Predicting Pigment Color Degradation with Time Series Models

Irina-Mihaela Ciortan [1], *Tina Grette Poulsson* [2], *Sony George* [1], *Jon Yngve Hardeberg* [1]; [1] *NTNU - Norwegian University of Science and Technology, Gjøvik, Norway;* [2] *National Museum of Norway, Oslo, Norway*

Abstract

The colors of pigments and dyes are affected by light exposure. Light-induced color change has an impact on various industrial and artistic applications where colored materials are frequently exposed to light throughout their life-cycle. For this reason, it is beneficial to understand the fading behaviour of pigments and simulate future degradation. In this article, we are proposing a method to forecast color change of pigments based on time series analysis. To begin with, we collect fading data from real objects with a microfadeometer, which records the color coordinates after every second of light exposure. Then, we treat this data as a time series, test for its stationarity and fit it with auto-regressive integrated moving average (ARIMA) models. Finally, using a train-test split, we validate the accuracy of the ARIMA models in predicting color degradation of pigments and dyes.

Introduction

While there are many biological, physical and chemical factors that contribute to the aging of pictorial materials, exposure to light is one of the most important. The reason to this is twofold. Firstly, light-induced damage is affecting the color attributes and so it has an immediate impact on the visual appearance. Secondly, light exposure is very difficult to avoid when a colored material is being utilized. If we take the cultural heritage field as an example application, there are centuries old paintings and drawings that have already been exposed to light so far and as a consequence, have an already deteriorated color. A famous example here are the paintings of Vincent van Gogh who used to a great extent red lake pigments that are known to be highly sensitive to light [5, 26]. The only way to completely protect artworks from light would imply to keep them out of public display and perhaps store them in a dark room. However, this doesn't stop the other aging mechanisms provoked by factors other than light. In addition, it might seem like a rather extreme measure, especially in today's age run by principles such as open access and art democratisation. To this purpose, museum curators, art conservators and conservation scientists have dedicated a lot of work in the recent years to design policies more suited to the individual objects or object groups, in order to protect the most sensitive artworks, but allow for longer display periods for artworks that are less sensitive.

The microfading measurement technique has played an essential role in the understanding of pigment color degradation which is, in its turn, crucial for developing the adequate lighting policies in museum exhibitions. In microfading, concentrated light is cast on points of submilimetric size and the colorimetric data is collected after every second of exposure. Given the small size of the measurement, microfading is considered a micro-destructive technique, which enables its use for real artworks. While microfading has a direct application on assessing the lightfastness of pigments, another relevant application can be the prediction of future change of the pigments beyond the measured data.

In this article, we investigate the fading behavior of pigments from a 19^{th} century pastel painting based on microfading colorimetry. Furthermore, we treat the temporal-varying color observations as a time series from a statistical perspective and use univariate auto-regressive integrated moving average models (ARIMA) to simulate the future change for each L*, a*, b* color coordinate individually.

Related Work

Microfading. Chan et al. [7] studied the lightfastness of pigments in Edward's Munch "The Scream (1910)" with a microfadeometer and discovered that the sky region painted with vermilion was the most sensitive to light. Their research served as motivation to enforce a new display strategy for the painting, where the painting is kept inside a black case that opens every several hours a day, on rotation with other paintings of similar sensitivity. Grimstad et al. [12] carried out a batch microfading measurement, collecting 176 data points from 27 paintings. The authors then used the Blue Wool Scale ISO standard [9] to qualitatively characterize the fading rate of all the pigments. They discovered that 5 points had a color change rate faster than Blue Wool 1, which is the most sensitive sample from the standard. The qualitative comparison with the Blue Wool standard is a *de facto* practice in fading analysis and is widely used by the scientific community [1, 13, 22]. Nevertheless, microfaded data may also be used directly without referring to the Blue Wool standard, as the measurement gives a concrete indication of the color difference in relation to light exposure.

Digital restoration and forecasting of color degradation. A couple of methods were proposed to predict future changes or to reconstruct the original color appearance of an artwork with or without accelerated light-induced aging experiments. A general drawback of these methods is given by the lack of a fully visual ground-truth. In most cases, there is no photographic documentation of a painting immediately after it was created and before the fading process occurred. Likewise, future simulations have not yet been attained to facilitate a direct comparison. Even so, the lack of perfect ground-truth is sometimes accounted for by art experts' validation [15, 4, 3] and by using existing knowledge of the artists' materials, obtained either from scientific analysis and/or from written records [17, 15]. For instance, Kirchner et al. [17] used both Vincent van Gogh's letters and knowledge of the painter's pigment palette to reconstruct the original colors of "Field with Irises near Arles". The authors created a mockup of base pigments using historical recipes similar to van Gogh's technique. Then, they computed concentration maps for every base

pigment in the painting using Kubelka-Munk theory. Finally, they recovered the color appearance of the original painting by combining the concentration maps with the pigment properties derived from the unaged mockup.

