

Analyzing long-term performance of the Keck-II adaptive optics system

Emily Ramey^{a,*}, Jessica R. Lu^a, Ruoyi Yin^a, Steve Robinson^a,
Peter Wizinowich^b, Sam Ragland^b, Jim Lyke^b, Siyao Jia^a,
Shoko Sakai^c, Abhimat Gautam^c, Tuan Do^c, Matthew Hosek Jr.^c,
Andrea Ghez^c, Mark R. Morris^c, Eric Becklin^c and Keith Matthews^d

^aUniversity of California, Astronomy Department, Berkeley, California, United States

^bW. M. Keck Observatory, Kamuela, Hawaii, United States

^cUniversity of California, Department of Physics and Astronomy, Los Angeles, California, United States

^dCalifornia Institute of Technology, Division of Physics, Mathematics and Astronomy, Pasadena, California, United States

Abstract. We present an analysis of the long-term performance of the W. M. Keck observatory laser guide star adaptive optics (LGS-AO) system and explore factors that influence the overall AO performance most strongly. Astronomical surveys can take years or decades to finish, so it is worthwhile to characterize the AO performance on such timescales in order to better understand future results. The Keck telescopes have two of the longest-running LGS-AO systems in use today, and as such they represent an excellent test-bed for processing large amounts of AO data. We use a Keck-II near infrared camera 2 (NIRC2) LGS AO survey of the Galactic Center (GC) from 2005 to 2019 for our analysis, combining image metrics with AO telemetry files, multi-aperture scintillation sense/differential imaging motion monitor turbulence profiles, seeing information, weather data, and temperature readings in a compiled dataset to highlight areas of potential performance improvement. We find that image quality trends downward over time, despite multiple improvements made to Keck-II and its AO system, resulting in a 9 mas increase in the average full width at half maximum (FWHM) and a 3% decrease in the average Strehl ratio over the course of the survey. Image quality also trends upward with ambient temperature, possibly indicating the presence of uncorrected turbulence in the beam path. Using nine basic features from our dataset, we train a simple machine learning (ML) algorithm to predict the delivered image quality of NIRC2 given current atmospheric conditions, which could eventually be used for real-time observation planning and exposure time adjustments. A random forest algorithm trained on this data can predict the Strehl ratio of an image to within 18% and the FWHM to within 7%, which is a solid baseline for future applications involving more advanced ML techniques. The assembled dataset and coding tools are released to the public as a resource for testing new predictive control and point spread function-reconstruction algorithms. © 2022 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JATIS.8.2.028004](https://doi.org/10.1117/1.JATIS.8.2.028004)]

Keywords: adaptive optics; machine learning; predictive modeling.

Paper 21116 received Sep. 13, 2021; accepted for publication Feb. 22, 2022; published online May 30, 2022.

1 Introduction

Since its inception, the use of laser guide star adaptive optics (LGS-AO) in astronomy has vastly improved scientific image quality, but it has not yet matched the performance of space telescopes, which are limited only by the instrument performance and the diffraction of their primary apertures. A new wave of ground-based telescopes is currently being planned for the near future,¹⁻³ but further research must be done on the performance of current AO systems to both avoid known pitfalls and improve upon current performance levels. Once put into practice,

*Address all correspondence to Emily Ramey, emily_ramey@berkeley.edu

Table 1 Most relevant parameters used in the analysis of the Keck-II AO system. Display names are in bold.

Source	Variable name	Units	Description
NIRC2 image	strehl	n/a	Strehl Ratio : ratio of peak intensity of a point source to that of a diffraction-limited point source of the same magnitude
	FWHM	mas	FWHM intensity of the point spread function of a point source
NIRC2 header	az	deg	Telescope Azimuth
	Airmass	n/a	Integrated air density along the line of sight
	aolbfwhm	arcsec	Low-bandwidth wavefront sensor FWHM (LB-FWHM)
	lsamppwr	Watts	Laser Power of the guide star
	lgrmswf	nm	RMS WF Residual on the low-bandwidth wavefront sensor
	aoaomed	ADN	WFS AOA camera Median Subaperture Light
	tubetemp	°C	Telescope Tube Temperature : average of 15 sensors on the telescope tube/structure
MKWC seeing	MASS	arcsec	Integrated MASS seeing
	MASSPRO	$m^{1/3}$	Average C_n^2 seeing profile
	MASSPRO_half, _1, _2, _4, _8, _16	$m^{1/3}$	C_n^2 profile at 0.5, 1, 2, 4, 8, and 16 km above the telescope pupil (MASSPRO 0.5 km, 1 km, etc.)
	DIMM	arcsec	DIMM seeing
MKWC weather (CFHT)	wind_speed	m/s	Mean Wind Speed
	wind_direction	deg	Mean Wind Direction in degrees from north
	Temperature	°C	Air temperature at the CFHT weather station
	relative_humidity	%	Relative Humidity as a fraction of the saturation point
	Pressure	millibar	Mean atmospheric pressure
K2AO temperature	k2ao_nirc2_temp	°C	AO bench temperature near NIRC2 (AO Bench: NIRC2).
	k2ao_wfs_temp		AO bench temperature near the WFS (AO Bench: WFS).
K2ENV temperature	k0_temperature		Outdoor Air temperature at Keck Observatory
	k2_dome_temp		Dome Air temperature at Keck-II
	k2_primary_temp		Temperature of the Keck-II Primary mirror
	k2_secondary_temp		Temperature on the Keck-II Secondary mirror structure

such research could improve the long-term stability of AO Systems in general, which is crucial for fields or applications that require observations to take place over the span of years or decades. Any long-term astronomical survey making use of AO could benefit greatly from improved stability, with specific applications including microlensing,⁴ exoplanet imaging,⁵ observations of low-mass binary stars,⁶ and Galactic Center (GC) research.⁷ Though the precise results of this analysis will be specific to the Keck-II telescope at the W. M. Keck Observatory, broader trends in our findings could shed light on other AO systems in use today.

The Keck-II telescope is one of the largest ground-based telescopes currently in operation, and it also has one of the longest-running astronomical applications of an LGS-AO system in the world,^{8,9} which presents a good opportunity to study the performance of an AO system on a long timescale and under a multitude of different weather and seeing conditions. The goals of this project are twofold: the first is to analyze the performance of the Keck-II LGS-AO system to determine potential problems or improvements, and the second is to develop a preliminary predictive tool for the AO-corrected image quality of an observation given real-time weather and seeing conditions. To accomplish these goals, we use LGS-AO observations of the GC taken with Keck-II's near-infrared camera (NIRC2), along with the corresponding weather and seeing information from the Mauna Kea Weather Center (MKWC),¹⁰ AO telemetry data from Keck servers, and temperature measurements from internal Keck data to characterize current AO performance. We then use weather and seeing information to train and test basic performance prediction algorithms for accuracy, with the NIRC2 Strehl ratio and full width at half maximum (FWHM) used as targets for image quality.

Similar analyses of AO performance have been done on the Gemini Planet Imager (GPI) (see Refs. 11 and 12) specifically looking at the effects of certain variables on the resulting image quality. In another study (see Ref. 13), machine learning (ML) techniques are used to predict Gemini image quality given environmental and observational factors. Though exact results are specific to each telescope, these papers provide a starting point for the analysis in this research. As in Ref. 13, our strategy is to first find correlations and connections between environmental factors in the data and the delivered image quality, and then to use a subset of this data to train and test quality prediction algorithms. Though there have been other in-depth studies of AO performance,^{14–16} at the time of this paper, and to the authors' knowledge, there has been no tool developed to predict the quality of LGS-AO observations in real-time for the Keck-II telescope. If successfully implemented, such a tool would be very useful for observation planning, allowing the prioritization of targets with higher predicted image quality on a given night over those which might yield poorer scientific results.

2 Observations

2.1 NIRC2 Data

The primary data used in this project are images of the GC obtained using the Keck-II LGS-AO system and the NIRC2 camera (PI: K. Matthews), which is specifically designed for use with AO. The observations are part of a long-running survey to measure the orbits of stars around the GC over a period of roughly 14 years (2005–2019).¹⁷ Images used in this research are of the central 10'' region of the GC, approximately centered on Sagittarius A*^{7,17,18} (ICRS Coordinates: RA 17 45 40.03599 DEC -29 00 28.1699). All images are taken in the Kp-band (center $\lambda = 2.124 \mu\text{m}$ and bandpass = $0.351 \mu\text{m}$) and in the narrow-field mode of NIRC2. With these settings, NIRC2 has a field of view of $\sim 10'' \times 10''$. The plate scale was 9.952 mas/pixel until 2014, after which the AO system and NIRC2 camera were realigned, changing the plate scale to 9.971 mas/pixel.¹⁹ We use cleaned frames processed through the standard NIRC2 reduction pipeline, which corrects for geometric distortion and differential atmospheric refraction. The Strehl ratio and FWHM are calculated from the image and header data using Keck Observatory's LGS-AO data reduction and analysis software.²⁰ These values are good measures of the delivered image quality and are used frequently as indicator/target variables in both our analysis of the full dataset and in ML testing. More information about the observational

parameters and processing of NIRC2 images can be found in Refs. 8, 9, 18, 19, 21–26. A summary of the dataset can be found in Table 2.

2.2 CFHT Data

The meteorological data for this study come from the weather tower at Canada–France–Hawaii Telescope (CFHT) and can be found on the MKWC website,¹⁰ with samples taken every minute. We gather metrics such as pressure, temperature, wind speed and direction, and relative humidity, and use these in both analyzing current system performance and in training the predictive algorithm.

