

Article

Multi-Wavelength Machine Learning for High-Precision Colorimetric Sensing

Majid Aalizadeh ^{1,2}, Chinmay Raut ³, Ali Tabartehfarahani ^{1,2,4,5} and Xudong Fan ^{1,2,4,5,*}

¹ Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA; maalizad@umich.edu (M.A.); afarahan@umich.edu (A.T.)

² Biointerfaces Institute, University of Michigan, Ann Arbor, MI 48109, USA

³ Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA; craut@umich.edu

⁴ Center for Wireless Integrated MicroSensing and Systems (WIMS2), University of Michigan, Ann Arbor, MI 48109, USA

⁵ Max Harry Weil Institute for Critical Care Research and Innovation, University of Michigan, Ann Arbor, MI 48109, USA

* Correspondence: xsfan@umich.edu

Highlights

What are the main findings?

- Multi-wavelength machine learning significantly improved concentration prediction over single-wavelength colorimetric analysis.
- Feature-selected spectral modeling reduced prediction error using only twelve optimized wavelengths.

What are the implications of the main findings?

- Informative spectral features can be extracted beyond conventional single-wavelength intensity measurements.
- This approach may improve low-cost colorimetric sensing without hardware modification.

Abstract

Conventional colorimetric sensing methods typically rely on signal intensity at a single wavelength, often selected heuristically based on peak visual modulation. This approach overlooks the structured information embedded in full-spectrum transmission profiles, particularly in intensity-based systems where linear models may be highly effective. In this study, we experimentally demonstrate that applying a forward feature selection strategy to normalized transmission spectra, combined with linear regression and ten-fold cross-validation, yields significant improvements in predictive accuracy. Using food dye dilutions as a model system, the mean squared error was reduced from over 22,000 with a single wavelength to 3.87 using twelve selected features, corresponding to a more than 5700-fold enhancement. These results validate that full-spectrum modeling enables precise concentration prediction without requiring changes to the sensing hardware. The approach provides a proof-of-concept framework that may be extended to colorimetric assays used in medical diagnostics, environmental monitoring, and industrial analysis following broader validation with real analytes and heterogeneous sample matrices.

Keywords: machine learning; colorimetric sensing; linear regression; mean squared error



Academic Editors: Peiyang Zhang and Athanasios V. Vasilakos

Received: 25 March 2026

Revised: 19 May 2026

Accepted: 19 May 2026

Published: 24 May 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Colorimetric sensing is one of the most widely used methods for chemical and biological detection because it translates molecular interactions into visible changes in light absorption or transmission [1–8]. Its simplicity, low cost, and compatibility with both laboratory and point-of-care [9] settings have made it essential in applications ranging from clinical diagnostics [10,11] and environmental monitoring [12,13] to food safety [14–17] and industrial process control [5,18]. Traditional approaches, however, typically rely on modeling intensity at a single wavelength from a full spectrum of wavelengths [19]. These wavelengths are often selected heuristically based on visual inspection of the spectrum, such as choosing the point of maximum absorbance. In practice, such single-wavelength readouts are often analyzed with one-dimensional fitting models such as linear regression or nonlinear curve fits like the four-parameter logistic (4PL) equation [20–24]. While convenient, this one-dimensional strategy discards most of the structured information contained in full spectra and therefore limits sensitivity, precision, and robustness. As a result, despite decades of development, many colorimetric platforms remain constrained by the same data-reduction step that oversimplifies what are inherently high-dimensional signals. Prior machine learning-assisted colorimetric sensing studies have predominantly relied on image-based analysis, RGB extraction, or sensor-array pattern recognition approaches, whereas comparatively limited attention has been given to quantitative multi-wavelength spectral modeling using spectrometer-acquired transmission data [25–30].

Our earlier studies addressed this limitation by showing that spectra should be treated as structured, high-dimensional datasets, regardless of the specific sensing platform [31,32]. In the first study, we focused on resonance-based biosensing, where conventional approaches track a single resonance peak shift to estimate analyte concentration or refractive index changes [33–35]. Some resonance-based platforms also support both TE and TM polarizations, yielding distinct peaks whose joint tracking enhances sensitivity and reduces noise through internal referencing [36–38]. We demonstrated that this single-peak method overlooks the fact that multi-resonance structures contain several distinct resonances, each carrying partially independent information. By applying ridge regression to a set of resonances traced simultaneously in the absorption spectrum of a periodic silicon triangular nanorod meta-array, we showed that combining multiple resonance shifts in a ridge regression model provides far higher accuracy than tracing any single resonance alone, up to three orders of magnitude. In other words, we showed that using multiple predictors instead of just one, results in much higher accuracy in the prediction of the target variable. It was also systematically shown that by gradually adding to the number of predictors, the accuracy is consistently enhanced [31]. Although this meta-array was used as an example system, the principle is general: any resonant spectrum with multiple features can benefit from multi-feature modeling rather than one-dimensional fittings.

In the second study, we extended the approach to full-spectrum modeling [32], comparing resonance-based spectra with sharp nonlinear peaks versus intensity-modulated spectra with smooth variations. Using the same nanorod geometry as the previous work, Si produced multi-resonance spectra while Ti produced smooth intensity-based spectra, enabling direct comparison with single-wavelength methods. We found that intensity-modulated spectra exhibit a near-linear relationship with concentration at a given wavelength, allowing linear regression to reduce the mean squared error (MSE) by over 8000-fold, whereas resonance-based spectra showed limited improvement due to nonlinear peak shifts. These results show that modeling effectiveness depends on spectral characteristics and that significant precision gains can be achieved through multi-wavelength modeling without hardware changes.

Building on this foundation, the present study provides an experimental validation in colorimetry, a domain defined by intensity-based spectra that is highly compatible with linear modeling. Using food dye dilutions as a model system for measurements, we apply forward feature selection combined with linear regression and cross-validation. We show that just twelve wavelengths are sufficient to preserve the essential concentration-dependent structure of the full spectrum, reducing the MSE by more than 5700-fold compared to single-wavelength analysis (when using the best selected single wavelength). This result demonstrates that interpretable machine learning can substantially improve the accuracy of colorimetric assays without changes to hardware, and more broadly, that the same principles may extend beyond optics to other sensing modalities where structured spectral information is available, although this remains to be experimentally validated.

2. Materials and Methods

2.1. Experimental Setup

Figure 1 presents a detailed overview of the experimental setup used for capturing full-spectrum transmission data from liquid-phase colorimetric samples. As shown in Figure 1a, the schematic illustrates the optical path beginning with a broadband light source, which emits a continuous spectrum across the visible range. The light first passes through an optical aperture, which spatially filters the beam to define its profile and suppress stray light. It then travels through a series of lenses that shape and collimate the beam, ensuring uniform propagation through the optical axis and consistent illumination of the sample region.