Another body of works implements artistic color change prediction by manipulating digitized versions of artworks in a commercial image editing software [15, 3, 4]. The manipulations are done at various degrees and then psycho-physical experiments are conducted where art historians, museum curators and conservation scientists rank the digital simulations according to their feasibility from a perceptual point of view. Hendriks et al. [15] employed this approach towards virtually restoring the unaged colors of van Gogh's "The Bedroom". With a related processing method, Brokerhof et al. [3] simulated future fading of a collection of maps to investigate the threshold for critical change awareness with a perceptual experiment. In a similar way, Brokerhof et al [4] operated a commercial image editing software to simulate fading of a series of Japanese woodblock prints for a psycho-physical experiment. As input to the digital simulations, they used color differences of prints on-display relative to prints kept in darkness.

There are few attempts in the literature that use the microfading data for predicting future changes beyond measured observations. Thomas et al. [25] applied classical regression on the color difference time series and simulate changes for both ends of history: past and future. They compensated for the lack of ground-truth visualizations, by closely working with art and conservation experts and through a vast knowledge of the pigment palette. Riutort et al. [21] proposed an interpolation method based on Gaussian process that achieves a spatio-temporal extension of the microfaded colorimetric data collected from rock art paintings in Spanish prehistoric caves. For the temporal dimension, they enforced two constraints by including the first derivative information in their model: long-term stabilization and monotonicity. The former is based on the assumption that color change will eventually plateau, when the pigments have become stable to light exposure and thus no further fading happens. The latter restriction refers to the observed behaviour of the authors that color changes will be either sequentially increasing or decreasing, without being able to bounce to a reverse direction. Cross-validation is used to test the performance of their approach [21].

ARIMA. Auto-regressive integrated moving average (ARIMA) model was proposed in 1970 by Box and Jenkins [2] as an effective statistical technique for analyzing time-series and forecasting future data. ARIMA is a generalized expression of many forecasting models, including exponential models, random walk and stationarized regression models [18]. For this reason, it has the advantage of being applicable to a wide range of temporal distributions. Indeed, ARIMA models were successfully used for predicting data in numerous applications of economical, environmental and civil nature. For instance, ARIMA is used by Zhang et al. [27] to simulate GDP growth of Sweden, and by Jere et al. [16] to assess the trend of second-car imports in Zimbabwe. Similarly, ARIMA were shown to be a good fit for simulating daily transportation habits in the United States by Pavyluk et al. [19] and for estimating the carbon dioxide emissions in Bangladesh by Rahman et al. [20]. In addition, there are several studies on using auto-regressive models for predicting the incidence of criminal activities [8, 6].

While the list of applications mentioned here is by far non-exhaustive, it is important to highlight that, to the best of our knowledge, there is no other work that analyzes colorimetric data with ARIMA models.

Materials and Method

In this work, we collected colorimetric data of real objects with a microfadeometer, characterized their fading behaviour and searched for a good fit with ARIMA models towards predicting future alteration.

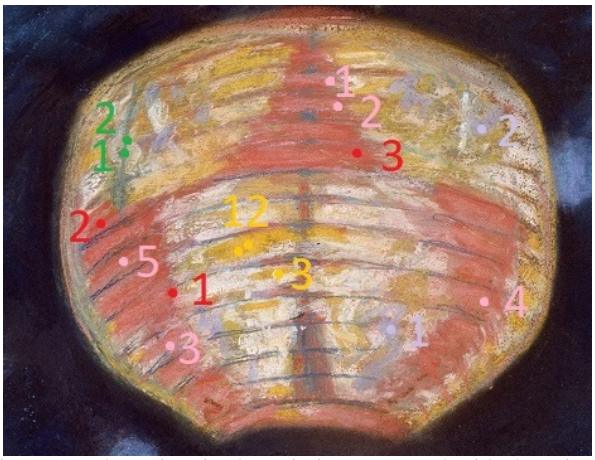

Figure 1: The points in the painting measured with the microfadeometer, corresponding to five colour groups: pink, red, green, orange, violet blue.

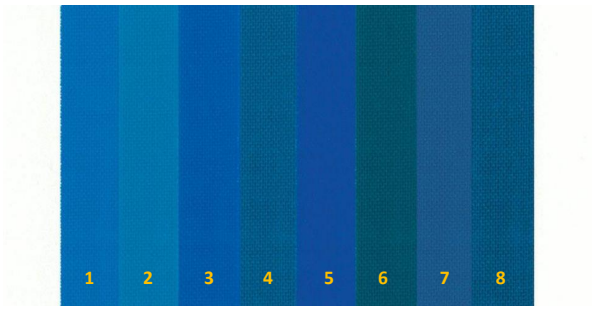

Figure 2: The ISO Blue Wool Standard used for material lightfastness assessment. The standard is made of blue-dyed wool stripes, presenting an 8-level scale of light sensitivity, where 1 indicates higher sensitivity.