2.3 Seeing Data

The seeing data studied here also come from the MKWC archives.¹⁰ This seeing information includes both multiaperture scintillation sensor (MASS) and differential imaging motion monitor (DIMM) readings, which measure the atmospheric seeing, or turbulence strength, in different ways. The seeing is recorded as the atmospheric induced image blurring, in arcseconds of angular resolution, at a wavelength of 500 nm and at zenith.²⁷ The MASS instrument makes measurements based on the principle that the spatial scale of intensity fluctuations changes with the height of the atmospheric layer that produces them. It uses four concentric apertures to measure the intensity fluctuations for a star near the zenith and fits a six-layer atmospheric profile, $C_n^2(z)$, to the data by assuming a Kolmogorov turbulence model and that the scale of the turbulence is small. The results are then corrected for zenith angle (ζ) and reported at $\zeta = 0$. It should be noted that the MASS instrument does not report $C_n^2(z)$ directly, but instead the integrated turbulence in each layer in the free atmosphere (>0.5 km)²⁷

$$\text{MASS}_i = \int_{z_i, \min}^{z_i, \max} C_n^2(z) w(z) dz, \quad (1)$$

where i is the layer measured, z is the elevation, and $w(z)$ is the idealized response function of the MASS instrument. Turbulence measurements for each layer are reported in units of $m^{-1/3}$, and an aggregate MASS statistic, which represents the free atmosphere seeing, is reported in arcseconds.

The DIMM instrument makes a much more straightforward measurement of the seeing, producing one value for the entire integrated atmosphere (up to ≈ 30 km). It measures the AO centroid fluctuations of a single star (not at zenith) and makes a correction for the zenith seeing.²⁷ Results are reported as the FWHM of a point source at a wavelength of 500 nm (in arcseconds). When the MASS and DIMM instruments are in use, measurements are recorded every few minutes. This data are used for both analysis and prediction, as current conditions can be easily obtained from the MKWC website.¹⁰

2.4 Temperature Data

We were able to obtain more detailed temperature measurements from Keck Observatory for the years from 2000 to 2020, from which we extracted three sets of temperature measurements relevant to Keck-II: K2AOtemps (from the AO bench and electronics room), K2L4temps (from the laser enclosure), and K2envMet (from the dome, telescope, and surrounding environment). Note that the temperature sensor labeled for the secondary mirror (see Table 1) is actually on the telescope structure behind the mirror, and thus more closely tracks the structure/sky temperature rather than the temperature of the mirror itself. In addition, the NIRC2 and wavefront sensor (WFS) temperatures come from sensors on the Keck-II AO bench near these instruments rather than the instruments themselves. This data are used in the analysis portion of the project to follow up on our initial findings in Sec. 3.2. However, such detailed temperature readings are not readily available to observers at Keck, and as such must be omitted from the training of the predictive algorithm.

Table 2 Table of observational data, calculated from the analysis sample. Strehl ratio, FWHM, MASS, and DIMM represent the median quantities for the observing run.

Date (UT)	Strehl ratio	FWHM (mas)	MASS (as)	DIMM (as)
July 26, 2004	0.36	58.50	None	None
July 30, 2005 and July 1, 2005	0.21	60.88	None	None
June 30, 2005	0.34	60.81	None	None
July 17, 2006	0.36	58.11	None	None
June 20, 2006 and June 21, 2006	0.31	61.67	None	None
May 3, 2006	0.34	58.50	None	None
May 21, 2006	0.24	67.03	None	None
August 10, 2007 and August 12, 2007	0.26	62.86	None	None
May 17, 2007	0.34	59.64	None	None
July 24, 2008	0.29	63.30	None	None
May 15, 2008	0.30	53.66	None	None
July 24, 2009	0.24	66.80	None	None
May 1, 2009; May 2, 2009; May 4, 2009	0.32	63.35	None	None
September 9, 2009	0.32	61.51	None	None
August 15, 2010	0.30	58.54	0.11	0.34
July 6, 2010	0.31	62.10	0.15	0.39
May 4, 2010 and May 5, 2010	0.30	63.19	0.14	0.55
August 23, 2011 and August 24, 2011	0.33	60.82	0.20	0.49
May 27, 2011	0.26	66.42	None	None
July 24, 2012	0.33	60.13	0.10	0.48
May 15, 2012 and May 18, 2012	0.29	61.97	None	0.41
April 26, 2013 and April 27, 2013	0.23	72.52	0.21	0.52
July 20, 2013	0.34	59.82	None	None
July 31, 2013	0.09	106.28	None	None
April 18, 2014	0.18	88.94	None	1.33
April 19, 2014	0.10	123.12	1.29	1.27
August 4, 2014	0.26	66.08	0.40	0.94
August 6, 2014	0.35	57.23	0.08	0.74
July 3, 2014	0.23	74.09	0.19	0.57
July 4, 2014	0.19	78.90	0.41	0.73
Mar 19, 2014	0.10	102.46	0.57	1.04
Mar 20, 2014	0.22	68.27	0.59	0.55
May 11, 2014	0.25	68.37	0.26	1.02

Table 2 (*Continued*).

Date (UT)	Strehl ratio	FWHM (mas)	MASS (as)	DIMM (as)
May 19, 2014	0.29	65.15	0.18	0.60
August 9, 2015, August 10, 2015, and August 11, 2015	0.35	59.79	0.18	0.56
March 31, 2015	0.20	76.61	0.22	0.38
April 1, 2015 and April 2, 2015	0.22	72.90	0.18	0.43
May 14, 2015	0.13	93.24	0.60	0.52
May 3, 2016	0.31	64.08	0.20	0.47
May 17, 2016	0.25	68.79	0.28	0.55
August 8, 2017, August 9, 2017, August 10, 2017, and August 11, 2017	0.31	60.57	0.18	0.54
August 23, 2017, August 24, 2017, and August 26, 2017	0.28	66.07	0.23	0.64
July 18, 2017	0.27	65.10	0.17	0.36
July 27, 2017	0.13	92.90	None	0.59
May 4, 2017 and May 5, 2017	0.28	66.19	None	0.50

3 Methods

3.1 Data Compilation

As NIRC2 FITS files are the primary observations for this study, all other data are secondary and must be cross-matched with these images. Telemetry files are unique for each observation, but weather, seeing, and temperature data are taken periodically on different timescales, and as such their timestamps do not always match up exactly with NIRC2 observations. For a first pass at the data, we match each NIRC2 file with a telemetry file (if available) and pull the closest data points available within a day for all other secondary observations. The relevant data columns from each observation type are then extracted to a separate file as a pandas dataframe.²⁸ Any data columns that cannot be found for a specific NIRC2 file are initially recorded as null.

To clean the data, we first remove rows with negative Strehl ratios or with FWHM values below 50 mas or above 150 mas, as these are nonphysical or very poor quality for NIRC2 LGS-AO images of the GC in the near-IR. Some NIRC2 observations are unable to be paired with secondary readings (e.g., observations that predate the seeing/temperature instruments, had problems with the weather tower, or have no associated telemetry), and this results in null values in weather, seeing telemetry, or temperature fields. Additionally, any secondary observations taken >5 min from the primary observations are considered invalid and changed to null when filtering the data table. Rows with a number of invalid or missing secondary observations are still considered in our analysis of system performance (Sec. 3.2), but due to the nature of our chosen ML models (described in Sec. 3.3), any entries with missing values in weather or seeing columns must be taken out or replaced before training the predictive algorithm. Histograms of the most relevant variables, indicating which observations compose the analysis and prediction samples, are shown in Fig. 1.

3.2 Performance Analysis

We first plot the nightly Strehl ratio and FWHM values over the full GC survey in Fig. 2, with green arrows indicating relevant entries in the general Keck AO engineering logs. As the figure

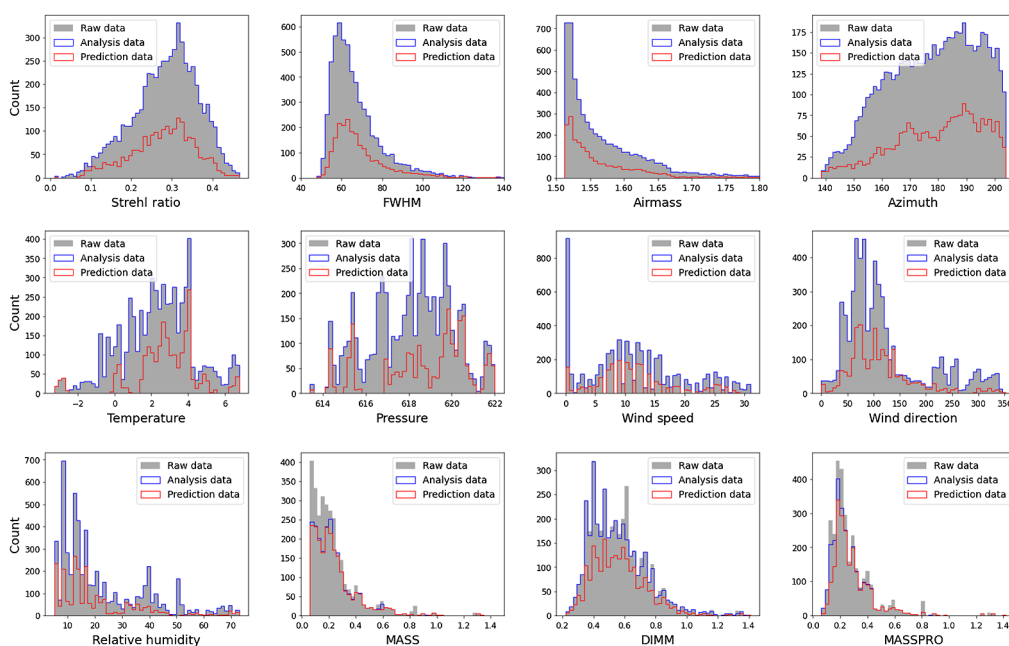


Fig. 1 Histograms of feature data before and after filtering. The analysis sample has all invalid values replaced with null, while the prediction sample has all rows with null or invalid values removed. Note: the MASS, DIMM, and MASSPRO panels represent a subsample (40% to 60%) of the full analysis dataset where these values were available.

