. Swings (Dordrecht: D. Reidel Publishing Co.), Vol. 7, 833
Koornneef, J., & Code, A. 1981, ApJ, 247, 860
Krist, J., & Hook, R. 2004, The Tiny Tim User's Guide, Version 6.3 (Baltimore: STScI) http://www.stsci.edu/software/tinytim
Kruk, J. W., et al. 1997, ApJ, 482, 546
Landolt, A. U. 1992, AJ, 104, 340
Landolt, A. U., & Uomoto, A. K. 2007, AJ, 133, 768
Low, F. J., Rieke, G. H., & Armstrong, K. R. 1973, ApJL, 183, L 105
Ludwig, H.-G., Behara, N. T., Steffen, M., & Bonifacio, P. 2009, A&A, 502, L 1
McClintock, W. E., Snow, M., & Woods, T. N. 2005, Sol. Phys., 230, 259
McGraw, J. T., et al. 2010, Proc. SPIE, 7739, 773929
Megessier, C. 1995, A&A, 296, 771
Melendez, J., & Ramirez, I. 2007, ApJ, 669, L 89
Mészáros, Sz., et al. 2012, AJ, 144, 120
Mill, J. D., et al. 1994, J. Spacecraft Rockets, 31, 900
Moos, H. W., et al. 2000, ApJL, 538, L 1
Mountain, C. M., Selby, M. J., Leggett, S. K., Blackwell, D. E., & Petford, A. D. 1985, A&A, 151, 399
Nichols, J. S., & Linsky, J. L. 1996, AJ, 111, 517
Oke, J. B., 1990, AJ, 99, 1621
Oke, J. B., & Schild, R. E. 1970, ApJ, 161, 1015
Opal, C. B., Moos, H. W., Fastie, W. G., Bottema, M., & Henry, R. C. 1968, ApJL, 153, 179
Pickles, A. J. 1998, PASP, 110, 863
Pickles, A. 2010, in STScI Calibration Workshop, ed.S. Deustua, & C. Oliveira (Baltimore: STScI), 614
Price, S. D. 2004, Space Sci. Rev., 113, 409
Price, S. D., Paxson, C., Engelke, C., & Murdock, T. L. 2004, AJ, 128, 889
Rauch, T., Werner, K., Bohlin, R. C., & Kruk, J. W. 2013, A&A, 560, A 106 (RWBK)
Reach, W. T., et al. 2005, PASP, 117, 978
Rieke, G. H., Lebofsky, M. J., & Low, F. J. 1985, AJ, 90, 900
Rieke, G. H., et al. 2008, AJ, 135, 2245
Selby, M. J., et al. 1980, MNRAS, 193, 111
Selby, M. J., Mountain, C. M., Blackwell, D. E., Petford, A. D., & Leggett, S. K. 1983, MNRAS, 203, 795
Sibthorpe, B., et al. 2010, A&A, 518, L 130
Sirianni, M., et al. 2005, PASP, 117, 1049 S
Smith, A. W., Woodward, J. T., Jenkins, C. A., Brown, S. W., & Lykke, K. R. 2009, Metrologia, 46, 219
Snow, M., Reberac, A., Quemerais, E., Clarke, J., McClintock, W. E., & Woods, T. N. 2013, in ISSI Scientific Report Series, Vol. 13, Cross-Calibration of Far UV Spectra of Solar System Objects and the Heliosphere, ed. E. Quemerais, M. Snow, & R. M. Bonnet (New York: Springer), 191
Stecher, T. P. 1970, ApJ, 159, 543
Stritzinger, M., et al. 2005, PASP, 117, 810
Strongylis, G. J., & Bohlin, R. C. 1979, PASP, 91, 205
Stubbs, C. W., & Tonry, J. L. 2012, in Proc. Calibration and Standardization of Large Surveys and Missions in Astronomy and Astrophysics Conf., in press (astro-ph/1206.6695v1)
Su, K. Y .L., et al. 2005, ApJ, 628, 487
———. 2006, ApJ, 653, 675
———. 2013, ApJ, 763, 118
Sullivan, M., et al. 2011, ApJ, 737, 102
Thompson, G. I., Nandy, K., Jamar, C., Monfils, A.Houziaux, L., Carnochan, D. J., & Wilson, R. 1978, The Science Research Council, U.K
Thuillier, G., et al. 2003, Solar Physics, 214, 1 (Th03)
Tonry, J. L., et al. 2012, ApJ, 750, 99
Tremblay, P. E., & Bergeron, P. 2009, ApJ, 696, 1755
Tremblay, P.-E., Bergeron, P., & Gianninas, A. 2011, ApJ, 730, 128
Turnshek, D. A., Bohlin, R. C., Williamson, R. L., Lupie, O. L., Koornneef, J., & Morgan, D. H. 1990, AJ, 99, 1243
Walker, R. G. 1969, R. Soc. London Philosophical Trans. Ser. A, 264, 209
Werner, K., et al. 2003, in ASP Conf. Ser. 288, Workshop on Stellar Atmosphere Modeling, eds. I. Hubeny, D. Mihalas, & K. Werner (San Francisco: ASP), 31
Wesselius, P. R., van Duinen, R. J., de Jonge, A. R. W., Aalders, J. W. G., Luinge, W., & Wildeman, K. J. 1982, A&AS, 49, 427
Witteborn, F. C. 1999, AJ, 117, 2552
Yanny, B., et al. 2009, AJ, 137, 4377
Zimmer, P. C., et al. 2010, Proc. SPIE, 7735, 77358 D
Zong, Y., Brown, S. W., Johnson, B. C., Lykke, K. R., & Ohno, Y. 2006, Appl. Opt., 45, 1111