Measured objects

The first real world object in our study is represented by the pastel painting "A Japanese Lantern" by Oda Krohg, from the collection of the National Museum of Norway (inventory number NG.M.00879). The painting dates back to the second half of the 19^{th} century and is made on canvas. In this article, only a fragment of the painting is studied (displayed in fig. 1). The second object in our study is the ISO Blue Wool Scale (BWS) standard[14, 9] (see fig. 2), which is used as reference for assessing the fading rate of the samples in the painting.

Microfading experiment

A total of 18 points (15 in the paintings and 3 for the BWS, stripe 1, 2, 3), highlighted in fig. 1 and 2, were acquired with a microfadeometer provided by Instytut FotoNowy [10]. The instrument has a 5500 K white LED source and an aperture size of 0.5 mm. The measurements were gathered at a 0°/45° geometry (see fig. 3), for a light dosage of 12.585 Mlux and 0.0034 Watts. The L*, a*, b* color coordinates were recorded after every second of light exposure for D65 illuminant and the CIE 1931 2° standard observer. The accelerated aging experiment was terminated once the ΔE_{00} reached 2 units or after 600 iterations in order to make sure that the actual damage on the painting is kept to a minimum. Therefore, after the microfading experiment, 54 time series were obtained, i.e. the individual alterations of the L*, a*, b* coordinates for all samples.

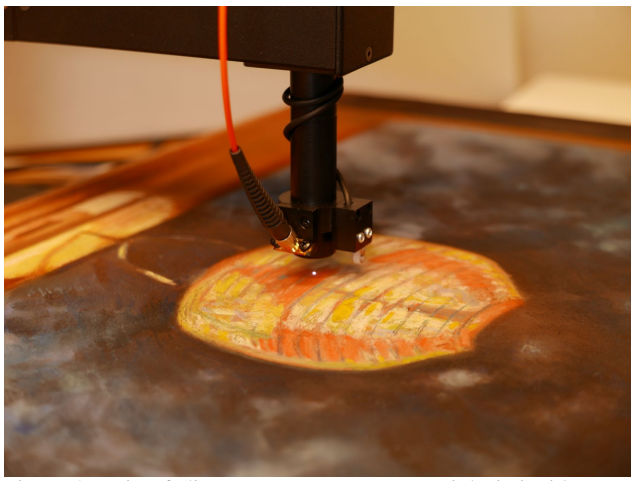

Figure 3: Microfading measurement setup. Light is incident on the surface at 0° and the colorimeter is collecting the signal at 45°.

ARIMA modelling

ARIMA gathers instances of several forecasting models into a single, generical formulation, and is suitable for stationary time series. Mathematically, it is represented by a sum of two submodels: auto-regressive (AR) and moving average (MA) [24]. The principle of AR models is that given a stationary time series, we can predict future values of the dependent variable as a linear combination of its past observations. Similarly, the MA component assumes that the residuals also linearly depend on previous errors. The integrated (I) feature of the ARIMA method refers to the order of the derivative operation required to be applied on the original data to enforce stationarity.

A stationary time series is the one where no strong increasing or decreasing trend is visible over time, implying temporally constant mean and variance. In order to assess the stationarity of the time series in a quantitative way, statistical tests, such as the augmented Dickey Fuller (ADF) [11] test can be verified. The null hypothesis of the ADF test is that the series is non-stationary. The output of the test, called the ADF statistic should be negative and the more negative it is, the better the interpretation, as it means that the time series is stationary. The test includes reference critical values, that can be checked against the output statistic. The critical values slightly increase with the number of observations in the time series. For the cardinality of our dataset, the critical values vary between -3.58 and -3.43 at 99% confidence level, and between -2.93 and -2.87 at 95% confidence level. Moreover, a p-value lower than 0.05 indicates stationarity as it significantly rejects the null hypothesis. If an original time-series is not stationary, there are several operations that could remove an existing trend in the data, such as the first derivative, also known as first-order differencing.

Thus, an ARMA (AR + MA) model without an I component is a model that can be applied where the original time-series is stationary and needs no differencing operation. Given a stationary time series y_t with t total data points, by modelling it with $ARMA(p,q)$ we are solving the following equation:

$$\begin{aligned} y_t &= c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + \alpha_p y_{t-p} \\ &+ \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} ... + \beta_q \varepsilon_{t-q} \end{aligned} \quad (1)$$

order p is the number of past observations of y that are input to the model and order q the number of previous residuals that influence the error ε at the current time step (also called lag) t. c, the constant term, together with α and β are the polynomial coefficients that are estimated during the fitting process, using an iterative maximum loglikelihood method. More precisely, the conditional sum of squares likelihood is maximized and its values are used to initialize the model.