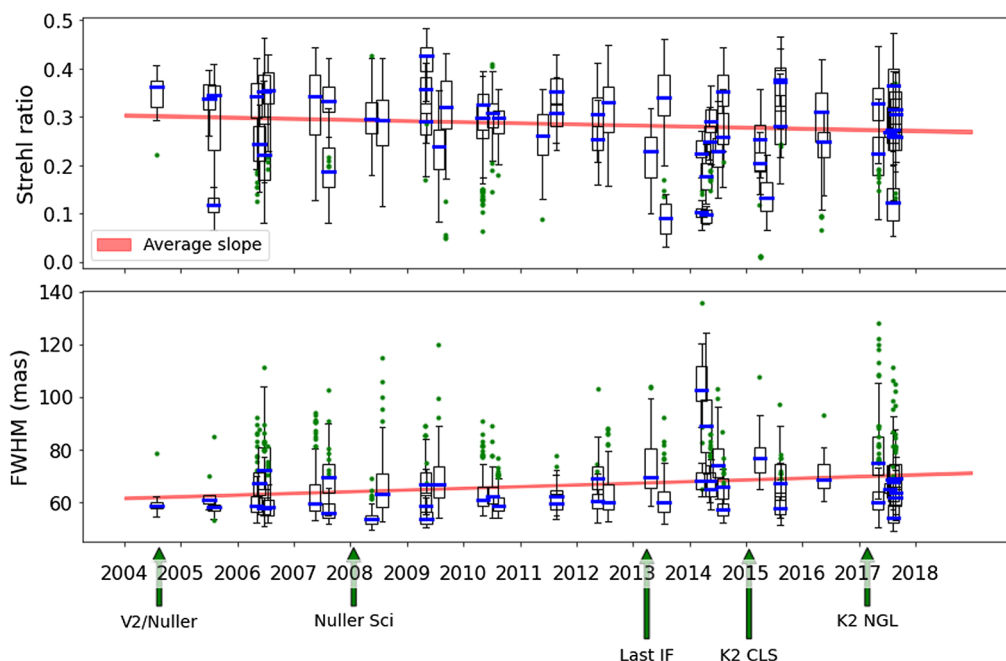


Fig. 2 Image quality on each observing night over the full GC survey. Boxes indicate the first and third quartiles, with green points showing outliers and medians marked in blue. Annotations indicate entries in the general Keck AO engineering logs over the period of observation (see Table 5 for log entry details). The red slope lines and error regions represent a simple linear least-squares fit of the data using `numpy.polyfit`. These show that average image quality decreases over the survey period.

shows, there has been a gradual trend toward poorer image quality since the start of the survey, with more outliers in recent years compared with previous observations (see Sec. 4.1 for more detailed information). This indicates that the performance of the system has degraded over the past decade, despite multiple improvements being made to instrumentation^{29–33} (see Figs. 2 and 9 for AO engineering log entries). The origin of this phenomenon is unknown, but potential causes include long-term changes in atmospheric seeing or weather, issues with telescope alignment and phasing operations, or degradation of AO system performance over time. This will be explored further in Sec. 4.1. This long-term trend should additionally be verified with other instruments on Keck-II, or with data from other surveys, to confirm the scope and significance of the issue. We also perform a correlation analysis of all relevant variables with the Strehl ratio and the FWHM, shown in Figs. 3 and 7. All correlation values represent Pearson coefficients [Eq. (2)] unless otherwise specified

$$\text{Pearson } r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (2)$$

High-altitude seeing measurements, measured by the MASS and MASSPRO, show a strong correlation with the image quality. This is expected, as turbulence in higher layers of the atmosphere is more difficult for AO systems to correct. However, we can also see that the outdoor air temperature and telescope tube temperature are correlated with image quality ($r \sim +0.3$ for Strehl ratio), which is an unexpected result, as this ranks just behind the seeing in terms of correlation strength. We confirm these correlations visually by plotting the relevant variables versus image quality in Figs. 4–6. The LGS-AO system should theoretically be able to correct

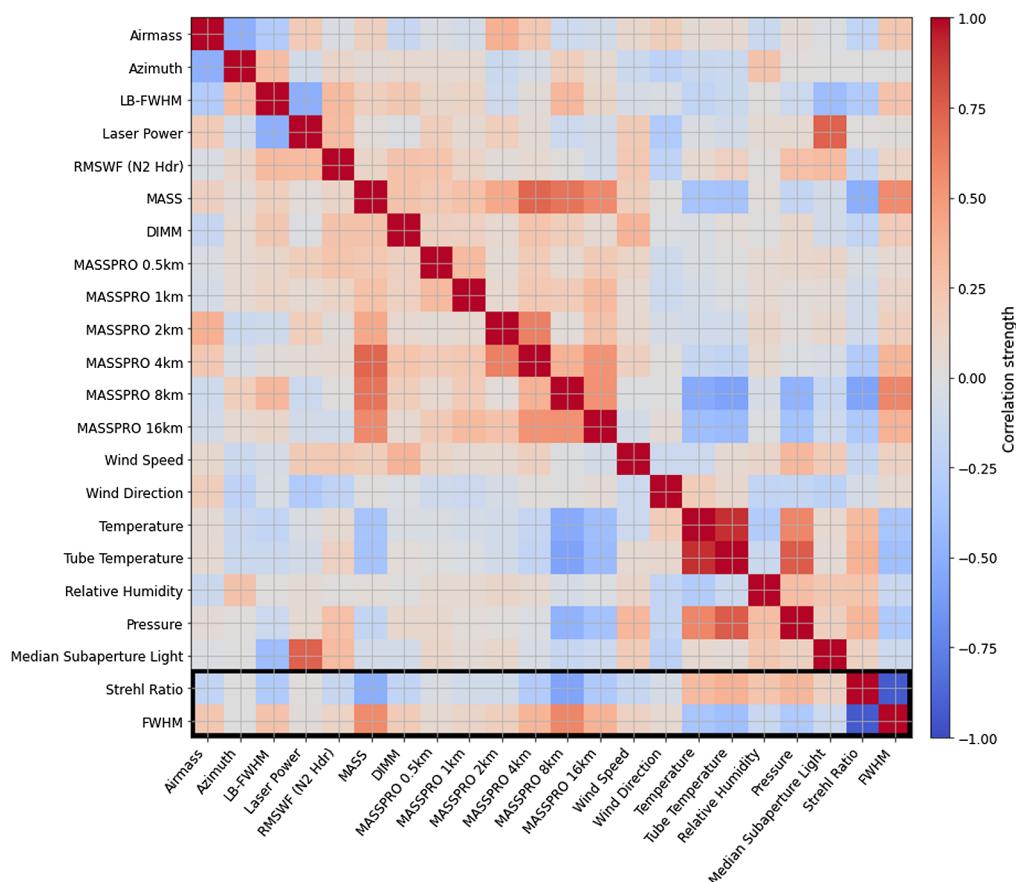


Fig. 3 Correlation plot of relevant variables. Strong positive correlations are shown in red while strong negative correlations are shown in blue. The strongest correlators with the Strehl ratio and FWHM are the MASS/MASSPRO seeing, outdoor temperature, and tube temperature.

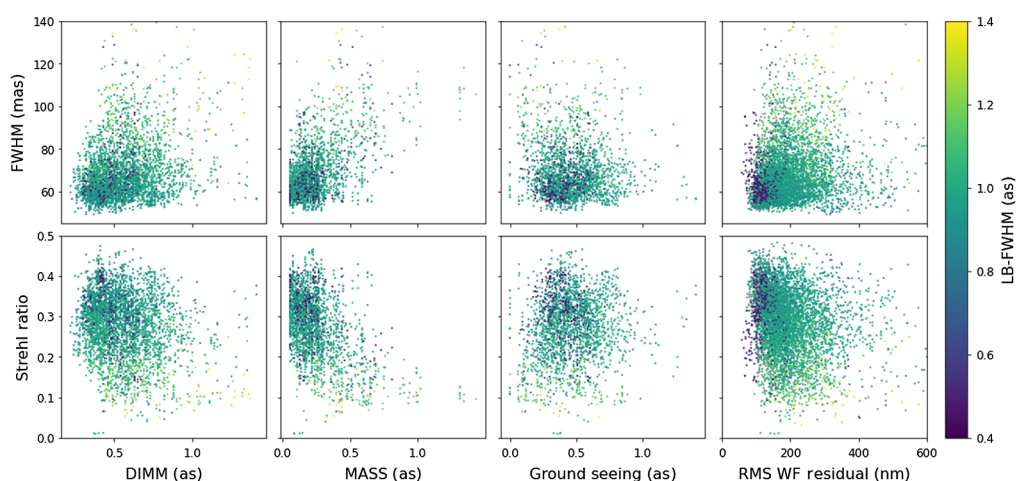


Fig. 4 Plot of variables related to seeing and AO correction. As expected, most seeing variables correlate with image quality.

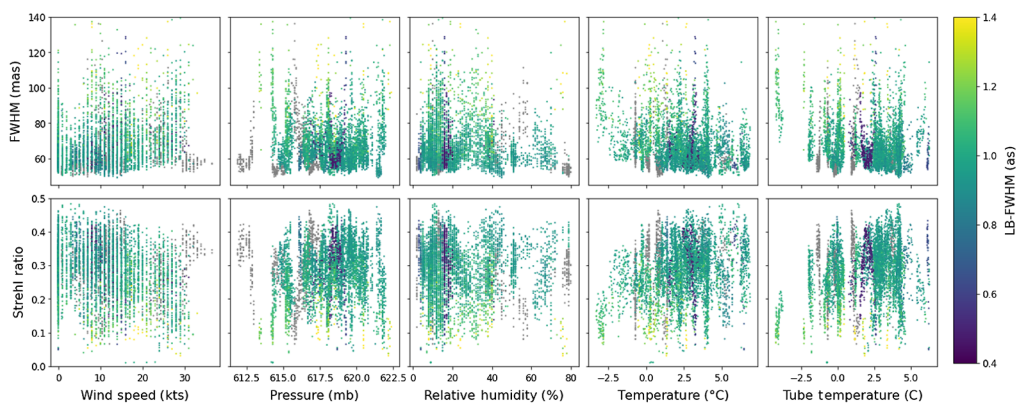


Fig. 5 Plot of variables relating to the weather conditions atop Mauna Kea during sample observations. The image quality shows a moderate linear dependence on both the outdoor and telescope tube temperatures ($r \sim +0.3$ for Strehl, -0.3 for FWHM).

