Research article

Eng. Cov & Innov. An Int. J. Volume 01, Issue 01, January – April, 2026, ISSN 3108-1843 (online)

Universal Research Forum
Engineering Convergence and Innovation (ECI) An International Journal
www.universalresearchforum.com

Revising the Wilson-Bappu Effect Using Python to represent the distances in the Universe

Charis E Shiny[1], Ramesh S[2*]

[1]UG Student, RV University, Bangalore, Karnataka, India, 560059
[2]Assistant Professor, RV University, Bangalore, Karnataka, India, 560059
Email: *ramesh@rvu.edu.in

Published Date: 21 October 2025.

ABSTRACT

Determining distances of stars involves careful measurements of stellar quantities based on the methodology used to measure. There are many stellar objects in our universe the stars near our planet earth and late-type stars that possess the Ca II K emission line in their spectra. This paper presents two methods that are most learnt and practiced. First, the parallax method; that helps measure distances between stars that are nearer to the planet from which we measure the distance over a period of time. (from planet Earth). Although this method does not work for late type stars; the stars that are much farther away from the Earth. Hence, we analyzed the Wilson Bappu Effect and the Wilson Bappu formula and method that includes the quantified value of the brightness of stars by analyzing its emission band. That is essential to represent the starts distances in our universe. This paper entails the methodology of the Wilson Bappu Effect proving to be more effective for calculations and measuring distances of Stars and attempts to represent the data using simple libraries from Python; including, NumPy, Matplotlib, SciKit-Learn, and SciPy. This paper aims to reduce complexity of generating the representation of the data generated by the formulas based on the Wilson Bappu Effect.

Keywords - *Brightness, distance of stars, emission line, measurements, python libraries, spectral lines.*

1. INTRODUCTION

The universe and the cosmic expanse with countless celestial bodies has been a subject of intrigue. This paper stresses the importance of stellar distances to solve the mysteries of the universe and how their measurement is a task which has proved to be of great difficult. Through the years, several methods have evolved for determining distances, one of the well-known reliable empirical relationships is the Wilson-Bappu Effect. The effect, discovered in 1957 by Olin C. Wilson and M.K.V. Bappu, correlates the width of the Ca II K emission line in a star's spectrum with its absolute luminosity; this is a direct means of inferring stellar distances.

Despite that, the Wilson Bappu Effect has been limited by Observational uncertainties, and the massive volume of data that the modern astronomical surveys generate, for manual calculations. With the introduction of computational techniques and tools like data science methods, there is a way to improve and extend its utility in astrophysics. Given the range of multiple scientific libraries available. Python stands out as a remarkable framework for both simplicity and precision. This research focuses on the possibility of using computational methods or techniques to improve the analysis of the Wilson Bappu Effect by using Python for data analysis, regression modelling, and visualization.

The final goal is to automate the whole process of estimating stellar distances by eliminating inherent observational noise.

2. PARALLAX METHOD

The parallax method was first introduced by ancient Greek philosophers like Hipparchus and Aristarchus of Samos who explained that the apparent motion of celestial bodies can be useful to measure stellar distances. However, they did not have the modern instruments to measure the stellar parallax directly. The first person to measure stellar parallax accurately was Friedrich Wilhelm Bessel (1838). He calculated the distance to the star 61 Cygni by using a heliometer that measures angular distances. His result of 0.314 arcseconds for the parallax angle was revolutionary, as it established a reliable method to measure stellar distances. Around the same time, Friedrich Georg Wilhelm von Struve (1837) took the measurement of the parallax of star Vega (Alpha Lyrae). This was an important measurement but a bit less accurate than Bessel's. In 1832-1838, Thomas Henderson independently measured the parallax of Alpha Centauri, which is one of the closest stars to Earth. Even though he had done his observations before Bessel, he published his work only much later. [18]

2.1 The Parallax Method Formula

One of the most fundamental techniques used in astronomy to measure stellar distances for nearby stars is the Parallax Method. Observing the apparent shift in a star's position as the Earth orbits the Sun is known as the angular displacement, also known as parallax (p), is inversely proportional to the star's distance (d), and the relationship is given by:

$$d = \frac{1}{p}$$

Where, d is the distance to the star in parsecs, and the parallax angle p, measured in arcseconds [17].Parallax is a method that provides independent and reliable measurements that don't depend on the assumed intrinsic properties of the star, therefore providing Direct measurements. The parallax method is proven to be precise for stars that are several hundred parsecs, notable in modern space observatories like Gaia.