If we apply ARIMA model, then a new order d is introduced, which refers to the order of the derivative transformation, required to stationarize the original data, y_t. For example, an $ARIMA(p,d,q)$ model with $d = 1$ considers the difference between a time step and the one before it:

$$\begin{aligned} y_t &= c + \alpha_1 (y_t - y_{t-1}) + \alpha_2 (y_{t-1} - y_{t-2}) + ... \\ &+ \alpha_p (y_{t-p+1} - y_{t-p}) + \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} ... + \beta_q \varepsilon_{t-q} \end{aligned} \quad (2)$$

Orders p,d,q are positive, integer numbers. When defining an ARIMA model for a stationary time series, the order d is chosen as 0. If the series is non-stationary, then d is increased with increments of 1, where an order higher than 1 indicates a quadratic behaviour of the series. One systematic way to find the p and q orders is through graphical examination of the autocorrelation function (ACF) and partial autocorrelation (PACF) functions [24], which also carry the name of correlograms. Whereas ACF captures the self-correlation of the time series with itself, PACF captures the self-correlation after removing intermediate lags. In principle, the number of consecutive significant lags that exceed the 95% confidence level is chosen from ACF and PACF as the orders q and p, respectively.

Results

One straightforward result of the microfading experiment is the plot of the number of iterations as an indicator of lightfastness. Fig. 4 shows the samples ordered by the total number of seconds of exposure to light. It is easy to observe that Pink 1, 5, 4, and 2 are the samples most rapidly fading, with a sensitivity higher than BWS 1. On the other hand, the Green and Orange samples seem to be the most resistant, similar to BWS 3. However, in order to better perceive the rate of change, fig. 5 presents the evolution of ΔE_{00} color difference with respect to the original, unfaded sample

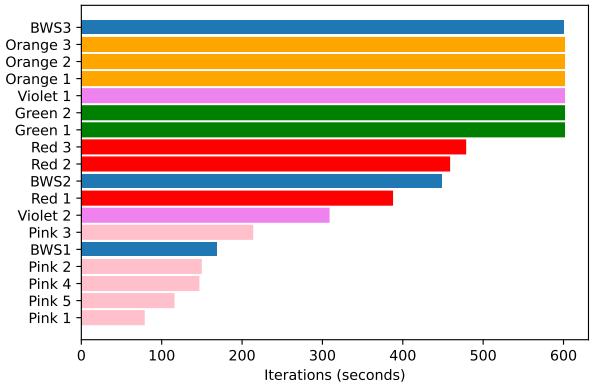

Figure 4: Light sensitivity of the pigment samples and the first 3 Blue Wool stripes (BWS1, BWS2, BWS3) quantified as number of seconds of exposure to light. In ideal experimental conditions, the smaller the bar size, the higher the sensitivity to light.

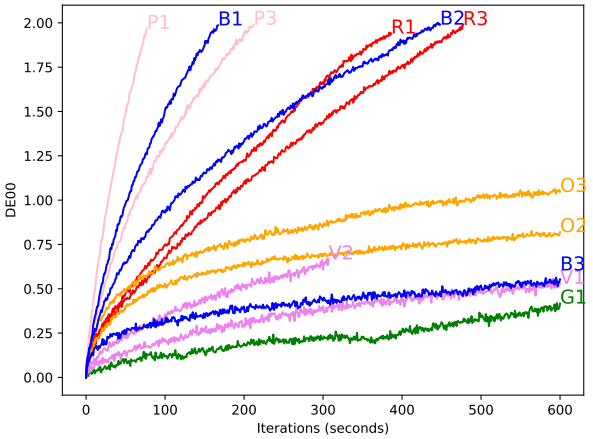

Figure 5: ΔE_{00} as a function of time for selected samples, which shows the rate of color change. The measurement of sample Violet 2 was unexpectedly interrupted, before 2 units of color change were reached, and before 600 iterations.

after every second of light exposure. In this plot, we can also notice that due to experimental errors, the measurement of the Violet 2 sample was unexpectedly stopped, even though the limits of $2\Delta E_{00}$ units, nor the cap of 600 iterations was reached. So in this case, the number of iterations is not a good clue for the fugitivity of the sample, and Violet 2 seems to have a lightfastness comparable to BWS3 rather than BWS1 or BWS2.

Because our samples have a wide range of light-induced sensitivities, we applied the ADF test to check for the stationarity of each time series. We treated each set of values of the three colorimetric L*, a*, b* coordinates as univariate time series. Table 1 reports the p-value and ADF statistic for the stationarity test. For several series, the p-value was higher than 0.05 and so, the null hypothesis (the time series is non-stationary) could not be rejected. In these cases, the first derivative was computed and another instance of the ADF test was run. After the first derivative transformation, all the original non-stationary series became stationary, as it can be observed from the results in Table 2. It is interesting to note that the sensitivity to light as characterized by the rate of color degradation, ΔE_{00}, is not immediately translated to the concept of stationarity. This is understandable, since the ΔE_{00} color difference is a global aggregate of the L*, a*, b* colorimetric attributes, but when taken individually, these coordinates might manifest a different degradation pattern.