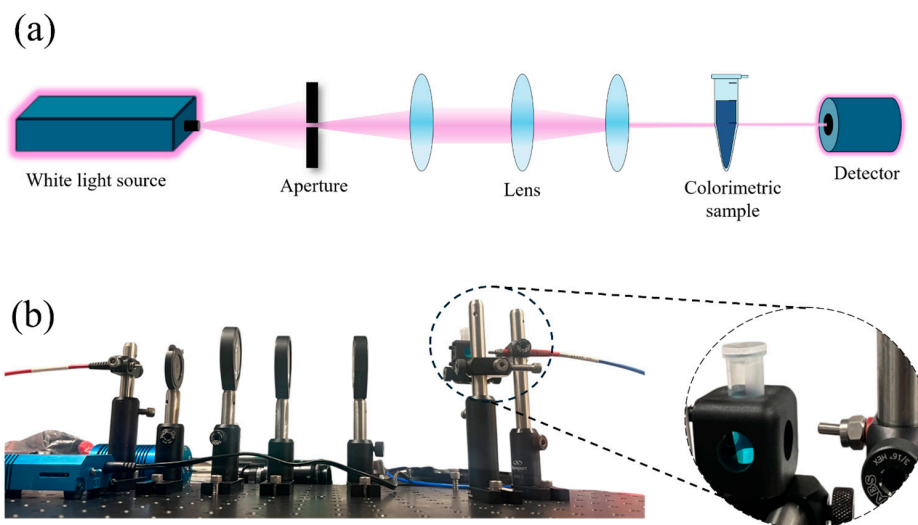


Figure 1. Experimental setup for transmission-based colorimetric measurements. (a) Schematic diagram of the optical path, showing the broadband light source, aperture, beam-shaping lenses, sample vial, and spectrometer aligned along the optical axis. (b) Photograph of the complete optical setup on a breadboard, including a zoomed-in view of the sample holder securing the vial at the beam center for stable and repeatable transmission measurements.

The colorimetric sample is contained within a standard microcentrifuge tube filled with dyed liquid and mounted vertically in the beam path. This configuration enables direct transmission measurements through the liquid column. As the beam traverses the sample, different spectral components are absorbed to varying degrees depending on the dye concentration, resulting in a wavelength-dependent attenuation. The remaining transmitted light is collected by a fiber-coupled spectrometer positioned on-axis downstream of the sample. The spectrometer records the full transmission spectrum from each sample, which serves as the raw input for all subsequent analysis and machine learning modeling.

Figure 1b shows a photograph of the physical system assembled on an optical breadboard. The components are mounted using standard post holders and translation stages, providing mechanical rigidity and fine control over alignment. The modular layout facilitates clear optical access to each section of the setup, particularly the sample region, while maintaining stable geometry across repeated measurements. Adjustable lens mounts allow the beam to be precisely collimated and focused for optimal transmission through the vial. An Ocean Optics LS-1 tungsten halogen lamp (Ocean Optics Inc., Dunedin, FL, USA) is used as a broadband light source. An Ocean Optics USB4000 spectrometer (Ocean Optics Inc., Dunedin, FL, USA) is used for recording the spectra using the OceanView 2.0 software.

A magnified view of the sample holder is shown in Figure 1b, highlighting the central alignment of the microcentrifuge tube within a 3D holder. This holder secures the vial in a fixed position along the beam axis, ensuring that the incident light is transmitted directly through the colored liquid without significant scattering or misalignment. The surrounding hardware is designed to enable rapid sample exchange while preserving alignment consistency across the dataset. Each measurement is repeated three times by removing and placing the sample vial back in the holder, and the result is averaged out in an attempt to remove the manual measurement errors. Details of the measurement setup are included in Supplementary Table S1.

2.2. Measurements

Figure 2 presents the complete colorimetric sample set and the corresponding raw transmission spectra acquired across a wide range of dye concentrations. Figure 2a displays the full set of food dye solutions, prepared by serial dilution from a 1000-unit stock solution. Each concentration was generated by precise volumetric mixing with deionized water and scaled to a final volume of exactly 1 milliliter to maintain uniform optical path length across all vials. This consistency ensured that transmission differences could be attributed solely to dye concentration, not variations in sample geometry.

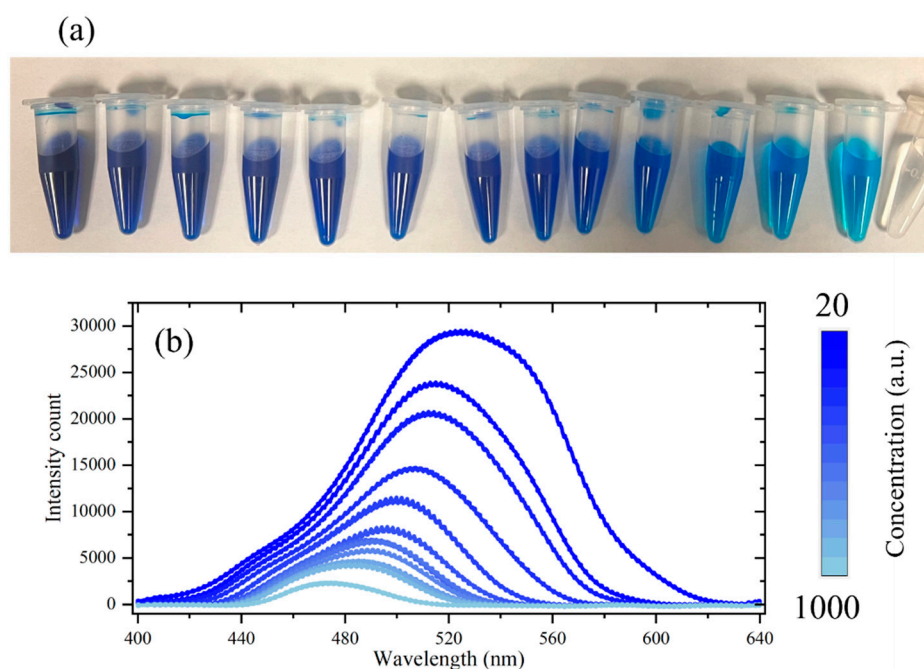


Figure 2. (a) Food dye samples ranging from 20 to 1000 concentration units (right to left), prepared by serial dilution from a 1000-unit stock and scaled to 1 mL to ensure consistent optical path length. (b) Raw transmission spectra collected for each concentration, showing decreasing intensity with increasing dye content. Spectra were averaged over three repeated placements per sample and lightly smoothed to reduce high-frequency noise while preserving spectral shape.

All spectral measurements were conducted in a dark environment to eliminate ambient light contamination. As mentioned above, to improve robustness while introducing realistic experimental variability, each sample vial was manually repositioned three times within the sample holder. The transmission spectrum was collected after each placement, and the resulting spectra were averaged to obtain a representative curve for each concentration level. A light smoothing filter was applied post-acquisition to reduce high-frequency spectrometer noise without distorting the spectral envelope or relative intensity profile.

All measurements were performed using the same brand and type of microcentrifuge tube to ensure consistent optical properties across the dataset. Although the tube material and transparency influence the absolute transmission intensity, these effects remain effectively constant across samples and therefore cancel out during relative spectral comparisons between concentrations.

Figure 2b shows the resulting transmission spectra for concentrations ranging from 20 to 1000 units. As expected, lower concentrations result in higher overall transmission across the visible spectrum due to reduced dye absorption. The spectral shape remains consistent, with a dominant transmission peak near 480–530 nm range and systematic amplitude changes as dye concentration increases. Notably, the visual modulation appears most pronounced in the 500–520 nm range, a region commonly targeted in traditional colorimetric assays. While conventional methods often rely on a single visually selected wavelength, later sections will demonstrate how performance varies dramatically across the spectrum and how greedy feature selection can identify combinations of wavelengths that dramatically improve prediction accuracy of the unknown concentration.

3. Results

3.1. Single-Wavelength Modeling

While full-spectrum data offers a wealth of information for concentration prediction, most traditional colorimetric systems still operate using just a single measurement wavelength. This is often done out of simplicity, legacy practice, or the assumption that the most visibly modulated part of the spectrum must also be the most predictive. To assess the validity and limitations of this assumption, we carried out a detailed investigation of one-dimensional linear models across a range of individual wavelengths. These models attempt to map transmission intensity at a single fixed wavelength to analyte concentration using one-dimensional linear fitting.