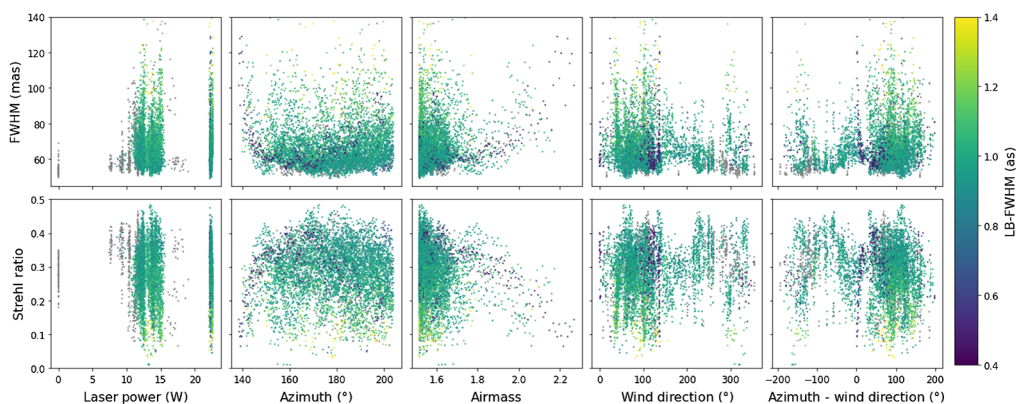


Fig. 6 Plot of observational parameters versus Strehl ratio and FWHM. None of these correlate very strongly with image quality.

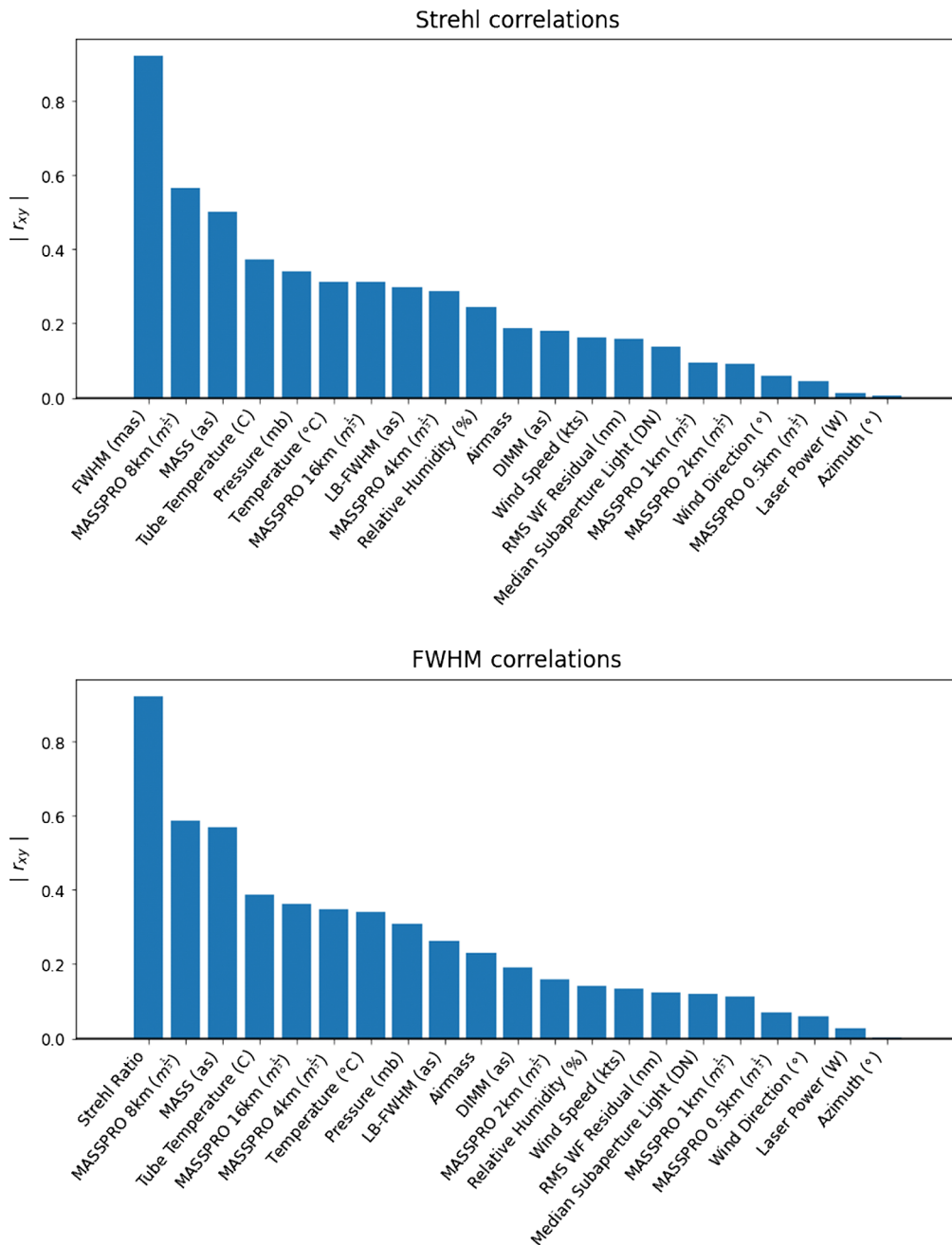


Fig. 7 Flattened Pearson correlation plot between image quality metrics and studied variables.

for outdoor turbulence at any temperature, so there could be an internal effect causing the system performance to degrade at lower temperatures. This will be explored further in Secs. 4.2 and 5. Figure 6 includes the difference between the telescope's azimuth and the wind direction in order to check whether directional wind effects (vibrations, eddy currents, etc.) could be causing performance issues. Keck-II currently does not have accelerometers installed, but if vibrational effects were an issue, we would expect to see a stronger correlation between the telescope's angle relative to the wind and the image quality in the farthest right plots of Fig. 6. As there is not a strong correlation between these two variables, we choose to focus on other potential issues affecting the image quality for the majority of this paper. We search for other correlations of note in Fig. 3, which shows the Pearson coefficient of each pair of variables as a heat map, but do not find any others that appear surprising or out of place.

Table 3 Feature and target variables used to train predictive algorithms.

Features	Targets
Wind speed	Strehl ratio
Wind direction	FWHM
Temperature	
Humidity	
Pressure	
Azimuth	
Airmass	
MASS seeing	
DIMM seeing	

3.3 Predictive Algorithm

The creation of an image-quality prediction algorithm necessitates different data processing methods than those used in Sec. 3.2: a predictive algorithm can only be trained on data available before an observation is made, with the exception of its targets (Strehl ratio and FWHM). For example, if we used telemetry files to train an algorithm, we would need to have relevant telemetry files available every time we wanted to predict the current image quality. As those files are only available after an observation is made, however, the prediction would be rendered useless, as the true Strehl ratio and FWHM would already be determined by the observation itself. As such, we are limited to data from the CFHT and MASS/DIMM files in building and training our ML model, since these can be obtained in real-time. From the weather and seeing data, we extract only those variables that we expect to have a high impact on overall image quality, shown in the left column of Table 3. This data then become the input to our ML algorithm, also known as a feature matrix. We continue to use the Strehl ratio and FWHM as metrics for the image quality, which become the target variables for our algorithm.

3.3.1 Data transformation

We first standardize the data, which set the mean of each feature/column to zero and the standard deviation to one. This is a way to save processing time in models that use a method known as gradient descent in training, and it is an essential step before a principal component analysis (PCA).³⁴ The PCA algorithm returns a basis set of orthogonal components, found by taking the eigenvalues of the feature correlation matrix. Components are listed by their *explained variance*, or how to spread out the data are along with the given component, and the *explained variance ratio* [Fig. 8(a)] refers to the explained variance of one component divided by the sum total of explained variances among all the components. While this cannot predict anything from the data on its own, as it does not use the target variables, it has the potential to condense the feature matrix into only a few of its components without losing a significant amount of information. In addition, certain ML algorithms rely on input features being uncorrelated, which can also be achieved through PCA.

We start with a linear PCA—using the feature data as-is, rather than transforming it into a higher-dimensional space first. To eliminate a component, we need its explained variance ratio to be below the desired threshold (we choose 1%). As shown in Fig. 8(a), none of the components has an explained variance ratio below this threshold, as removing the last component would still eliminate >1% of the information contained in our data. This implies that the data are already fairly uncorrelated in the linear regime, although higher-order dependencies are still possible.

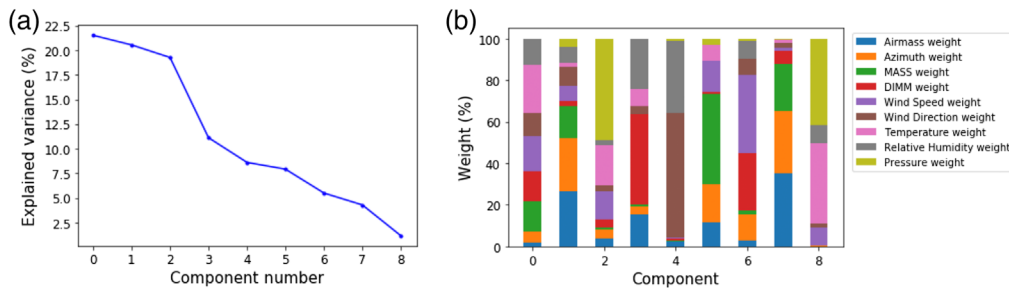


Fig. 8 (a) Explained variance ratios and (b) composition of components for a linear PCA on feature variables. No component explains <1% or >25% of the feature variance, making it difficult to remove any from consideration. The composition of the components reveals that temperature, wind speed/direction, and MASS/DIMM seeing are all large factors in the principal component (component 0), meaning they account for the most significant variation in feature data.