Table 1: Stationarity check for the L*, a*, b* coordinates using Augmented-Dickey Fuller test. A p-value lower than 0.05 and a negative test statistic (adf_stat) lower than approximately -3 mark stationarity (S) of the time series, while the opposite indicates non-stationarity (NS).

	L*			a*			b*		
	p-val	**adf_stat**		**p-val**	**adf_stat**		**p-val**	**adf_stat**	
Pink 1	p<0.001	-5.92	S	p<0.001	-6.47	S	0.002	-3.89	S
Pink 2	p<0.001	-6.97	S	p<0.001	-6.44	S	p<0.001	-4.69	S
Pink 3	p<0.001	-7.51	S	p<0.001	-9.04	S	0.001	-4.06	S
Pink 4	*0.42*	*-1.71*	*NS*	p<0.001	-4.12	S	p<0.001	-4.28	S
Pink 5	0.004	-3.70	S	p<0.001	-7.44	S	p<0.001	-4.94	S
Red 1	p<0.001	-5.11	S	p<0.001	-4.79	S	*0.09*	*-2.62*	*NS*
Red 2	p<0.001	-5.70	S	p<0.001	-6.95	S	*0.17*	*-2.31*	*NS*
Red 3	p<0.001	-5.03	S	p<0.001	-7.47	S	*0.24*	*-2.12*	*NS*
Green 1	p<0.001	-4.44	S	*0.91*	*-0.41*	*NS*	*0.96*	*0.04*	*NS*
Green 2	p<0.001	-5.10	S	*0.56*	*-1.45*	*NS*	0.03	-3.02	S
Violet 1	p<0.001	-5.30	S	*0.80*	*-0.86*	*NS*	*0.22*	*-2.16*	*NS*
Violet 2	p<0.001	-6.53	S	*0.29*	*-2.00*	*NS*	*0.70*	*-1.15*	*NS*
Orange 1	p<0.001	-7.24	S	*0.10*	*-2.57*	*NS*	p<0.001	-5.46	S
Orange 2	p<0.001	-9.45	S	p<0.001	-5.16	S	0.001	-4.03	S
Orange 3	0.002	-3.90	S	p<0.001	-5.01	S	0.003	-3.83	S
BWS 1	p<0.001	-6.45	S	p<0.001	-5.38	S	*0.07*	*-2.70*	*NS*
BWS 2	p<0.001	-6.76	S	p<0.001	-6.38	S	0.02	-3.22	S
BWS 3	*0.94*	*-0.15*	*NS*	*0.98*	*0.38*	*NS*	p<0.001	-6.19	S

Table 2: Augmented Dickey Fuller test for the L*, a*, b* coordinates after applying the first derivative transformation on the original non-stationary series. Now, the previously non-stationary series pass the stationarity test.

	1st derivative L*			1st derivative a*			1st derivative b*		
	p-val	**adf_stat**		**p-val**	**adf_stat**		**p-val**	**adf_stat**	
Pink 4	0.01	-3.34	S						
Red 1							p<0.001	-9.78	S
Red 2							p<0.001	-11.27	S
Red 3							p<0.001	-15.04	S
Green 1				p<0.001	-9.84	S	p<0.001	-6.22	S
Green 2				p<0.001	-9.97	S			
Violet 1				p<0.001	-10.60	S	p<0.001	-11.67	S
Violet 2				p<0.001	-7.77	S	p<0.001	-9.17	S
Orange 1				p<0.001	-10.28	S			
BWS 1							p<0.001	-4.16	S
BWS 2									
BWS 3	p<0.001	-11.16	S	p<0.001	-10.90	S			

After this initial step, the d order of the $ARIMA(p,d,q)$ as expressed in eq. 2 was defined as either 0 (for the originally stationary data, in which case eq. 1 is applied) or 1 (for the originally non-stationary data), for each color coordinate. Subsequently, we checked the patterns in the ACF and PACF to identify the p and q orders. For the series where the ACF showed many significant lags slowly declining to zero and a PACF with only few p significant lags, the q was identified as 0, because this is typical of a auto-regressive model, with no moving average component as explained in [24]. Fig. 6 displays an example of such pattern, representing Pink 1 L* time series. Arguably, such ACF pattern, showing many positive correlations, could also be interpreted as a clue for non-stationarity still contained in the data [18]. While Pink 1 L* time series passed the ADF stationarity test, in future work, it can be explored whether a model that considers a first derivative transformation would give a better fit. For the rest of the series, where the ACF shows only one significant lag, we are dealing with a mixed auto-regressive and moving average model, with $q = 1$ and p equal to the number of spikes in the PACF plot. For instance, the correlograms of Violet 2 b* sample indicate a $p = 5$, as it can be seen in fig. 7.