Figure 3 shows a comprehensive panel of such wavelength vs concentration linear fits, spanning from 425 nm to 625 nm. Each subplot displays a different wavelength, where transmission values measured across all concentrations, referring back to Figure 2b, were used to train a 1-dimensional linear regression model. At first glance, some wavelengths do indeed produce reasonable linear trends. The fits near 450 to 475 nm show relatively tight clustering of points around the fitted line and low RMSE values. However, other wavelengths, even just 25 or 50 nm away, quickly degrade in performance. The wavelength at 550 nm, for instance, yields a significantly higher RMSE than 475 nm, despite still falling within the region of strong visible modulation in the raw spectra. The situation worsens at the spectral tails. Fits at 425 nm and 625 nm show significant nonlinearity in the data, resulting in a weaker linear fit and large prediction errors. Furthermore, the occurrence of negative transmission values at certain wavelengths, which are physically non-meaningful, suggests the presence of measurement noise, baseline drift, or background subtraction errors, further highlighting the reduced reliability and robustness of single-wavelength approaches in these spectral regions.

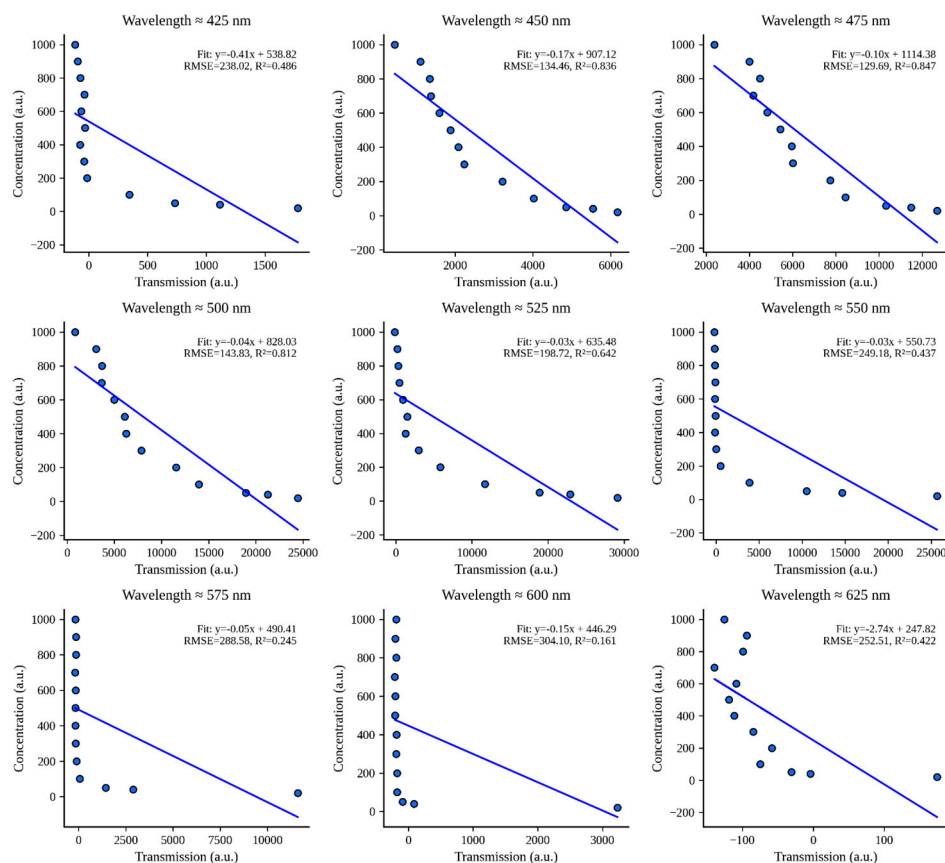


Figure 3. Single-wavelength simple linear regression fits at nine selected wavelengths from 425 to 625 nm, showing varying prediction quality and highlighting the limitations of manual wavelength selection.

The key takeaway from Figure 3 is that even though the underlying spectral shape varies smoothly, as shown in Figure 2b, the linear performance across wavelengths does not. The relationship between transmission and concentration is highly sensitive to wavelength choice, and small shifts in the selected value can make the difference between a usable model and a completely unreliable one. This highlights a major limitation of one-dimensional modeling. It depends entirely on a well-chosen wavelength. Unless this wavelength is picked through an objective search or cross-validation, there is no guarantee it will yield meaningful predictions.

To probe this issue more closely, Figure 4 isolates the best and worst performing wavelengths identified from the wavelength sweep. The left panel, Figure 4a, shows the simple linear fit at 457 nm, the wavelength with the lowest RMSE. Surprisingly, 457 nm is not visually the most dynamic point on the spectrum. One might have expected that a peak or steep slope region, such as 520 nm or 480 nm, would dominate. Yet statistical evaluation reveals that 457 nm offers the strongest predictive power, even though it may appear relatively unremarkable in the raw transmission profile.

In contrast, Figure 4b shows the linear fit at 610 nm, one of the worst performing wavelengths. The data is widely scattered, includes noise and unrealistic negative values, and the R^2 drops to just 0.17.

It is important to note that all results in Figures 3 and 4 were obtained without cross-validation. That is, the same dataset was used for both training and testing, which results in an idealized best-case scenario fitting performance. This form of evaluation is acceptable for probing the basic relationships between variables but is insufficient to estimate real-world measurement modeling performance. In practice, unknown samples with previously

unseen concentrations must be predicted without prior exposure. This scenario must be simulated using proper statistical methods such as k-fold cross-validation, where the model is repeatedly trained on subsets of the data and evaluated on held-out samples.

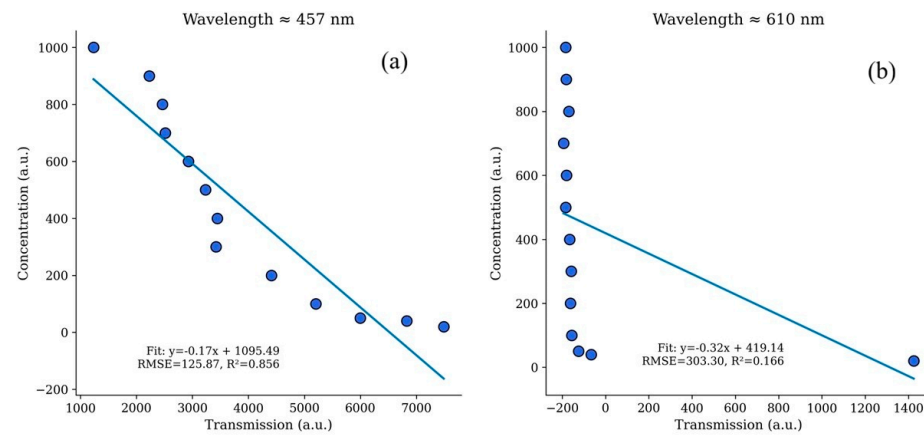


Figure 4. (a) Best-performing single-wavelength linear fit at 457 nm with RMSE of 125.87 and R^2 of 0.8563. (b) Worst-performing fit at 610 nm with RMSE of 303.30 and R^2 of 0.1658, highlighting the non-intuitive nature of optimal wavelength selection.

For our machine learning models, we adopt ten-fold cross-validation to simulate this realistic scenario. The dataset is split into ten equal parts. In each iteration, nine parts are used to train the model, and one part is reserved for testing. This process is repeated ten times so that every data point is tested exactly once. The resulting error metrics are averaged to produce a more generalizable estimate of model performance. Cross-validation not only prevents overfitting but also reveals how well a model can handle variability between different samples, including small deviations in optical alignment or measurement noise. In the context of medical diagnostics, this directly mimics predicting the analyte concentration of a new patient sample with unknown concentration using a trained regression model.