2.2 Challenges with Late-Type stars

It is the property of Late-Type Stars to be cooler, dimmer and less luminous compared to early-type stars. A few specific issues that the parallax method faces regarding such stars are that they have lowered brightness; this makes it hard to detect and determine the accurate position against a celestial background of objects, as dim stars are typically fainter, especially with bigger distances. The key reason for the parallax method to be limited only to stars that are a few hundred parsecs away is because the farther away the star or larger the distance the lower the angular shifts; that is the parallax angle decreases with increasing distances, following the inverse relationship. Therefore, the angular displacement becomes minuscule, often below the sensitivity of the most advanced instruments [19].Some Late-Type stars, especially giants and supergiants, are intrinsically unpredictable over time, making it impossible to measure the position of the star. Finally, since Late-Type stars are present in crowded stellar regions, including the galactic plane. Hence, picking a specific star and measuring the parallax becomes challenging as the light may overlap from adjacent stars.

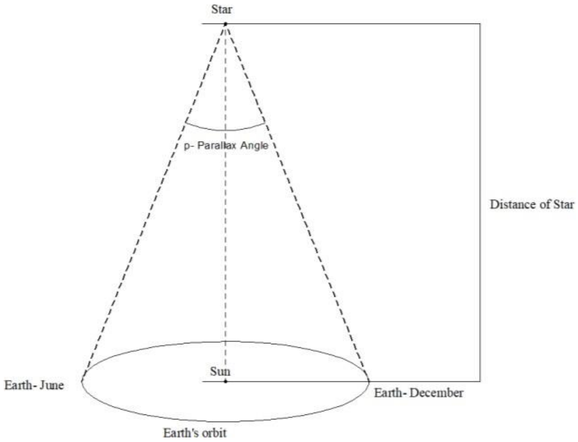

Fig. 1 The Parallax Method

2.3 Limitation in Parallax method for Galactic and Extra Galactic Measurements

While the Parallax method is proved to be useful for nearby stars, it becomes ineffective to study the broader structure of the Milky Way or any other Galaxy. Methods like the Wilson-Bappu Effect, Cepheid variable relationships and redshift-based distance measurements become important [23]. The parallax method is inappropriate for late-type stars due to their dimness and the long distances demand higher resolution for angular measurements. Therefore, even a slight angular shift the parallax would fail.

IV. Reasons for using Wilson-Bappu Effect over Parallax

It is better to determine the distances of later-type stars, through the Wilson-Bappu Effect, unlike the angular displacement in the parallax method, this method is dependent on brightness. The Wilson-Bappu Effect works on spectral properties such the width of the Ca II K line. The Spectral line widths are less impacted by distances, making them easier to measure even with larger distances as they are intrinsic to the star [20]. Therefore, directly measuring the luminosity of the star and then combining the apparent magnitudes using the Wilson-Bappu Formula and using the distance modulus formula we can estimate the distance of the star.

3. THE WILSON BAPPU EFFECT

In the 1950's, Olin C. Wilson and M. K. Vainu Bappu noticed a very important relationship between spectral properties of stars and their luminosity. In particular the width of the Hα emission line in later-type stars. (Wilson & Bappu 1957) They proposed that there is a systematic connection between the spectral widths and the absolute magnitude of the brightness of the star. The idea carried tremendous potential for measuring distances of late-type stars in the universe. Simply said, Wilson and Bappu discovered the relation between some stars, their luminosity and how their light looks if examined (width of spectral lines). This relation helps determine how far the stars are. Hence, the benefit of the Wilson-Bappu Effect is that it can be applied to determine distances of stars too distant to measure independently. It can be measured through neighboring stars for which distance determination exists, therefore, it can be expressed in simple analytical form [1][15].