This same procedure was followed to identify the orders of

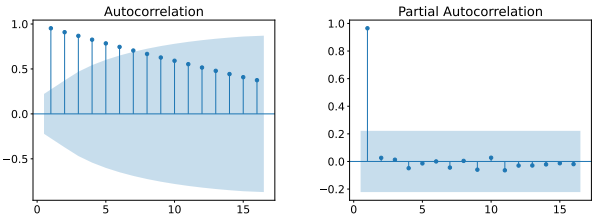

Figure 6: Correlograms for Pink 1 L* time series, where x-axis corresponds to time steps/lags and y-axis to correlation magnitude. Blue area depicts the 95% confidence interval. ACF exhibits many spikes slowly declining to zero, while PACF shows only one time step as significant. This is characteristic of an autoregressive model, in this case, an ARIMA (1,0,0) model.

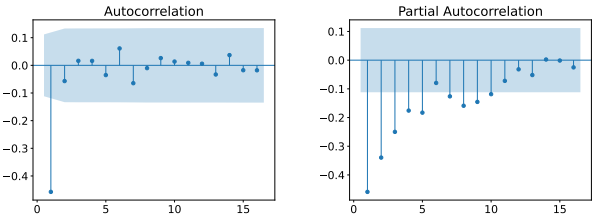

Figure 7: Correlograms for Violet 2 b* time series, where x-axis corresponds to time steps/lags and y-axis to correlation magnitude. Blue area depicts the 95% confidence interval. The original data was transformed to its first derivative to achieve stationarity. In ACF, only the first lag exceeds the confidence level, while in PACF we can count 5 consecutive significant lags. This indicates an ARIMA (5,1,1) model.

the ARIMA models (listed in Table 3) for all color coordinates corresponding to the 18 microfaded samples. It can be observed that in the case of Pink 4 L* and BWS 1 b*, no autoregressive and no moving average components were identified because of the lack of significant lags in the ACF and PACF. ARIMA(0,1,0) is an instance of a random walk model [18], i.e. the forecast of a variable equals its previous value in the time-series, which indicates a plateau of the data. The plateauing may be physically relevant, given that color change of pigments has previously been found [21] to reach a maximum point followed by no increment in value.

In order to assess the performance of the ARIMA model, the data was split into a training and test set using a 80%-20% ratio. Table 3 reports the following indices of performance: the Bayesian Information Criterion (BIC) [23] of the trained model and the sum of squared residuals computed for the test data. A lower value of both these metrics correspond to a better performance. Overall, the BIC values are very low suggesting that the trained models are a good fit for the input data. Similarly, the sum of squared residuals is in most cases under unity. There are few samples, such as Red 1 a*, Red 3 a* and BWS 2 b* that have high sums of squared residuals. To improve the fit, more search could be done in the p,d,q parameter space. However, the scope of this article is not to find the best fit in absolute terms, as it is more to prove the validity of ARIMA in modelling the temporal evolution of colorimetric attributes.

Fig. 8 shows one the prediction of some of the models with the best error metrics, corresponding to the Violet 2 color coordinates. It is interesting to note that in case of the Violet 2 sample, the predictions of the ARIMA models for the training set are less noisy than the original data, especially for a* and b* series, showing the capability of the model to learn the stable characteristics of the data and to separate the unstable, erroneous component. The smoothness is preserved for the forecasts of the test data as well. A similar case is that of Green 1 sample, plotted in fig. 9, where there is a lot of noise present in the measured time series of the a* coordinate. Nevertheless, ARIMA(8,1,1) manages to discard the noise and capture the essence of the change behaviour. A particular good result is obtained for Green 1 b* coordinate that has a very unstable temporal behaviour. Considering only 3 observations from the past, the ARIMA model learn well to the degradation trend. Another result, similar in terms of instability is that of Violet 2 a* (fig. 10). Actually, this sample is corrupted by high instability as well as a big amount of noise. Both these factors are surpassed by the ARIMA(5,1,0) model.

Fig. 11 visually describes the performance of the model in predicting both the training and test observations for the Red 3 a* time series. While the residuals of the training set closely follow the original values, the forecast converges to a constant value, that even though potentially physically relevant [21], fails to capture the specific variation of the test set. From one perspective, this suggests that the model is not over-fitting. Nonetheless, this happens at the cost of higher residual values. In future analysis, it would be relevant to add an extensive statistical investigation of the residuals and the model coefficients towards improving the currents fit.

Conclusion

In this article we showed the potential of ARIMA models in predicting the temporal variation of pigment color coordinates provoked by light exposure. We achieved good results by analyzing each of the L*, a*, b* coordinates as univariate time series. Nonetheless, as future work, it would be interesting to approach this problem as a multivariate time series and forecast the change of the three color dimensions simultaneously. Our work can be useful for any application that requires future simulation of colorimetric temporal change. Specifically, in this paper, we proved the direct application of ARIMA models for the cultural heritage field, where the understanding and visualization of how an artwork deteriorates with time is essential for its conservation and preservation.