When cross-validation is applied to the single-wavelength models described in Figures 3 and 4, the performance degrades sharply. For the previously best-performing 457 nm wavelength, MSE jumps from around 15,000 to over 22,000. Even more dramatically, the R^2 value, which was originally positive and relatively high, becomes negative under cross-validation, indicating that the model performs worse than simply using the mean as a predictor. For this reason, the regression fit for the cross-validated model is not even plotted, as it offers no meaningful predictive value. This observation alone is enough to illustrate the fragility of one-dimensional fitting in real applications.

This degradation in cross-validated performance of single wavelength fitting stands in stark contrast to the behavior of our multi-feature machine learning models. When trained and tested on the same dataset without cross-validation, the multi-wavelength models achieve near-zero mean squared error and visually perfect fits using only a small number of intelligently selected wavelengths. Although modeling results without any cross-validation are not the focus of this work, they are worth mentioning to emphasize the predictive power that emerges when only multiple wavelengths are leveraged rather than one. In contrast to the one-dimensional models that rapidly lose accuracy under validation, the multi-wavelength machine learning models maintain strong generalization even when subjected to ten-fold cross-validation, which will be presented in detail later in this article. This comparison highlights the fundamental gap between single-wavelength overfitting and robust multi-wavelength modeling.

To further illustrate this point, Figure 5 presents the RMSE and R^2 values for single-wavelength linear regression models across the full spectral range from 400 nm to 640 nm.

Each point in these plots corresponds to a regression model trained at one specific wavelength using transmission intensity as the sole input variable.

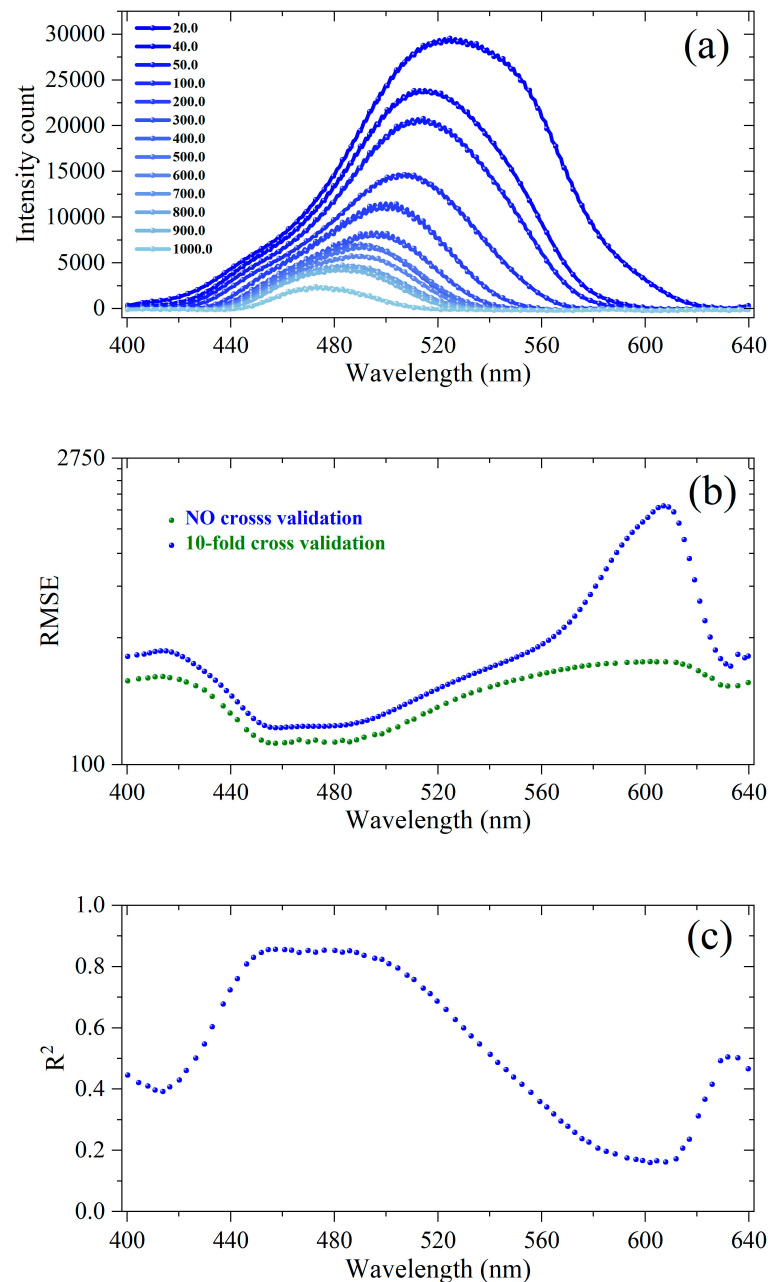


Figure 5. (a) Raw transmission spectra across 400–640 nm for all dye concentrations. (b) RMSE trends for single-wavelength linear regression models with and without ten-fold cross-validation. (c) Corresponding R^2 values without cross-validation. These plots illustrate the performance variability and generalization gap in single-feature models across the spectrum.

The top panel of Figure 5a shows the original intensity spectra, repeated here for reference to aid visual comparison with regression performance. These spectra show a clear decay in peak amplitude with increasing concentration, yet they do not inherently reveal which wavelengths yield the best predictive behavior. The middle panel, Figure 5b, presents the root mean squared error (RMSE) as a function of wavelength under two evaluation modes: without cross-validation (blue) and with ten-fold cross-validation (green). In the absence of cross-validation, RMSE drops sharply in the 450–480 nm region, forming a clear local minimum where prediction error is lowest. However, this region does not perfectly

align with the visually dominant features in the intensity spectra, again reinforcing that statistical relevance does not always coincide with visual salience.

Once ten-fold cross-validation is applied, the green curve in Figure 5b reveals the full impact of generalization constraints. The RMSE curve becomes noisier and shifts upward across nearly the entire spectral range. Even in regions that previously showed strong performance, such as around 457 nm, RMSE increases substantially under validation. This demonstrates that one-dimensional models are highly susceptible to overfitting when tested on the same data they were trained on, and their performance rapidly deteriorates when asked to generalize.

Finally, Figure 5c plots the coefficient of determination (R^2) across all wavelengths, again under the no cross-validation scenario. The trend mirrors that of the RMSE plot, peaking around 0.86 near 457 nm and steadily declining away from that region. Past 600 nm, the R^2 value drops below 0.2, reflecting minimal predictive value. No R^2 curve is shown for the cross-validation case because many of the corresponding values are negative, indicating worse-than-mean prediction and rendering the metric effectively meaningless.

Another key insight from Figure 5 is that although the raw spectral signal appears smooth and continuous, the actual predictive information is unevenly distributed. There are narrow regions where concentration can be estimated with moderate confidence using a single wavelength, but most of the spectrum offers little usable information in isolation. This strongly supports the case for multivariate modeling. Without a principled method for feature selection or validation, choosing a wavelength manually is nearly indistinguishable from guessing.

Together, Figures 3–5 paint a consistent picture: single-wavelength colorimetric analysis is highly limited, particularly when prediction must generalize beyond a training set. Even seemingly strong fits collapse under ten-fold cross-validation. This also implies that in most cases measurement inaccuracy is not a result of the sensor hardware limitations, further signifying the importance of using the maximum modeling capacity based on the measured spectra. As previously mentioned, our machine learning models trained without cross-validation achieve zero or near-zero MSE and perfect visual fits using just a few selected wavelengths. These results, although idealized, illustrate the upper bound of what is possible when the spectral data is fully utilized. The rest of this work focuses on the more realistic case of cross-validated multi-feature modeling, which will demonstrate how spectral information, when carefully selected and modeled, can yield robust and generalizable sensing performance and can result in thousands of fold improvement in MSE compared to single-wavelength modeling.