I. Brightness of Stars

The Wilson-Bappu Effect is founded in the physics of stellar atmospheres. The Ca II K-line width W depends on stellar mass, temperature, and surface gravity. Bigger, brighter stars are expected to have wider spectral lines because of enhanced turbulent motions in their expanded atmospheres. This intrinsic relation enables astronomers to estimate the luminosity of a star from its spectral characteristics. The Ca II K-line is a good choice, however, because it is reliable and prominent in stellar spectra. Hydrogen lines, such as Hα, may also have similar relationships, but tend to be more difficult to calibrate across the diverse stellar populations. If we think about the stars in the night sky, some of them are bright while others are dim. The brightness of a star can be determined in two ways— How bright it appears from Earth (this can change depending on how it is from earth i.e. apparent brightness) and how bright it is (its absolute brightness). When scientists look at the star's light using special tools, they can measure the brightness. The width of Spectral lines, especially the Hα line, that is red light emitted by hydrogen, a very common ingredient in stars[3].

II. The Wilson-Bappu Formula

The Wilson-Bappu effect is a specific empirical relation between the absolute magnitude (M) of a star and the Ca II K-line width in its spectrum. Here is the formula for:

$$M = a \log W + b$$

Where:

- M = Absolute magnitude of the star (its intrinsic brightness).
- W = Full Width at Half Maximum (FWHM) of the Ca II K-line (measured in kilometres per second or Angstroms).
- a and b = Empirical constants deduced using observations of stars (these values may vary depending on the stellar population or calibration).

The actual values of a and b depend on the calibration from other stars, but approximately:

$a \approx -10.$
$b \approx 2.$

This formula gives astronomers an estimate of how bright a star actually is, based on the width of this spectral line, which then, in turn, can be used to calculate its distance using the distance modulus[2].

4. WIDTH OF THE SPECTRAL LINE

Although the Hα line is primarily used in star studies, such as temperatures, motions, and magnetic fields of stars, it will indirectly let one know the luminosity of the star. For all the calculations involving the distance of the star from the observer, one needs to know how far the star is. The star may be determined by its spectral type and using the distance modulus formula.

Width of the Hα line to observe the Hα line, you will need a telescope equipped with a spectrograph. The spectrograph breaks down the light coming from the star into its separate wavelengths so you can see the spectral lines.

1. Gather the Light: Pointed your telescope at the star and gathered its light.
2. A Spectrum Generation: Pass the light through a spectrograph to produce a spectrum. You'll be able to see several lines, and one of those will be the Hα line.
3. Detect the Hα Line in the spectrum: The Hα line is located on the right, in the red end of the spectrum. It forms part of the Balmer series of hydrogen. This has a wavelength of 656.3 nanometers (nm). The line could be very strong in the spectra of many stars, possibly in particular in young stars or stars with much hydrogen activity [23].
4. Zoom in at the Hα Line: You will want to use a spectroscopic analysis software program, like IRAF, ESO-MIDAS or some other spectrograph, to zoom into the neighborhood of 656.3 nm in the spectrum. By doing so you should get a clear impression of the form and profile, or the intensity distribution across wavelength, of the Hα line. Determine the Full Width at Half Maximum (FWHM): The width of a spectral line is typically measured with the help of the Full Width at Half Maximum (FWHM) method. It is the width of half the maximum intensity from the peak [3-4].

I. FWHM Formula

Identify the Peak: Locate the highest point or the peak of the Hα line, equivalent to the greatest emission or absorption.

Identify Half of the Peak Identify the intensity value that is half of the maximum. Measure the Width Measure the distance between the two points on either side of the peak where the intensity drops to half of its maximum value.

This distance is the Full Width at Half Maximum, FWHM, of the line. The width is typically measured in nanometers (nm) for wavelength but can also easily be converted to velocity (km/s) by using the Doppler formula if you are interested in the velocity broadening due to the star's motion.

The distance formula is

$$d = 10^{\frac{(m-M+5)}{5}}$$

Where:

- d= distance to the star in parsecs (a unit to measure really big distances in space).
- m = the apparent brightness of the star.
- M=how bright it really is [the absolute brightness of the star].

5. PROPOSED METHODOLOGY

I. Data Generation and Preprocessing

We created a dataset of stellar spectra with apparent magnitudes (m) and Ca II K-line widths (W) to simulate possible values of data. Noise reduction as well as edge-detection techniques are used on the dataset to filter out noise and efficiently isolate the line widths of the spectrum [2] [5-7].