References

[1] Marthe Aambø, Magdalena Godzimirska, Emma Chan, Tomasz Lojewski, and Irina C.A. Sandu. Light Sensitivity of Pigments in Edvard Munch's Works on Paper. In *MUNCH2022: Understanding Munch and the Art at the turn of the Centuries - between the Museum and the Laboratory*, page 100, Oslo, March 2022.

[2] George EP Box, Gwilym M Jenkins, and G Reinsel. Time series analysis: forecasting and control. *Holden Day, San Franciso*, 1970.

[3] Agnes Brokerhof, Pieter Kuiper, and Steph Scholten. Spread or sacrifice: Dilemma for lighting policies. *Studies in Conservation*, 63(sup1):28–34, Aug 2018.

[4] Agnes W. Brokerhof, Margrit Reuss, Fiona MacKinnon, Frank Ligterink, Han Neevel, Farideh Fekrsanati, and Graeme Scott. Optimum Access at Minimum Risk: The

Table 3: Orders of the ARIMA models and performance metrics (Bayes Information Criterion and sum of squared residuals) for every L*, a*, b* coordinate.

	L*					a*					b*				
	p	**d**	**q**	**BIC**	**ssr test**	**p**	**d**	**q**	**BIC**	**ssr test**	**p**	**d**	**q**	**BIC**	**ssr test**
Pink 1	1	0	0	-271	0.21	1	0	0	-175	1.42	1	0	0	-163	0.24
Pink 2	1	0	0	-594	0.34	1	0	0	-486	1.16	1	0	0	-368	0.24
Pink 3	1	0	0	-905	0.45	1	0	0	-748	2.24	1	0	0	-553	1.27
Pink 4	0	1	0	-692	0.27	1	0	0	-479	1.42	1	0	0	-337	0.35
Pink 5	1	0	0	-446	0.21	1	0	0	-327	1.37	1	0	0	-270	0.85
Red 1	1	0	0	-1757	1.21	1	0	0	-1444	7.20	3	1	1	-1001	1.25
Red 2	1	0	0	-2055	1.50	1	0	0	-1681	4.62	3	1	1	-1134	0.70
Red 3	1	0	0	-2176	1.26	1	0	0	-1762	10.57	6	1	1	-1224	0.89
Green 1	3	0	0	-2923	0.04	8	1	1	-2478	0.03	3	1	1	-1985	0.27
Green 2	3	0	0	-2968	0.04	6	1	1	-2505	0.06	5	0	0	-1985	0.31
Violet 1	3	0	0	-2906	0.01	5	1	1	-2660	0.03	5	1	1	-2185	0.16
Violet 2	2	0	0	-1443	0.03	5	1	1	-1350	0.02	5	1	1	-1128	0.06
Orange 1	2	0	0	-2808	0.08	7	1	1	-2853	0.05	4	0	0	-1652	1.60
Orange 2	2	0	0	-2881	0.14	3	0	0	-2742	0.08	4	0	0	-1431	2.14
Orange 3	2	0	0	-2768	0.21	4	0	0	-2746	0.06	4	0	0	-1345	3.47
BWS 1	1	0	0	-665	0.49	1	0	0	-492	0.18	0	1	0	-518	5.08
BWS 2	1	0	0	-2046	1.20	2	0	0	-1550	0.76	1	0	0	-1516	6.34
BWS 3	4	1	1	-2744	0.08	5	1	1	-2084	0.18	2	0	0	-2131	2.54

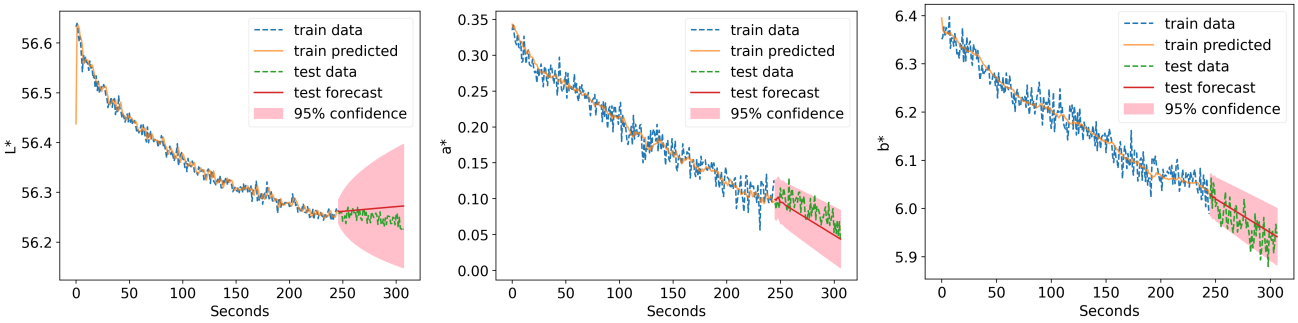

Figure 8: Prediction of the change of the L* (left), a* (middle) and b* (right) coordinates for the Violet 2 sample. The ARIMA models used for the forecast give very low sum of squared residuals for this sample.