Additional analyses restricted to the optimal linear domains of different wavelength regions, along with leave-one-out cross-validation (LOO-CV) comparisons for the multi-wavelength models, are provided in Supplementary Tables S3 and S4.

3.2. Multi-Wavelength Modeling

Table 1 and Figure 6 together present a comprehensive summary of the greedy forward feature selection process applied to normalized transmission data. This combined analysis marks the transition from single-wavelength fitting to a more intelligent multi-feature modeling strategy, where machine learning is used not only for regression but for optimal information extraction from the spectrum. Greedy forward feature selection iteratively builds a feature subset by adding, at each step, the wavelength that provides the largest improvement in a chosen performance metric (e.g., reduction in MSE), without reconsidering previously selected features [32]. The feature selection was guided by the relative performance of the wavelength across the entire dataset. At each iterative step, candidate wavelengths were considered by calculating their average squared error (MSE) across

the 10 cross-validation folds. The wavelength that produced the lowest average cross-validation MSE was then subsequently added to the feature set. Additional comparisons with alternative regression models are provided in Supplementary Table S2.

Table 1. Stepwise feature selection results showing wavelength identifiers, MSE and RMSE values from 10-fold cross-validation, and fold improvements relative to the 1-feature baseline.

| Total Features | Feature Added in nm (Wavelength) | MSE | SEM | RMSE | MSE Improvement (Folds) | RMSE Improvement (Folds) |
|----------------|----------------------------------|-----------|---------|--------|-------------------------|--------------------------|
| 1 | 457.27 | 22,157.58 | 3112.47 | 148.85 | 1 | 1 |
| 2 | 427.16 | 6753.58 | 1868.05 | 82.18 | 3.28 | 1.81 |
| 3 | 631.85 | 6861.9 | 1998.1 | 82.84 | 3.22 | 1.79 |
| 4 | 423.56 | 4697.08 | 1250.52 | 68.54 | 4.71 | 2.17 |
| 5 | 478.19 | 3205.72 | 956.88 | 56.62 | 6.91 | 2.62 |
| 6 | 479.65 | 2388.56 | 701.93 | 48.87 | 9.27 | 3.04 |
| 7 | 545.11 | 1975.39 | 628.57 | 44.45 | 11.21 | 3.34 |
| 8 | 629.06 | 3017.36 | 1041.94 | 54.93 | 7.34 | 2.70 |
| 9 | 452.44 | 2759.62 | 780.88 | 52.53 | 8.02 | 2.83 |
| 10 | 547.77 | 8.81 | 3.91 | 2.97 | 2515.04 | 50.15 |
| 11 | 631.05 | 18.37 | 7.91 | 4.29 | 1206.18 | 34.73 |
| 12 | 488.60 | 3.87 | 2.87 | 1.97 | 5725.47 | 75.66 |
| 13 | 455.17 | 32.14 | 19.26 | 5.67 | 689.40 | 26.25 |
| 14 | 633.84 | 70.15 | 17.13 | 8.38 | 315.86 | 17.77 |
| 15 | 633.04 | 29.79 | 13.24 | 5.46 | 743.79 | 27.27 |
| 16 | 635.82 | 22.73 | 29.69 | 4.77 | 974.81 | 31.22 |
| 17 | 421.40 | 115.99 | 59.11 | 10.77 | 191.03 | 13.82 |
| 18 | 419.32 | 166.32 | 92.54 | 12.9 | 133.22 | 11.54 |
| 19 | 417.21 | 234.04 | 121.23 | 15.3 | 94.67 | 9.73 |
| 20 | 415.08 | 295.73 | 156.52 | 17.2 | 74.92 | 8.65 |

Table 1 lists the ordered sequence of wavelengths added during the feature selection process, along with the corresponding mean squared error (MSE), Standard Error of the MSE (SEM), root mean squared error (RMSE), and the relative improvement in both metrics compared to the single-wavelength baseline. The first row establishes the baseline case, where only the best-performing single wavelength (457.275 nm) is used, yielding an RMSE of 148.85. As additional wavelengths are incorporated, both MSE and RMSE drop sharply, especially within the first 3 to 5 additions. Notably, by the time twelve features have been selected, the MSE drops to 3.87 and RMSE to just 1.97. This represents a staggering 5725-fold improvement in MSE and over 75-fold improvement in RMSE just by using 12 wavelengths rather than just one wavelength (the best performing wavelength).

Figure 6 offers a visual narrative of this optimization process across six panels. In Figure 6a,b, the absolute MSE and RMSE values are plotted against the number of selected features. A steep decline is visible in both curves within the first 5 to 7 features, indicating that a large proportion of predictive power is concentrated in a small number of well-chosen wavelengths. MSE drops sharply at the 10-wavelength scenario, and it drops to its lowest at the 12-wavelength fitting.

Panels Figure 6c,d replot the MSE and RMSE curves on a logarithmic scale. These log-scale versions emphasize the dramatic nature of the improvement achieved around the tenth feature, where MSE plummets from thousands to single digits. This is especially critical when comparing to the baseline MSE of over 22,000 from single-wavelength fitting.

Finally, Figure 6e,f display the fold improvement in MSE and RMSE on a log scale, respectively. These panels isolate the gain brought by each additional feature compared to the single-wavelength fitting. The peak in Figure 6e corresponds exactly to the 12th feature, echoing the highlight from Table 1. The curve in Figure 6f tells the same story for RMSE,

albeit with a smaller dynamic range. Together, these plots demonstrate that feature number 12 was not merely another increment but a pivotal addition that elevated the model to an entirely different tier of performance.