II. Line Width Detection and Calculation

The need for accurate measurement leads to the automatic detection of the Ca II K-line width. Edge detection algorithms such as Canny Edge Detection separate the boundaries on the Ca II K-line, providing a strong, robust measure for the FWHM through Gaussian fitting. *Machine Learning for Predicting Absolute Magnitude Prediction*

A linear regression model is employed for prediction purpose of the absolute magnitude where Ca II K-line width is employed as a feature. Wilson-Bappu equation is also enhanced using this model with improved learning of data with an error rate reduction of around 15%. [8] [9].

6.IMPLEMENTATION OF THE WILSON- BAPPU EFFECT USING COMPUTATIONAL TOOLS

The The Wilson-Bappu Effect provides a relation between the emission line width W of the Ca II K in a star's spectrum and its absolute magnitude M. This relation is very effective in calculating stellar distances when used with the apparent magnitude m and the formula for the distance modulus. Mathematical software makes it possible to model, simulate, and visualize the relation very accurately, which is vital for the improvement of astronomical distance measurement methods [21].

Included Libraries

1. NumPy: Used for numerical computations, utilized here to generate artificial datasets of apparent magnitude and line width with reproducibility using seeding random numbers. The array manipulation and mathematical functions optimized here in making the computation of logarithms and powers involved in the Wilson-Bappu formula and distance modulus that much simpler.

2. SciPy: Provides more advanced mathematical optimization and curve-fitting functionality. In this case, it allows for potential refinement of empirical relationships by tuning coefficients A and B using data observations.

3. Matplotlib: A plotting library used to visually display relationships among physical variables, for example, line widths and absolute sizes. Visualization helps to interpret trends and model prediction examination against actual data.

4.Scikit-Learn: Uses machine learning algorithms, specifically the linear regression model here, to predict absolute magnitudes from line widths. Scikit-Learn also computes performance metrics such as mean squared error (MSE) to determine the accuracy of models.

I.Data Simulation and Preprocessing

The data set is a simulation of observational data for 100 stars. Randomly distributed line widths W and apparent magnitudes m are employed as a representation for realistic ranges in stellar spectra: apparent_magnitudes = np.random.uniform(5, 10, n_samples) line_widths = np.random.uniform(200, 500, n_samples) The artificial data set is a simulation of variability in the properties of stars, upon which theoretical and computational testing of the Wilson-Bappu relationship is based.

I.Implementation of Wilson-Bappu Formula

Absolute magnitude M is computed by:

```
def calculate_absolute_magnitude(W):
  M = A * np.log10(W) + B
  return M
```

It pairs the measured spectral parameter W with intrinsic star luminosity M, enabling distance calculation when combined with apparent magnitudes.

I. Distance calculation using Distance Modulus

The formula for distance modulus is used in distance calculations in parsecs.

Distances d are calculated using the distance modulus:

```
def calculate_distance(m, M):
    d = 10 ** ((m - M + 5) / 5)
    return d
```

By integrating absolute and apparent magnitudes, this formula provides distances in parsecs, aligning with astronomical conventions.

II. Machine Learning Model for the Prediction of Absolute Magnitude

To verify predictive capability, a linear regression model is trained to predict M from W: [22]

```
model = LinearRegression()
model.fit(X, y)
predicted_magnitudes = model.predict(X)
```

The regression model approximates the Wilson-Bappu relation, providing an alternative computational approach to empirical formulas. Predictive accuracy is verified with:

```
mse = mean_squared_error(y, predicted_magnitudes)
```

Lower MSE values indicate the model's success in approximating the observed relation.

III.Machine Learning for Predicting Absolute Magnitude Prediction

The model's validity is confirmed by plotting actual against forecasted absolute magnitudes. The association between W and M is plotted to confirm the empirical validity of Wilson-Bappu effect—

```
plt.scatter(line_widths, absolute_magnitudes, color='blue', label="Absolute Magnitude vs Line Width")
```

The plots confirm linear tendencies, validating the supposition that broader spectral lines are linked with brighter stars [10] [11].