The composition of the components in terms of the original feature data is shown in Fig. 8(b). The principal component (component 0) shows that the temperature, wind speed/direction, and MASS/DIMM seeing all play large roles in the feature variance, which confirms the correlations found in Sec. 3.2. Based on the PCA, we do not eliminate any features from our data in this iteration of the project, although we do use these nine components in the training of ML models. More advanced methods of feature selection are possible, such as using a Kernel PCA instead of a linear one. However, as this is meant to be a relatively simple test of basic ML models, we leave these to be explored in future work.

3.3.2 Models tested

As we do not know what form the final algorithm should take, we test several ML models, largely chosen for their simplicity and/or adaptability to our problem, as well as their value as proofs-of-concept. We have ~6000 data points to start with, which reduce to ~2000 after the data set is cleaned. The amount of data is much greater than the number of features, so overfitting is less of a problem, and we are less limited in our choice of models as a result. We select three algorithms to start with: random forest regression, support vector regression, and an artificial neural network (ANN).

*Random forest regression*³⁵ is particularly suited to our problem, as it is able to model non-linear relationships between feature and target data with relative ease. It relies on a simple decision tree as its most basic predictor, and each decision tree is trained on a slightly different subset of the data to predict the image quality. The individual results from each decision tree are then averaged to give the final target value. This is a method that is generally very accurate, while also being fairly robust to outliers and overfitting, so it serves as a good test case for our research goals.

*Support vector regression*³⁶ is an adaptation of a simple linear regression, but it makes use of a mathematical technique to maximize the “margins” between the fitted function and the spread of the data. This makes it more robust than a simple linear regression, although both are straightforward to implement with the sklearn package. This model is potentially useful because of a mathematical construct known as the *kernel trick*, which allows it to fit nonlinear functions with a relatively short computation time. One potential drawback to this method, however, is that, to use the kernel trick, one must know the approximate shape the solution should take to get the best results (e.g., a sine wave, a fourth-order polynomial, etc.).

*ANNs*³⁷ are a very powerful tool in ML. They use a series of layers to reduce data, with each layer containing individual “neurons.” Each layer has a series of weights, which are learned in the training process, and a propagation function to transmit data between neurons. Once the desired performance has been reached, the weights are fixed and the network is used only for prediction. This method is intended to solve problems the way a human or animal brain does and is thus one of the more complicated ML models available. Though it is very

powerful, it has the potential to pick up on correlations that result from noise or bad data, and great care must be taken in the training and testing process in order to avoid overfitting. With the right controls, however, it is possible to get an algorithm that will generalize well to real-world problems.

3.3.3 Model selection and validation

To compare and evaluate these algorithms, we test the accuracy of each one on data that was not used in the training process. We first split the data into training and testing sets and train each model using the training set only. The model is then used to predict the targets for the test set, and the results are compared with the true values to get an estimate of the *test error*, also known as the out-of-sample error or the generalization error, which approximates how the model will perform in reality. We can take this one step further by using the *cross-validation error*, which involves splitting the data into a number of equal sections (typically 5) and training/testing the algorithm on each resulting train/test split. The results of each test run are then averaged to get the final cross-validation error. This is generally a more robust validation method than simply using the test error, as it is less susceptible to random chance and can provide an estimate of the variability of the algorithm's performance. However, all validation methods generally tend to overestimate the true error slightly, as they rely on artificially limiting the size of the training set, which tends to make the algorithm less accurate. This introduces a trade-off in computation time versus precision of results, as using more splits in cross-validation is statistically more informative but also more complex.

To compare between models, we start by reporting the results of the most successful version of each algorithm on a simple 80% to 20% train–test split (see Table 8). Finding the most successful version of each type of predictor is done using a grid search algorithm.³⁸ We use the GridSearchCV function from the sklearn package,³⁹ which internally implements *k*-fold cross-validation. A grid search is used to determine the optimal *hyperparameters* for each algorithm, training multiple different versions on slightly different parameter sets until the test error is minimized. For example, a random forest algorithm includes such hyperparameters as the number of decision trees used, the maximum depth of each decision tree, and the maximum number of features that can be considered by any individual tree, and changing any of these can impact the accuracy of the algorithm. A list of the hyperparameter values searched for each algorithm studied here can be found in Table 4. The Grid Search method returns the version of each algorithm with the lowest cross-validation error after it has searched all sets of hyperparameters. Once the best version of each algorithm is found, we test it on the remaining 20% of data and evaluate the results (see Table 8).

Table 4 Grid search parameters for each ML algorithm tested. More information can be found in scikit-learn's API Ref. 40.

Algorithm	Parameter	Values	Step
Random forest	n_estimators	5 to 500	10
	max_depth	5 to 100	5
	max_features	1-#features	1
SVR	deg	1 to 10	1
	C	0.1 to 5	0.1
ANN	hidden_layer_sizes	(1) to (10), (1,1) to (20,20)	1

Table 5 Description of Keck AO engineering terms from Figs. 2 and 9.

AO log entry	Description
V2/Nuller	Nulling interferometer integration with AO system
Nuller Sci	Interferometer used for science
Last IF	Interferometer no longer used
hline Rotator	Field derotator (K-mirror)
IR camera	Refers to the IR boresight camera on the secondary mirror
SFP	Simulator fiber positioner
WYKO	Phase-shifting interferometer to measure DM flatness
LBWFS	Low-bandwidth wavefront sensor
ACAM	Optical acquisition camera for the AO system
WFS	Fast wavefront sensor
NGL	Next-generation laser
CLS	Center launch laser
TOPTICA laser	Next-generation sodium laser (from TOPTICA photonics)

4 Results

4.1 Trends in Image Quality

As shown in Fig. 2 and Table 6, Strehl and FWHM values on an average observing night have degraded over time. Based on a simple linear fit (`numpy.polyfit`⁴⁰), the average Strehl has decreased by roughly 3% and the average FWHM has increased by roughly 9 mas over the course of this 14-year survey. This degradation in quality happens despite multiple system improvements being made to Keck-II and its AO system^{29–33} in the intervening years. As this data have been taken by one research group, on one instrument, and with one target, it is ideal for exploring long-term trends in AO performance. However, the lack of data diversity (e.g., in guide star brightness, target location, etc.) also makes it difficult to determine the source of this

Table 6 Slopes and slope errors of relevant image quality and seeing parameters over time. Strehl, FWHM, MASS, and DIMM, all show degradation, although temperature increases.

	Slope	Slope error
Strehl ratio (year^{-1})	−0.0023	0.0003
FWHM (mas/year)	0.6385	0.0454
MASS (as/year)	0.0117	0.0016
DIMM (as/year)	0.0186	0.0012
Temperature (deg)	0.0776	0.0071
Airmass (year^{-1})	0.0027	0.0003

apparent degradation. A closer examination of the available data may still allow us to propose several potential underlying causes as a starting point for further research.

4.1.1 Instrumental changes

If the cause of the degradation is instrumental, it could be reflected in Keck AO engineering log entries, which we have included in Fig. 2 as green arrows. More detailed logs are included in Fig. 9, with a zoomed-in portion of the timeline, for reference. As the installation of Keck-II's center-launch system (CLS) in 2014³⁰ occurred near a period of increased variance in image quality, it may be a potential cause of the degradation. While a CLS provides a more concentrated and better-shaped guide star than a side-launch laser, the signal-to-noise ratio may be reduced due to laser light scattering back into the telescope.³⁰ The CLS also requires more optical components than a side-launch laser, theoretically making the laser's return power lower than it would be otherwise.³⁰ This is confirmed in an internal study of the CLS performance done at Keck,⁴¹ which shows a dimming of the laser by 0.5 magnitudes. While this decrease in return power was mitigated by the installation of a more powerful laser (TOPTICA) in 2016,³¹ there was not a proportionate increase seen in image quality compared with data taken pre-CLS.

Another internal study done after the new laser was installed⁴² only compares the on-sky performance from post-CLS observations, and reports nearly the same average image quality after the CLS and TOPTICA installations as that found in a similar study in 2008.^{32,43} (All images were taken on NIRC2 and in the K-band, though there may have been differences in methods.) It is possible, therefore, given the data used in this study and that found in previous performance reviews, that the upgrades installed on Keck-II since 2007 have not only failed to improve the on-sky performance, but may have even degraded it in some cases. If this is the case, a solution may be just around the corner. Current estimates show that an upgrade to the real-time

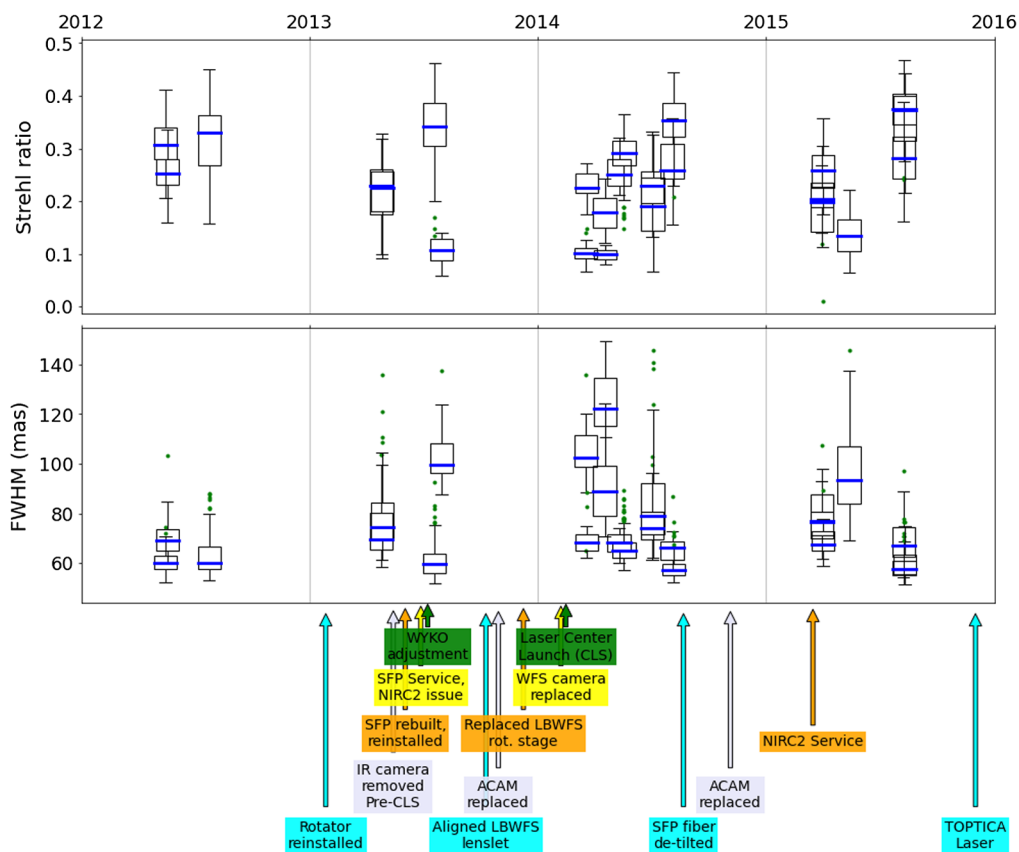


Fig. 9 Image quality per observing night over the full data sample. Arrows and annotations represent entries from Keck Observatory's detailed AO engineering logs.

controller (RTC) may be necessary to fully realize the gains of the brighter TOPTICA laser, and this will likely increase the on-sky performance beyond its current level. An RTC upgrade is currently underway at Keck-II⁴⁴ and is planned to be completed in 2022.