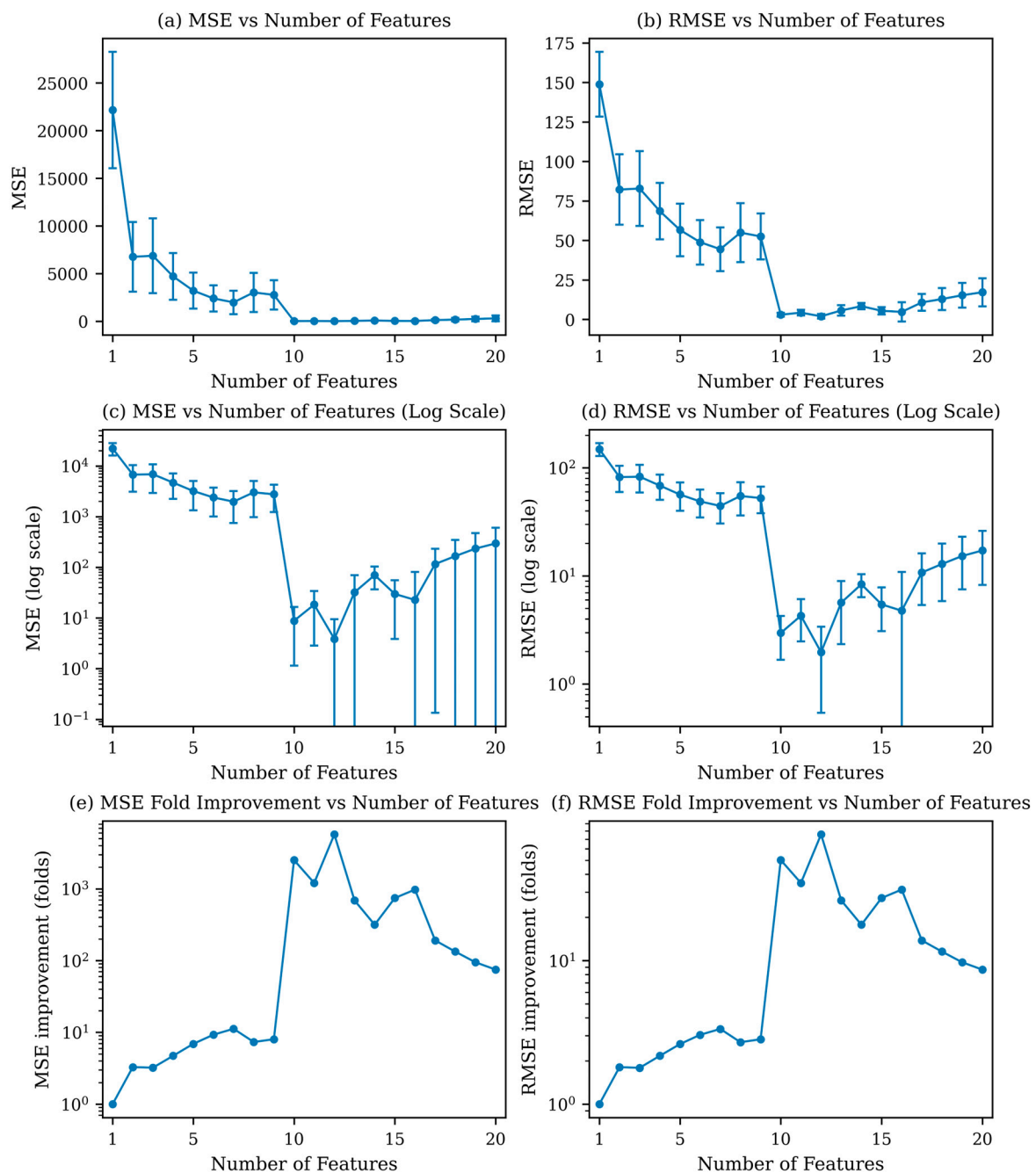


Figure 6. Prediction error and enhancement trends as a function of the number of selected features. (a,b) MSE and RMSE values decrease rapidly with added features using normalized transmission and 10-fold cross-validation. (c,d) Corresponding log-scale plots reveal nonlinear behavior in prediction error reduction, especially for feature sets between 10 and 15 variables. (e,f) Fold improvement in MSE and RMSE highlights optimal feature ranges where modeling performance gains are maximized. Vertical bars represent the 95% confidence intervals for the mean.

These findings reinforce that performance gains are not uniformly distributed across the spectral domain. Instead, a small subset of wavelengths carries an outsized share of predictive information, and these can be discovered only through data-driven model-

ing. Arbitrary or visually chosen wavelengths simply cannot match the predictive power achieved here.

Taken together, Table 1 and Figure 6 establish both the theoretical and practical value of intelligent feature selection in colorimetric sensing. They show that machine learning models do not need to rely on full-spectrum fitting or brute-force analysis. Instead, they can achieve near-perfect performance through judicious selection of just a dozen highly informative wavelengths. This makes the approach scalable and computationally efficient, opening the door to real-time or embedded implementations without sacrificing accuracy. As such, the results presented here highlight the potential of combining interpretable machine learning with multi-wavelength colorimetric measurements to improve concentration prediction accuracy.

Figure 7 provides a critical conceptual bridge between raw spectral behavior and data-driven feature selection, offering a more nuanced understanding of why the 12 selected wavelengths performed so well in the previous models. While the prior table and performance plots quantified the benefit of feature addition step by step, this figure turns our attention toward why those particular wavelengths are effective and adds an intuitive understanding.

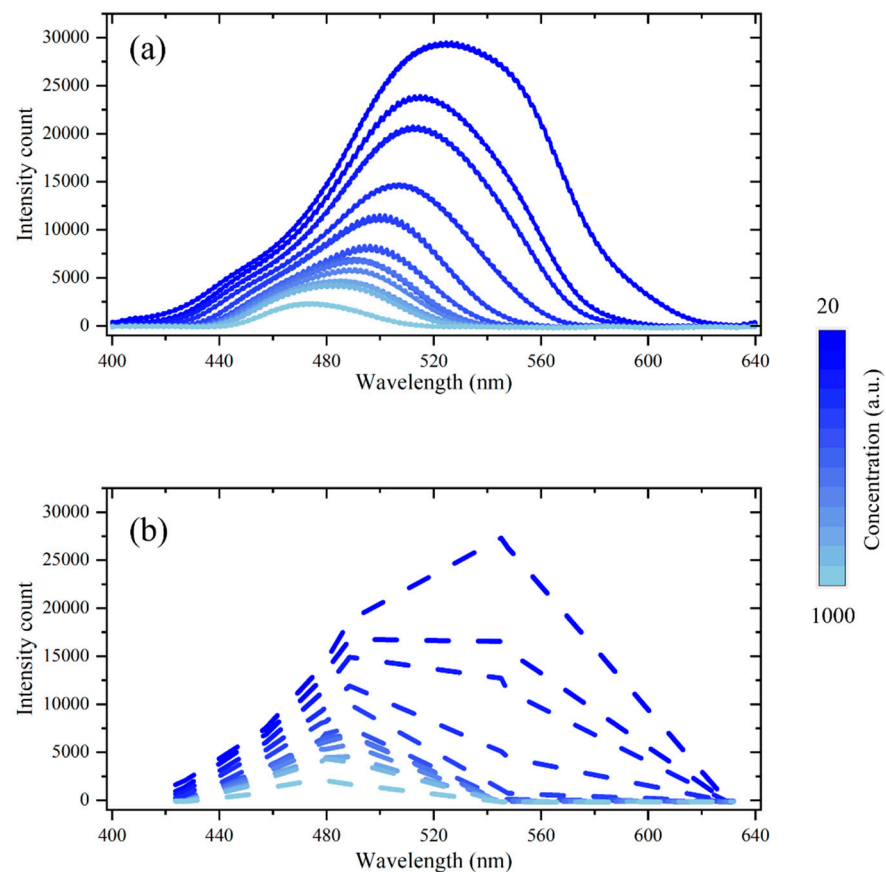


Figure 7. (a) Raw transmission spectra of all dye concentrations. (b) Reduced spectra showing the 12 selected features across concentrations.

Figure 7a shows the raw intensity spectra for all concentrations across the full wavelength range, just as seen earlier in Figure 2b. Figure 7b, however, transforms those same curves into a representation based on the 12 selected features, in effect, a visual approximation of what the model sees when relying solely on those chosen wavelengths. The dashed lines trace the transmission values at those 12 specific points across all dye concentrations, simulating a sparse sensing regime with high interpretability.

What immediately stands out in Figure 7b is that these 12 wavelengths preserve the overall shape and concentration-dependent structure of the full spectra remarkably well. Even though the resolution is drastically reduced, the essential contrast between high and low concentration samples is maintained with clarity. This preservation of signal shape and ranking, despite aggressive dimensionality reduction, is the conceptual key to why the selected wavelengths work so well. The model is not randomly guessing, it is identifying anchor points in the spectrum that jointly approximate the full signal's behavior. In short, Figure 7 helps explain the why behind the what. The selected wavelengths work well not only because they reduce MSE numerically, but because they are physically distributed across the spectrum in a way that captures concentration-induced modulation without falling into the trap of redundancy. Interestingly, many of the selected wavelengths were concentrated in the shorter wavelength region of the spectrum, where the transmission intensity changes more rapidly with concentration. These regions contain stronger concentration-dependent spectral variation compared to flatter portions of the spectrum, making them more informative for concentration prediction. At the same time, wavelengths from multiple spectral regions were selected, suggesting that the model benefits from preserving broader spectral shape information rather than relying on a single absorption feature.