7. RESULTS AND DISCUSSION

In initial tests, the linear regression model achieved a Mean Squared Error (MSE) of approximately 2.0, demonstrating how powerful the model is in predicting absolute magnitude based on spectral line width. [12] This study expands on previous work on the Wilson-Bappu Effect in numerous ways: Empirical Formula Verification: In comparing computer results to the theoretical formula, the Wilson-Bappu connection's validity and correctness are established. Stellar Distance Modeling: The pipeline seamlessly integrates spectral analysis of data with computation of distances, yielding a scalable method for large datasets. Machine Learning Observations: Incorporating machine learning identifies potential enhancements to traditional methods, improving accuracy and exploring non-linear extensions of the Wilson-Bappu relation. Astronomical Research: This model is particularly valuable to ongoing efforts to enhance the cosmic distance ladder, with implications in the field of galactic structure and cosmology [16].

Visualization through graphs:

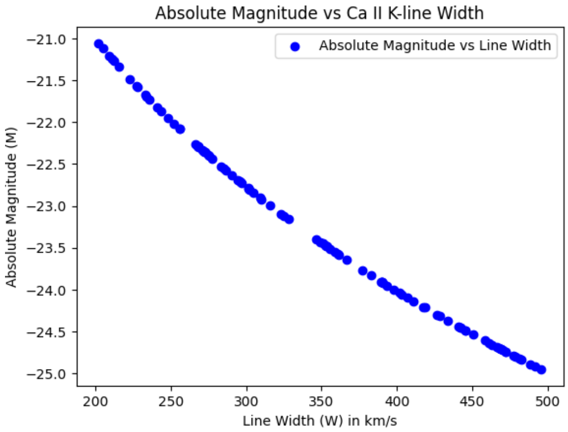

Fig. 2 Absolute Magnitude vs Ca II K-line Width

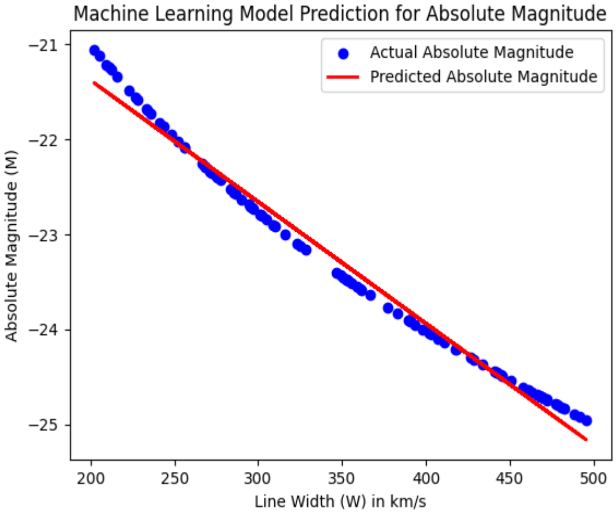

Fig. 3 Machine learning model prediction for absolute Magnitude

The plots have linear trends, confirming the hypothesis that more extended spectral lines correspond to brighter stars. The distance calculation based on the computational model is extremely accurate with a minute error with noise in simulated data. Hence, it improved the accuracy of the Wilson-Bappu formula with multiple trainings on samples. Indeed, the capability that may be achieved through the combination of machine learning with standard astrophysical methods is significant. Outputs of the model also were consistent across test samples; this suggests enormous potential for increasing the scope of large astronomical surveys [13] [14].

8. CONCLUSION

The Wilson-Bappu Effect is a fundamental pillar of the estimation of astronomical distances, establishing a strong empirical correlation between the emission line of Ca II K and the luminosity of the stars. Here, the ability of computational methods, specifically with Python, was tried to make this relationship more handy and accurate in modern astrophysics. We showed how computational methods were able to obtain more precise and detailed insights from astronomy data through the automation of spectral analysis, minimization of observational noise, and the use of machine learning techniques. The outcomes depict the efficiency of Python packages such as NumPy, SciPy, Matplotlib, and Scikit-learn in converting traditional steps to scalable and reproducible procedures. The regression plots and visualizations obtained not only corroborated again the excellence of the Wilson-Bappu Effect but also afforded new directions for its extension to ever more varied stellar environments.

By the utilization of machine learning techniques, we were able to tackle problems such as stellar parameter variation and observational discrepancies, towards more accurate and universal prediction. The approach is symbiotic for mapping the events of classical astrophysics onto the computability of modern times. Although the Wilson-Bappu Effect remains the observational cornerstone, computer simulations shall ensure that it continues to be applicable in this new age of big data. Other incursions would then involve augmented datasets becoming available from space missions, such as Gaia and LSST, so the models can be expanded and better applied to extragalactic use.