4.1.2 Atmospheric trends

It is possible, however, that this phenomenon is not related to instrumentation at all, but rather environmental factors. As shown in Table 6, the MASS and DIMM seeing have similar trends overtime to those seen in the Strehl Ratio and FWHM. The MASS seeing increases by 163 mas and the DIMM seeing increases by 260 mas over the course of the survey, meaning an observing run today is more likely to have poor seeing than an observing run at the start of the survey. As we know the image quality to be highly correlated with seeing, this could be the only required explanation for performance degradation over time, although it would be difficult to determine this solely from the data available. In fitting a line to the temperature over time, however, we find that temperature increases over the survey, which should cause the image quality to go up slightly if the correlations found in Sec. 3.2 are correct. A more formal analysis of temporal changes in image quality, controlling for seeing and temperature, would be able to shed more light on the situation, although we leave this investigation for future work.

4.1.3 Observational or experimental factors

Although there are over 5000 individual images in this sample, they represent <50 observing runs in total, all taken during the summer months. With such a relatively small number of truly independent samples, it is possible that several nights or weeks of bad seeing and weather, or simple human errors during observing could throw off the dataset as a whole and weigh the overall trend line toward poorer results. A similar analysis of other long-running image surveys (from Keck or another ground-based telescope) could circumvent this issue. However, it is difficult to find long-running surveys such as this one with controlled targets, exposure time, guide star brightness, etc., and it is the stability of all of these factors that have allowed us to perform such a detailed correlation analysis in this case.

4.2 Temperature Correlations

To confirm our previous findings from MKWC temperature readings, we perform a correlation analysis on more detailed temperature data from the Keck dome, telescope, and AO bench (described in Sec. 2.4). We plot the correlation matrix of the new temperature data in Fig. 10, including temperature differences with the outside air. The top 15 largest correlations with image quality, among individual temperatures or temperature differences, are shown in Table 7, and temperatures for several example nights of observation are shown in Fig. 11. There are several potential sources of turbulence that we investigated in this part of the project. Since the image quality improves as the outdoor and tube temperatures get warmer, it is likely that the issue stems from a source that is routinely warmer than its surroundings, although other scenarios are possible.

4.2.1 Dome seeing

If the air in the telescope dome were routinely warmer than the outdoor air, it would cause convection in the dome itself, which would be difficult to correct with AO. Based on Table 7, the image quality is more strongly correlated with the dome air temperature than it is with most other temperatures or temperature differences, including the dome-outdoor temperature difference. This indicates that the dome air plays a large role in the image quality, although the direction of the correlation is not as expected. As a general trend, the Strehl ratio improves as the dome air gets warmer, which indicates that dome seeing is not necessarily the problem here. As only the temperature correlations are taken into account in Table 7, underlying factors such as mirror or instrument temperatures could be driving this trend without having an obvious impact on image quality.

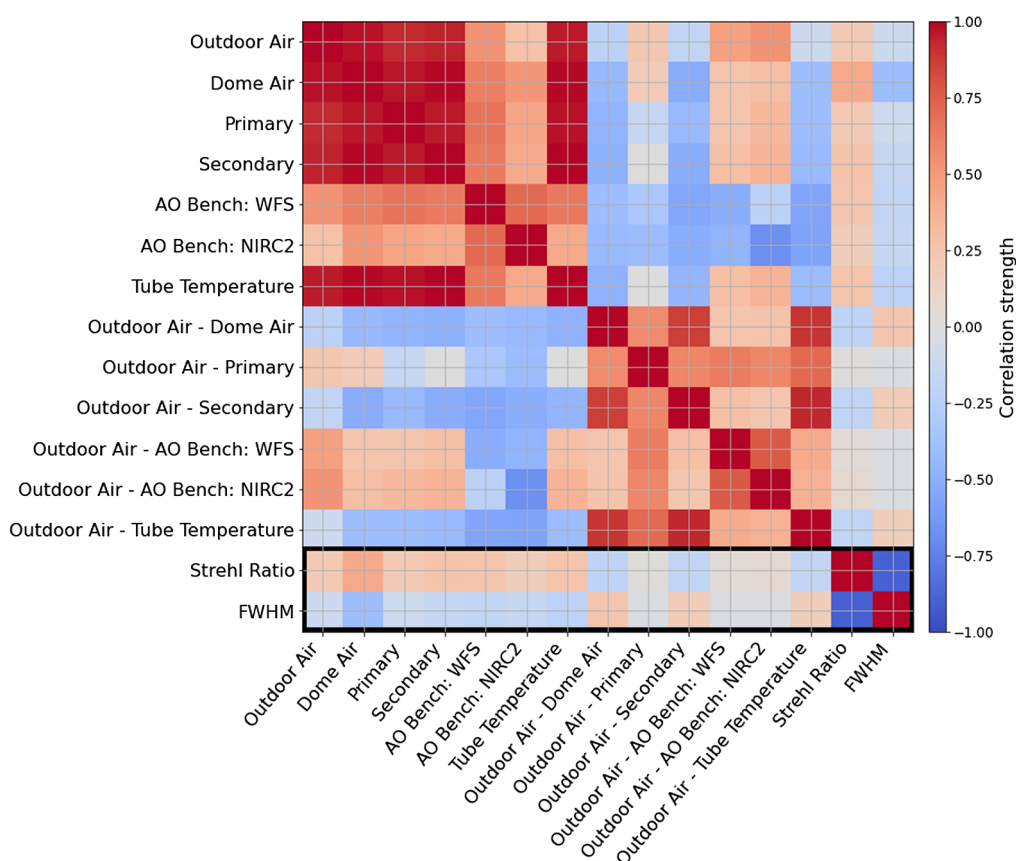


Fig. 10 Correlations among temperature variables from the Keck-II dome, telescope, and AO instruments. The largest correlation comes from the dome temperature.

4.2.2 Mirror seeing

The difference in temperature between the primary mirror and the dome air is also moderately correlated with the image quality, which could indicate mirror seeing as a potential issue. Due to its size, the primary mirror is more resistant to changes in temperature than the dome air, as we can see in Fig. 11. The primary temperature remains mostly constant throughout each night, while the dome and outdoor temperatures show larger fluctuations. This could lead to a scenario in which the primary mirror maintains a higher temperature than the dome air during certain parts of the night, causing turbulence across its surface and degrading the image quality. Indeed, in Table 7, the dome-primary temperature difference ranks in the top three correlations for both Strehl ratio and FWHM ($r = 0.26$ to 0.29). The correlation direction indicates that the image quality degrades when the primary mirror is warmer than the dome air, which supports the hypothesis that mirror seeing is an issue impacting the image quality.

4.2.3 Low-wind effect

Another potential source of turbulence is the low-wind effect⁴⁵ (LWE), in which the telescope structure radiates heat faster than the dome or mirrors and cools to lower temperatures during the night. This can cause differing air temperatures across the telescope pupil, which, in the absence of other wind or convection effects in the dome, can degrade the quality of AO images. Although the telescope body is routinely colder than its surroundings (Fig. 11), we have not found the image quality to be strongly correlated with the wind speed (Fig. 5) as we would expect it to be if this effect were taking place. The correlation between the outdoor-tube temperature difference and image quality is also fairly low ($r = 0.17$ to 0.19), and this is true for the dome-tube temperature difference as well ($r = 0.13$ to 0.15). However, there are currently no temperature sensors

Table 7 Top 15 highest Pearson correlation coefficients between temperature readings and image quality. All temperature differences between components are also included in this analysis.