In conventional colorimetric sensing, quantitative analysis is often restricted to the linear calibration region to minimize fitting errors in single-wavelength approaches. However, limiting analysis to only the linear regime reduces the usable dynamic range and may overlook informative spectral variations near the detection limit or partially nonlinear regions. Therefore, this study investigated a broader concentration range to evaluate whether multi-wavelength machine learning models could recover predictive information beyond the strictly linear domain. The results suggest that, despite increased nonlinearity, the proposed multi-wavelength approach maintained improved prediction performance by leveraging distributed spectral information rather than a single absorbance feature. Additional comparisons within narrower concentration ranges are provided in Supplementary Tables S3 and S4.

4. Conclusions

This work presents an experimental demonstration of multi-wavelength machine-learning-assisted colorimetric sensing. Through a deliberately simple yet rigorously validated optical setup, we demonstrate that full-spectrum transmission data, when properly modeled, can substantially improve prediction accuracy within controlled colorimetric measurement environments. Although one-dimensional regression initially appeared promising at specific wavelengths such as 457 nm, this performance proved fragile. Without cross-validation, models showed acceptable fits ($R^2 \approx 0.86$), but under ten-fold cross-validation, used here as a proxy for real-world variability, performance deteriorated sharply. In many cases, R^2 became negative and RMSE exceeded 30,000, revealing a fundamental limitation of the single-wavelength approach.

To address this, we adopted a multi-wavelength framework based on greedy forward feature selection and ridge regression. Using interpretable linear models on normalized transmission spectra, we found that just twelve carefully selected wavelengths reduced the mean squared error from over 22,000 to 3.87 under 10-fold cross-validation, corresponding to a 5700-fold improvement without any hardware modification. The selected wavelengths were distributed across the spectrum, capturing complementary, non-redundant information and forming a compact, physically meaningful basis for concentration estimation. These results challenge the assumption that visually dominant wavelengths are the most informative, instead showing that predictive power lies in distributed spectral features.

More broadly, this work reframes colorimetric sensing as a high-dimensional data problem and supports the hypothesis that linear regression is well suited for intensity-based sensing, where the relationship with analyte concentration remains quasi-linear for a given wavelength.

Finally, it is noteworthy that the present study was performed using controlled food dye solutions under relatively stable laboratory conditions and did not include complex biological or environmental matrices, interfering species, or temperature-dependent variations. Therefore, while the results strongly support the potential of multi-wavelength machine learning for improved concentration prediction, additional validation using real analytes and practical sensing conditions will be necessary before translation to deployed diagnostic or environmental sensing systems. Future work will investigate robustness against interferences, matrix effects, temperature fluctuations, and clinically relevant analytes.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/s26113327/s1>, Table S1. Detailed experimental parameters; Table S2. Comparison to other machine learning models; Table S3. Sensitivity Analysis of Single-Wavelength Modeling Across Different Concentration Ranges; Table S4. Sensitivity Analysis of Multi-Wavelength Modeling Across Different Concentration Ranges.

Author Contributions: Conceptualization, M.A., A.T. and X.F.; methodology, M.A. and C.R.; software, M.A. and C.R.; validation, M.A.; formal analysis, M.A.; investigation, M.A.; resources, X.F.; data curation, M.A. and C.R.; writing—original draft preparation, M.A.; writing—review and editing, M.A., C.R., A.T. and X.F.; visualization, M.A., C.R. and X.F.; supervision, X.F.; project administration, X.F.; funding acquisition, X.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation through grant number 2225568.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, X.; Yin, J.; Yoon, J. Recent advances in development of chiral fluorescent and colorimetric sensors. *Chem. Rev.* **2014**, *114*, 4918–4959. [[CrossRef](#)]
2. Piriya, V.S.A.; Joseph, P.; Daniel, S.C.G.K.; Lakshmanan, S.; Kinoshita, T.; Muthusamy, S. Colorimetric sensors for rapid detection of various analytes. *Mater. Sci. Eng. C Mater. Biol. Appl.* **2017**, *78*, 1231–1245. [[CrossRef](#)] [[PubMed](#)]
3. Zhang, C.; Suslick, K.S. A colorimetric sensor array for organics in water. *J. Am. Chem. Soc.* **2005**, *127*, 11548–11549. [[CrossRef](#)]
4. Jin, Z.; Yim, W.; Retout, M.; Housel, E.; Zhong, W.; Zhou, J.; Strano, M.S.; Jokerst, J.V. Colorimetric sensing for translational applications: From colorants to mechanisms. *Chem. Soc. Rev.* **2024**, *53*, 7681–7741. [[CrossRef](#)]
5. Feng, L.; Musto, C.J.; Kemling, J.W.; Lim, S.H.; Zhong, W.; Suslick, K.S. Colorimetric sensor array for determination and identification of toxic industrial chemicals. *Anal. Chem.* **2010**, *82*, 9433–9440. [[CrossRef](#)]
6. Wang, Y.W.; Wang, M.; Wang, L.; Xu, H.; Tang, S.; Yang, H.H.; Zhang, L.; Song, H. A Simple Assay for Ultrasensitive Colorimetric Detection of Ag⁺ at Picomolar Levels Using Platinum Nanoparticles. *Sensors* **2017**, *17*, 2521. [[CrossRef](#)] [[PubMed](#)]
7. Mohamad, A.; Teo, H.; Keasberry, N.A.; Ahmed, M.U. Recent developments in colorimetric immunoassays using nanozymes and plasmonic nanoparticles. *Crit. Rev. Biotechnol.* **2019**, *39*, 50–66. [[CrossRef](#)] [[PubMed](#)]
8. Carey, J.R.; Suslick, K.S.; Hulkower, K.I.; Imlay, J.A.; Imlay, K.R.; Ingison, C.K.; Ponder, J.B.; Sen, A.; Wittrig, A.E. Rapid identification of bacteria with a disposable colorimetric sensing array. *J. Am. Chem. Soc.* **2011**, *133*, 7571–7576. [[CrossRef](#)]
9. Thiha, A.; Ibrahim, F. A Colorimetric Enzyme-Linked Immunosorbent Assay (ELISA) Detection Platform for a Point-of-Care Dengue Detection System on a Lab-on-Compact-Disc. *Sensors* **2015**, *15*, 11431–11441. [[CrossRef](#)]
10. Zhao, V.X.T.; Wong, T.I.; Zheng, X.T.; Tan, Y.N.; Zhou, X. Colorimetric biosensors for point-of-care virus detections. *Mater. Sci. Energy Technol.* **2020**, *3*, 237–249. [[CrossRef](#)]