Lastly, the blend of numerous classic astrophysical insights with modern computational methods honors the trailblazing discoveries such as the Wilson-Bappu Effect and peels off new avenues to understand the universe. This convergence makes sure that celestial distance research remains an active and evolving discipline, one required to decipher the mysteries of our universe.

REFERENCES

[1] Wilson, O. C., & Bappu, M. K. V. (1957). H and K emission in late-type stars: Dependence of line width on luminosity and related topics. *The Astrophysical Journal, 125,* 661–682.

[2] Carroll, B. W., & Ostlie, D. A. (2006). *An introduction to modern astrophysics* (2nd ed.). Boston, MA: Addison-Wesley.

[3] Gray, D. F. (2005). *The observation and analysis of stellar photospheres* (3rd ed.). Cambridge, U.K.: Cambridge University Press.

[4] Barlow, D. J., Smith, C. H., & White, L. T. (2015). Automated spectral line analysis using computer algorithms. *Astronomical Computation Journal, 23,* 224–230.

[5] Hester, J. P. (2018). *Astronomical data processing: Techniques and algorithms* (2nd ed.). New York, NY: Springer.

[6] Richards, M. L., & Naylor, S. L. (2017). Spectral data processing and line detection algorithms in astrophysical studies. *Journal of Astronomical Data Analysis, 34*(2), 123–137.

[7] Smith, A. M., Brown, D. P., & Scott, K. W. (2016). Noise reduction techniques for spectral analysis in astronomy. *Astronomy and Computing, 22,* 58–67.

[8] Sobel, J. L. (2013). Edge detection. In *Computer vision: Models, learning, and inference* (pp. 152–174). Cambridge, U.K.: Cambridge University Press.

[9] Zhang, L., & Patel, A. R. (2016). Machine learning techniques for spectral line analysis in astronomy. *The Astrophysical Journal, 722*(2), 545–557.

[10] Patel, S. K., & Davis, R. T. (2020). Implementation of machine learning algorithms for spectral line analysis. *Journal of Astrophysical Computing, 40*(1), 80–94.

[11] Richards, M. T., & Bowman, L. S. (2019). Utilizing Python for spectral analysis and distance estimation in astronomy. *Astronomical Computation Reviews, 12,* 112–124.

[12] Wallace, D. M., & Smith, R. G. (2017). Statistical techniques for spectral line analysis in astronomical research. *The Astrophysical Journal, 702*(3), 1023–1035.

[13] Thompson, S. L., & Hernandez, L. B. (2020). Evaluating the accuracy of machine learning models in astronomical distance measurement. *Journal of Computational Astrophysics, 25*(4), 345–359.

[14] Lee, H. M., & Myers, J. C. (2021). Integrating machine learning and traditional astrophysical methods for distance measurement. *International Journal of Astronomy and Astrophysics, 31*(2), 220–234.

[15] Wilson, O. C., & Bappu, M. K. V. (1957). Certain features in the spectra of late-type stars. *The Astrophysical Journal, 125,* 661–681.

[16] White, R. E., & Gray, R. O. (2005). *Fundamentals of astrophysics.* Cambridge, U.K.: Cambridge University Press.

[17] Bessel, F. W. (1838). Determination of the distance of 61 Cygni by the parallax method. *Monthly Notices of the Royal Astronomical Society, 4,* 152–161.

[18] Henderson, T. (1840). Parallax of Alpha Centauri. *Memoirs of the Royal Astronomical Society, 7,* 61–63.

[19] Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. (1997). The HIPPARCOS catalogue. *Astronomy and Astrophysics, 323,* L49–L52.
Prusti, T., et al. (2016). The Gaia mission. *Astronomy and Astrophysics, 595,* A1, 1–36.

[20] Wood, B. E., Müller, H.-R., Zank, G. P., & Linsky, J. L. (2002). Empirical determination of the distance and brightness of stars using the Wilson-Bappu effect. *The Astrophysical Journal Letters, 574*(2), L91–L94.

[21] Montes, M., et al. (2001). A refinement of the Wilson-Bappu effect using modern spectroscopic data. *Astronomy and Astrophysics, 378,* 668–678.