Temp. location(s)	<i>r</i>
FWHM coefficients	
Dome air	−0.42
Dome air–primary	−0.29
Dome air–AO bench: NIRC2	−0.26
Primary–tube temperature	0.24
Outdoor air–dome air	0.24
Primary–secondary	0.23
Tube temperature	−0.21
Dome air–AO bench: WFS	−0.21
AO bench: WFS	−0.19
Outdoor air–secondary	0.18
Secondary	−0.17
Outdoor air–tube temperature	0.17
AO bench: NIRC2	−0.15
Dome air–tube temperature	−0.13
Outdoor air	−0.12
Dome air–secondary	0.12
Strehl ratio coefficients	
Dome air	0.42
Dome air–AO bench: NIRC2	0.27
Dome air–primary	0.26
Secondary	0.25
Primary–tube temperature	−0.25
Tube temperature	0.24
Primary–secondary	−0.24
AO bench: WFS	0.23
Dome air–AO bench: WFS	0.21
Outdoor air–dome air	−0.21
Outdoor air	0.21
Primary	0.20
Outdoor air–secondary	−0.19
Outdoor air–tube temperature	−0.19
AO bench: NIRC2	0.18
Dome air–tube temperature	0.15

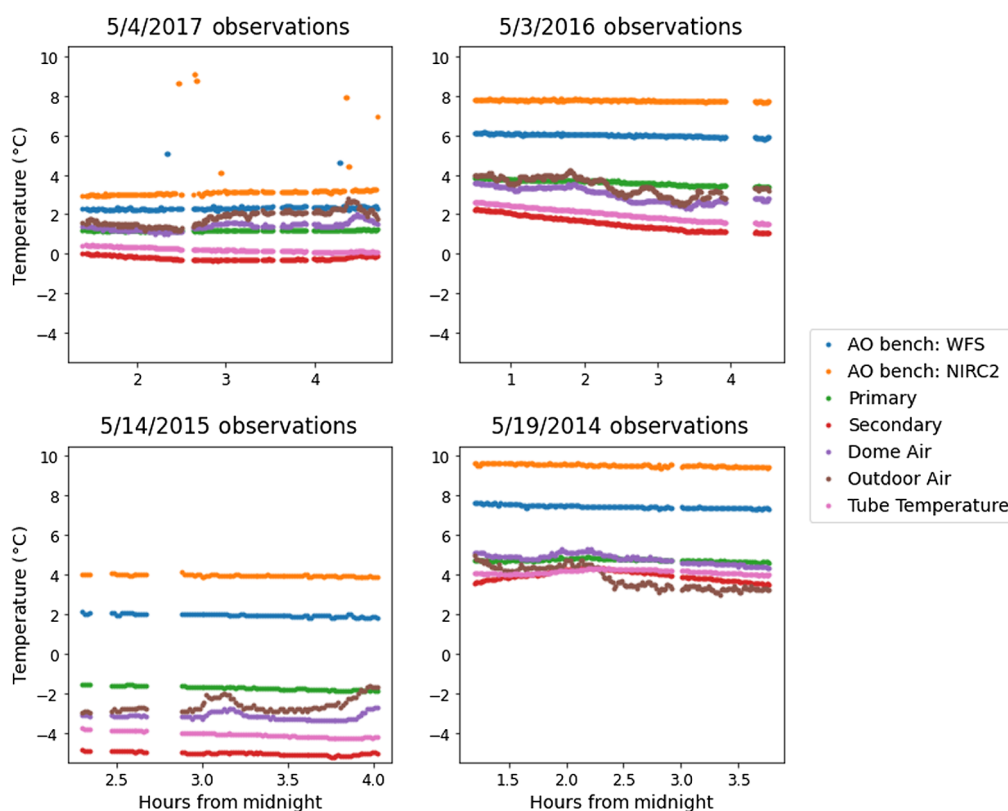


Fig. 11 Temperature measurements from the dome, outdoor air, and several main telescope components over the course of four observation nights on NIRC2.

on the telescope spiders themselves, which are typically thought to be the biggest factors in causing the LWE. There are other temperature sensors near the spiders, which could be a good approximation depending on their location. Measuring the temperature on or near the telescope spiders could help rule out this effect as a potential cause of performance issues. More active phase control of the primary mirror is currently being investigated, and could also help correct for the LWE if it is found to be a problem.

4.2.4 AO enclosure hatch

Large temperature differences between the dome and the AO enclosure could cause turbulent flows through the AO enclosure hatch, which connects the two. The temperatures recorded on the AO bench (near the WFS and NIRC2 instruments) are consistently higher than the dome and outdoor temperatures on the nights tested, lending credence to this hypothesis. The dome-NIRC2 temperature difference also ranks fairly high in Table 7, indicating that it is moderately correlated with image quality.

4.2.5 Summary

The correlation between AO image quality and temperature appears to be a complex one involving both the dome air and the differential temperatures between various telescope components. The overall trend is such that the image quality is higher when the dome air is warmer than both the primary mirror and the AO bench, and when the dome air is warmer in general. This indicates that mirror seeing, turbulence from the AO hatch, dome seeing, or some combination of all three could be causing a degradation in image quality. Further research and data collection are needed, both to confirm or refute these hypotheses and, eventually, to correct for this effect.

Table 8 Mean absolute errors (top of each row) and mean absolute percentage errors (bottom of each row) for each ML algorithm tested.

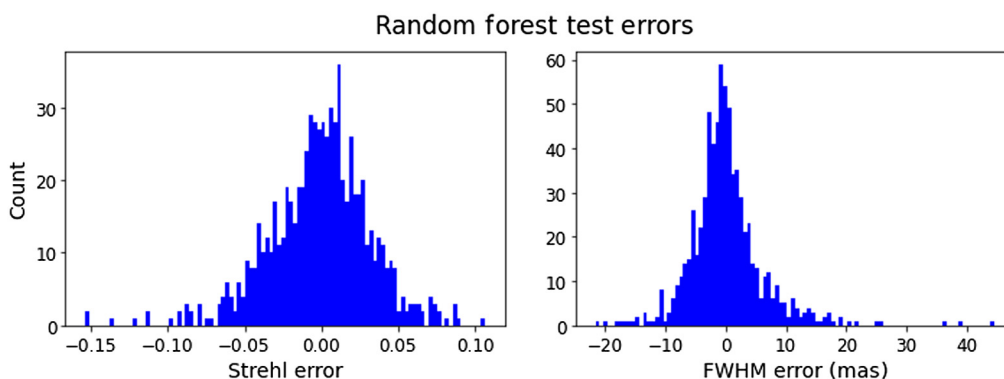
Algorithm	Best training error		Best testing error	
	Strehl	FWHM	Strehl	FWHM
Random Forest	0.01	1.96 mas	0.03	5.42 mas
	4.55%	2.49%	17.48%	6.69%
SVR	0.04	6.8 mas	0.04	7.33 mas
	19.27%	8.48%	25.29%	8.71%
ANN	0.04	6.38 mas	0.04	7.15 mas
	17.7%	8.22%	27.83%	8.89%

4.3 Predictive Algorithm

After performing a grid search over each algorithm's hyperparameters as described in Table 4, we list the results of the best-performing functions in each category in Table 8. We find that the algorithm with the best overall performance is the random forest regressor. This is logical for our relatively naive approach, as it is the most adaptable of the algorithms tested given the relatively simple feature set and model-selection method. The error distributions for both the Strehl ratio and FWHM are shown in Figs. 12 and 13 contains a plot of the Strehl and FWHM predicted values and percent errors against the true values. We see a sharp increase in Strehl percentage error as the Strehl ratio itself gets lower, as shown in Fig. 13. This could be due to the fact that Strehl calculations themselves become less accurate at lower Strehl,⁴⁶ which could interfere with the accuracy of our predictive algorithm in that range. These preliminary results are encouraging, however, in that predictions of image quality are more than 90% accurate on the tested FWHM data (80% for Strehl) with these nine parameters alone.

5 Discussion and Future Work

From our analysis, we find several areas for further investigation and potential improvement of the Keck-II LGS-AO system. The mean/median image quality of NIRC2 images has degraded slightly over time (Figs. 2 and 9 and Table 6), despite a series of improvements to the Keck-II telescope and AO system.^{29–33} It is unclear whether these problems stem from the NIRC2

**Fig. 12** Histogram of random forest Strehl and FWHM error values. Errors are distributed roughly evenly about zero, with heavier tails in one direction.

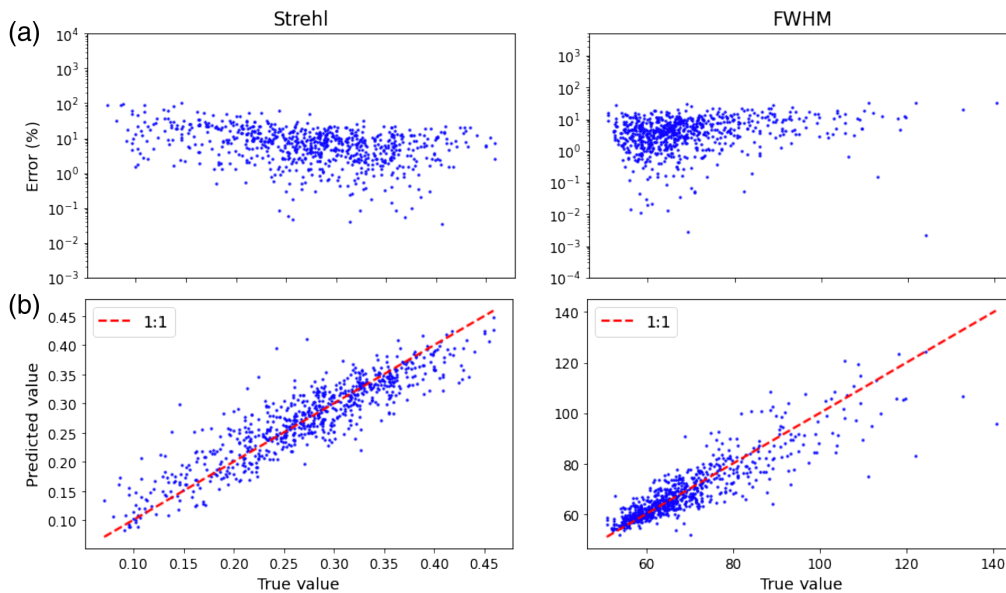


Fig. 13 (a) Percentage errors and (b) random forest Strehl and FWHM predicted values plotted against the true values. The predicted Strehl ratio and FWHM are less accurate at smaller values when accuracy is considered as a percentage of the true value.

instrument itself, the AO system overall, the Keck-II telescope, or other environmental or systematic factors, but a similar analysis using additional Keck-II instruments or data from other surveys could help to clarify this point. A study of the Keck AO logs and similar temporal trends in seeing point to the CLS or a broader environmental change in seeing as possible avenues of further research. As the decrease in image quality seems to happen more strongly in later years, studying the relationship between image quality and time while controlling for the effects of seeing and/or temperature may reveal more about the causes of this trend.