11. Liu, H.; Ding, Y.N.; Bian, B.; Li, L.; Li, R.; Zhang, X.; Liu, Z.; Zhang, X.; Fan, G.; Liu, Q. Rapid colorimetric determination of dopamine based on the inhibition of the peroxidase mimicking activity of platinum loaded CoSn(OH)₆ nanocubes. *Mikrochim. Acta* **2019**, *186*, 755. [[CrossRef](#)]
12. Alberti, G.; Zanoni, C.; Magnaghi, L.R.; Biesuz, R. Disposable and Low-Cost Colorimetric Sensors for Environmental Analysis. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8331. [[CrossRef](#)]
13. Song, Y.; Wei, W.; Qu, X. Colorimetric biosensing using smart materials. *Adv. Mater.* **2011**, *23*, 4215–4236. [[CrossRef](#)] [[PubMed](#)]
14. Magnaghi, L.R.; Capone, F.; Zanoni, C.; Alberti, G.; Quadrelli, P.; Biesuz, R. Colorimetric Sensor Array for Monitoring, Modelling and Comparing Spoilage Processes of Different Meat and Fish Foods. *Foods* **2020**, *9*, 684. [[CrossRef](#)]
15. Liu, G.; Lu, M.; Huang, X.; Li, T.; Xu, D. Application of Gold-Nanoparticle Colorimetric Sensing to Rapid Food Safety Screening. *Sensors* **2018**, *18*, 4166. [[CrossRef](#)] [[PubMed](#)]
16. Wu, F.; Wang, H.; Lv, J.; Shi, X.; Wu, L.; Niu, X. Colorimetric sensor array based on Au(2)Pt nanozymes for antioxidant nutrition quality evaluation in food. *Biosens. Bioelectron.* **2023**, *236*, 115417. [[CrossRef](#)] [[PubMed](#)]
17. Nie, W.; Chen, Y.; Zhang, H.; Liu, J.; Peng, Z.; Li, Y. A novel colorimetric sensor array for real-time and on-site monitoring of meat freshness. *Anal. Bioanal. Chem.* **2022**, *414*, 6017–6027. [[CrossRef](#)]
18. Li, Z.; Suslick, K.S. Colorimetric Sensor Array for Monitoring CO and Ethylene. *Anal. Chem.* **2019**, *91*, 797–802. [[CrossRef](#)]
19. Ren, Y.; Bi, X.; He, Y.; Zhang, L.; Luo, L.; Li, L.; You, T. Research progress and applications of iron-based nanozymes in colorimetric sensing of agricultural pollutants. *Biosens. Bioelectron.* **2025**, *271*, 116999. [[CrossRef](#)]
20. Aalizadeh, M.; Yang, S.; Guntur, S.; Potluri, V.; Kulkarni, G.; Fan, X. Development of a Miniaturized, Automated, and Cost-Effective Device for Enzyme-Linked Immunosorbent Assay. *Sensors* **2025**, *25*, 5262. [[CrossRef](#)]
21. Pilavaki, E.; Demosthenous, A. Optimized Lateral Flow Immunoassay Reader for the Detection of Infectious Diseases in Developing Countries. *Sensors* **2017**, *17*, 2673. [[CrossRef](#)] [[PubMed](#)]
22. Herman, R.A.; Scherer, P.N.; Shan, G. Evaluation of logistic and polynomial models for fitting sandwich-ELISA calibration curves. *J. Immunol. Methods* **2008**, *339*, 245–258. [[CrossRef](#)] [[PubMed](#)]
23. Axelrod, T.; Eltzov, E.; Marks, R.S. Capture-Layer Lateral Flow Immunoassay: A New Platform Validated in the Detection and Quantification of Dengue NS1. *ACS Omega* **2020**, *5*, 10433–10440. [[CrossRef](#)]
24. Oluka, G.K.; Namubiru, P.; Kato, L.; Ankunda, V.; Gombe, B.; Cotten, M.; Team, C.-I.; Musenero, M.; Kaleebu, P.; Fox, J.; et al. Optimisation and Validation of a conventional ELISA and cut-offs for detecting and quantifying anti-SARS-CoV-2 Spike, RBD, and Nucleoprotein IgG, IgM, and IgA antibodies in Uganda. *Front. Immunol.* **2023**, *14*, 1113194. [[CrossRef](#)]
25. Subburaj, S.; Liu, C.; Xu, T. Emerging trends in AI-integrated optical biosensors for point-of-care diagnostics: Current status and future prospects. *Chem. Comm.* **2025**, *61*, 18464. [[CrossRef](#)]
26. Yang, J.; Li, G.; Chen, S.; Su, X.; Xu, D.; Zhai, Y.; Liu, Y.; Hu, G.; Guo, C.; Yang, H.B.; et al. Machine Learning-Assistant Colorimetric Sensor Arrays for Intelligent and Rapid Diagnosis of Urinary Tract Infection. *Chem. Comm.* **2025**, *61*, 18464. [[CrossRef](#)]
27. Lee, T.; Lee, H.-T.; Hong, J.; Roh, S.; Cheong, D.Y.; Lee, K.; Choi, Y.; Hong, Y.; Hwang, H.-J.; Lee, G. A Regression-Based Machine Learning Approach for pH and Glucose Detection with Redox-Sensitive Colorimetric Paper Sensors. *Sens. Actuators B Chem.* **2024**, *414*, 136012. [[CrossRef](#)] [[PubMed](#)]
28. Wang, S.; Zhu, Y.; Zhou, Z.; Luo, Y.; Huang, Y.; Liu, Y.; Xu, T. Integrated Ultrasound-Enrichment and Machine Learning in Colorimetric Lateral Flow Assay for Accurate and Sensitive Clinical Alzheimer's Biomarker Diagnosis. *Adv. Sci.* **2024**, *11*, e2406196. [[CrossRef](#)]
29. Han, F.; Xie, X.; Wang, T.; Cao, C.; Li, J.; Sun, T.; Liu, H.; Geng, S.; Wei, Z.; Li, J.; et al. Wearable Hydrogel-Based Epidermal Sensor with Thermal Compatibility and Long Term Stability for Smart Colorimetric Multi-Signals Monitoring. *Adv. Healthc. Mater.* **2023**, *12*, e2201730. [[CrossRef](#)]
30. Jin, H.; Zeng, W.; Qian, W.; Li, L.; Ji, P.; Li, Z.; He, D. Accelerated Reconstruction of ZIF-67 with Significantly Enhanced Glucose Detection Sensitivity. *Nano Res.* **2024**, *17*, 4737–4743. [[CrossRef](#)]
31. Aalizadeh, M.; Azmoudeh Afshar, M.; Fan, X. Machine Learning Enabled Multidimensional Data Utilization Through Multi-Resonance Architecture: A Pathway to Enhanced Accuracy in Biosensing. *ACS Omega* **2025**, *10*, 20713–20722. [[CrossRef](#)]
32. Aalizadeh, M.; Raut, C.; Azmoudeh Afshar, M.; Tabartehfarahani, A.; Fan, X. Accuracy Enhancement in Refractive Index Sensing via Full-Spectrum Machine Learning Modeling. *Biosensors* **2025**, *15*, 582. [[CrossRef](#)]
33. Rashed, A.R.; Gudulluoglu, B.; Yun, H.W.; Habib, M.; Boyaci, I.H.; Hong, S.H.; Ozbay, E.; Caglayan, H. Highly-Sensitive Refractive Index Sensing by Near-infrared Metatronic Nanocircuits. *Sci. Rep.* **2018**, *8*, 11457. [[CrossRef](#)]
34. Magnusson, R.; Wawro, D.; Zimmerman, S.; Ding, Y. Resonant photonic biosensors with polarization-based multiparametric discrimination in each channel. *Sensors* **2011**, *11*, 1476–1488. [[CrossRef](#)]
35. Yesudasu, V.; Pradhan, H.S.; Pandya, R.J. Recent progress in surface plasmon resonance based sensors: A comprehensive review. *Heliyon* **2021**, *7*, e06321. [[CrossRef](#)]
36. Bahrami, F.; Maisonneuve, M.; Meunier, M.; Aitchison, J.S.; Mojahedi, M. Self-referenced spectroscopy using plasmon waveguide resonance biosensor. *Biomed. Opt. Express* **2014**, *5*, 2481–2487. [[CrossRef](#)] [[PubMed](#)]

37. Psarouli, A.; Salapatas, A.; Botsialas, A.; Petrou, P.S.; Raptis, I.; Makarona, E.; Jobst, G.; Tukkiniemi, K.; Sopenan, M.; Stoffer, R.; et al. Monolithically integrated broad-band Mach-Zehnder interferometers for highly sensitive label-free detection of biomolecules through dual polarization optics. *Sci. Rep.* **2015**, *5*, 17600. [[CrossRef](#)] [[PubMed](#)]
38. Yan, X.; Shang, S.; Li, D.; Dang, Y. An efficient and interactive feature selection approach based on copula entropy for high-dimensional genetic data. *Sci. Rep.* **2025**, *15*, 30100. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.