In addition, we find that the image quality in this survey tends to degrade when the ambient temperature is lower. Similar studies of the GPI (see Refs. 11 and 12), have found that temperature differences in the beam path can have a major impact on system performance, which could also be the case at Keck. At Gemini, mirror seeing was found to be causing atmospheric turbulence within the telescope dome. This may be a contributing factor to the performance degradation seen in this data, however, in this case, the image quality does not appear to be highly correlated with the outdoor-primary temperature difference, as it is in the GPI exoplanet survey studies data, although it does show a correlation with the dome-primary temperature difference. This indicates that another factor discussed in Sec. 4.2, such as AO hatch turbulence, dome seeing, or the LW, could be involved, and further investigation is needed to determine the exact cause. For example, the Keck-II telescope has many more temperature sensors available in and around the tube and dome than we have been able to obtain for this study. More specific and localized data could greatly benefit this research and could be requested for future iterations of this project.

In creating an image quality prediction algorithm, we were able to train a simple random forest model to predict the Strehl ratio to within 18% (0.03) and the FWHM to within 7% (5.42 mas), which is a good starting point for future iterations of this research. Our current algorithm performs worse at lower Strehl ratios and higher FWHM values, however, which is something that should be addressed before creating a formal observing tool. Other necessary improvements include gathering more data as inputs such as turbulence parameters (r_0 , τ_0 , θ_0), guide star brightness, etc., training the algorithm on a larger data set, or including observations of other targets. In Ref. 14, MASS and DIMM seeing are shown to have significant predictive power on the image quality, as is the case in this study, but the variable with the largest predictive power is found to be the Fried parameter, r_0 , which can be estimated from the seeing. Future work could significantly benefit from the inclusion of r_0 in the training of the predictive algorithm,

as it may improve the resulting accuracy. A higher level of accuracy for this algorithm, and the eventual development of an observing tool, would mean higher-quality imaging data from NIRC2 in the future, as this would enable better target selection and exposure time adjustments during each observing run.

Over the course of this project, we have identified several avenues for future research into the Keck-II LGS AO system, including potential long-term issues affecting AO performance, and have trained a preliminary algorithm for image quality prediction during observing runs. The full dataset is available from UC Berkeley⁴⁷ and the code library used to create it is located at this GitHub link: <https://github.com/jlyke-keck/KeckAoTelemetry>. It is our hope that the dataset and coding tools can be used by others to further improve Keck-II observations, so that the facility may be better used in pursuit of new research and discoveries going forward.

Acknowledgments

This paper has been previously published in SPIE conference proceedings.⁴⁸ The material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2018268910 and NSF Award AST-1518273. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We also acknowledge support from the W. M. Keck Foundation and the Heising Science Foundation.

References

1. A. H. Bouchez et al., “The Giant Magellan Telescope adaptive optics program,” *Proc. SPIE* **9148**, 300–318 (2014).
2. M. Le Louarn et al., “New challenges for adaptive optics: extremely large telescopes,” *Mon. Not. R. Astron. Soc.* **317**(3), 535–544 (2000).
3. B. Macintosh et al., “Extreme adaptive optics for the thirty meter telescope,” *Proc. SPIE* **6272**, 201–215 (2006).
4. J. R. Lu et al., “A search for Stellar-mass black holes via astrometric microlensing,” *Astrophys. J.* **830**(1), 41 (2016).
5. E. L. Nielsen et al., “The Gemini planet imager exoplanet survey: giant planet and brown dwarf demographics from 10 to 100 au,” *Astron. J.* **158**, 13 (2019).
6. M. C. Liu, T. J. Dupuy, and M. J. Ireland, “Keck laser guide star adaptive optics monitoring of 2MASS J1534-2952AB: first dynamical mass determination of a binary T Dwarf,” *Astrophys. J.* **689**(1), 436–460 (2008).
7. T. Do et al., “Relativistic redshift of the star S0-2 orbiting the Galactic center supermassive black hole,” *Science* **365**(6454), 664–668 (2019).
8. M. A. van Dam et al., “The W. M. Keck observatory laser guide star adaptive optics system: performance characterization,” *Publ. Astron. Soc. Pac.* **118**(840), 310–318 (2006).
9. P. L. Wizinowich et al., “The W. M. Keck observatory laser guide star adaptive optics system: overview,” *Publ. Astron. Soc. Pac.* **118**(840), 297 (2006).
10. S. Businger, T. Cherubini, and R. Lyman, “Mauna Kea Weather Center–home,” <http://mkwc.ifa.hawaii.edu/> (2021).
11. M. Tallis 2021 et al., “Air, telescope, and instrument temperature effects on the Gemini planet imager’s image quality,” arXiv:1807.07157 [astro-ph] (2018).
12. M. Tallis et al., “Effects of mirror seeing on high-contrast adaptive optics instruments,” *J. Astron. Telesc. Instrum. Syst.* **6**(01), 015002 (2020).
13. D. Savransky et al., “Mining the GPIES database,” *Proc. SPIE* **10703**, 107030H (2018).
14. V. P. Bailey et al., “Status and performance of the Gemini Planet Imager adaptive optics system,” *Proc. SPIE* **9909**, 99090V (2016).
15. D. J. Laidlaw et al., “Optimizing the accuracy and efficiency of optical turbulence profiling using adaptive optics telemetry for extremely large telescopes,” *Mon. Not. R. Astron. Soc.* **483**(4), 4341–4353 (2019).

16. M. van Kooten, N. Doelman, and M. Kenworthy, “Impact of time-variant turbulence behavior on prediction for adaptive optics systems,” *J. Opt. Soc. Am. A* **36**(5), 731 (2019).
17. S. Jia et al., “The Galactic Center: improved relative astrometry for velocities, accelerations, and orbits near the supermassive black hole,” *Astrophys. J.* **873**(1), 9 (2019).
18. S. Yelda et al., “Improving Galactic Center astrometry by reducing the effects of geometric distortion,” *Astrophys. J.* **725**(1), 331–352 (2010).
19. M. Service et al., “A new distortion solution for NIRC2 on the Keck II Telescope,” *Publ. Astron. Soc. Pac.* **128**, 095004 (2016).
20. W.M. Keck Observatory, “Keck LGSAO software,” 2016, <https://www2.keck.hawaii.edu/optics/lgsao/software/>.
21. A. Boehle et al., “An improved distance and mass estimate for SGR A* from a multistar orbit analysis,” *Astrophys. J.* **830**(1), 17 (2016).
22. A. M. Ghez et al., “The first laser guide star adaptive optics observations of the galactic center: Sgr A*’s infrared color and the extended red emission in its vicinity,” *Astrophys. J.* **635**(2), 1087 (2005).
23. A. M. Ghez et al., “Measuring distance and properties of the Milky Way’s Central supermassive black hole with stellar orbits,” *Astrophys. J.* **689**(2), 1044–1062 (2008).
24. J. R. Lu et al., “A disk of young stars at the galactic center as determined by individual stellar orbits,” *Astrophys. J.* **690**(2), 1463–1487 (2008).
25. S. Yelda et al., “Adaptive optics observations of the Galactic Center young stars,” *Proc. SPIE* **8447**, 84470A (2012).
26. S. Yelda et al., “Properties of the remnant clockwise disk of young stars in the Galactic Center,” *Astrophys. J.* **783**(2), 131 (2014).
27. A. Tokovinin and V. Kornilov, “Accurate seeing measurements with MASS and DIMM,” *Mon. Not. R. Astron. Soc.* **381**, 1179–1189 (2007).
28. The Pandas Development Team, “pandas-dev/pandas: Pandas” (2020).
29. S. M. Adkins et al., “New developments in instrumentation at the W. M. Keck Observatory,” *Proc. SPIE* **9147**, 914703 (2014).
30. J. C. Y. Chin et al., “Laser guide star facility developments at W. M. Keck Observatory,” *Proc. SPIE* **9148**, 914808 (2014).
31. J. C. Y. Chin et al., “Keck II laser guide star AO system and performance with the TOPTICA/MPBC laser,” *Proc. SPIE* **9909**, 99090S (2016).
32. E. M. Johansson et al., “Upgrading the Keck AO wavefront controllers,” *Proc. SPIE* **7015**, 70153E (2008).
33. S. Ragland et al., “Recent improvements to the Keck II laser guide star facility,” in *Adapt. Opt. for Extremely Large Telesc. 4—Conf. Proc.*, Vol. 1(1) (2015).
34. K. Pearson, “On lines and planes of closest fit to systems of points in space,” *Lond. Edinburgh, Dublin Philos. Mag. J. Sci.* **2**(11), 559–572 (1901).
35. T. K. Ho, “Random decision forests,” in *Proc. 3rd Int. Conf. Doc. Anal. Recognit.*, IEEE, Vol. 1, pp. 278–282 (1995).
36. C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.* **20**(3), 273–297 (1995).
37. W. S McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bull. Math. Biophys.* **5**(4), 115–133 (1943).
38. S. M. LaValle et al., “On the relationship between classical grid search and probabilistic roadmaps,” *Int. J. Rob. Res.* **23**(7–8), 673–692 (2004).
39. F. Pedregosa et al., “Scikit-learn: machine learning in Python,” *J. Mach. Learn. Res.* **12**(85), 2825–2830 (2011).
40. C. R. Harris et al., “Array programming with NumPy,” *Nature* **585**(7825), 357–362 (2020).
41. S. Ragland, “K2 CLS on-sky performance (KAON 1082),” unpublished (2015).
42. P. Wizinowich, “Keck adaptive optics note 1157: Keck II LGS AO performance based on nightly metrics,” unpublished (2016).
43. M. van Dam et al., “Keck adaptive optics note 489: performance of the Keck II AO system,” unpublished (2007).
44. R. Biasi et al., “Implementation and initial test results of the new Keck real time controller,” *Proc. SPIE* **11448**, 114482O (2020).

45. J. Milli et al., “Low wind effect on VLT/SPHERE: impact, mitigation strategy, and results,” *Proc. SPIE* **10703**, 752–771 (2018).
46. “Keck Strehl meter,” 2016, https://www2.keck.hawaii.edu/optics/aochar/Strehl_meter2.htm.
47. UC Berkeley Astronomy, <http://bhs.astro.berkeley.edu/keckao/>.
48. E. Ramey et al., “Analyzing long-term performance of the Keck-II adaptive optics system,” *Proc. SPIE* **11448**, 1010–1026 (2020).

Biographies of the authors are not available.