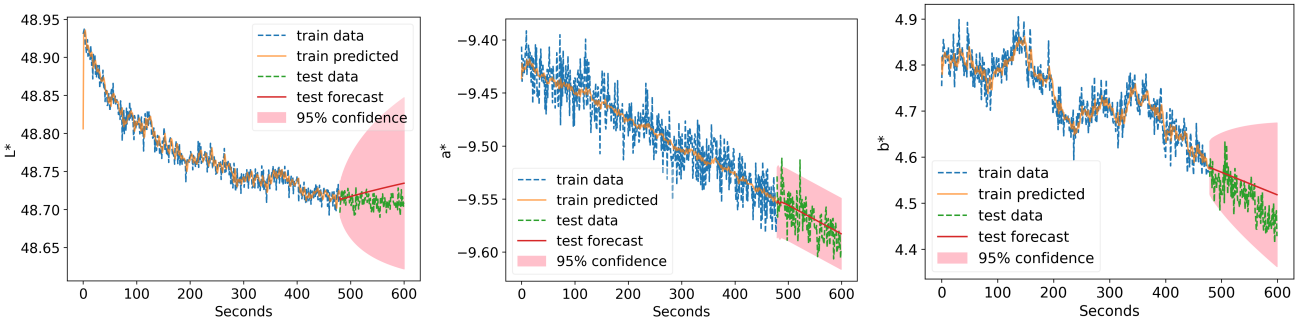

Figure 9: Prediction of the change of the L* (left), a* (middle) and b* (right) coordinates for the Green 1 sample. The b* coordinate is the very unstable. However, the ARIMA(3,1,1) model gives a smoother prediction, being able to discard the noise in the data.

Dilemma of Displaying Japanese Woodblock Prints. *Studies in Conservation*, 53(sup1):82–87, Jan 2008.

[5] Aviva Burnstock, Ibby Lanfear, Klaas Jan van den Berg, Leslie Carlyle, Mark Clarke, Ella Hendriks, and Jo Kirby. Comparison of the fading and surface deterioration of red lake pigments in six paintings by Vincent van Gogh with artificially aged paint reconstructions. In *Proceedings of the 14th Triennial Meeting of the ICOM Committee for Conservation Meeting in Den Haag (I. Vergier ed), Preprint book I, James and James, London. p. 459*, volume 466, 2005.

[6] Eugenio Cesario, Charlie Catlett, and Domenico Talia. Forecasting crimes using autoregressive models. In *2016*

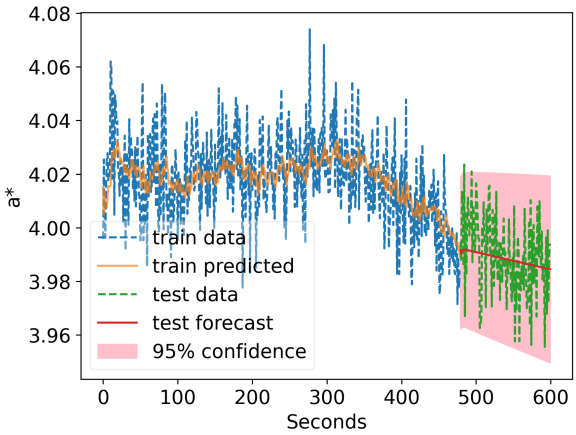

Figure 10: Forecast of the a* coordinate for the Violet 1 sample, fitted with ARIMA (5,1,0) model. Despite the instability of the data, the model manages to learn well the aging pattern.

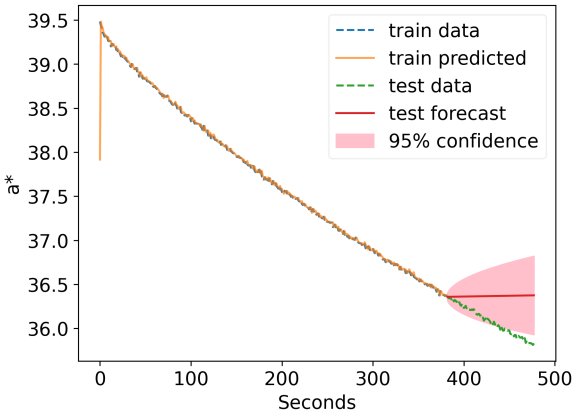

Figure 11: Forecast of the a* coordinate for the Red 3 sample, that has the highest sum of squared residuals based on ARIMA (1,0,0) model. We can notice that the forecast plateaus to a constant value, instead of adapting to the particularities of the test data.

IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), pages 795–802. IEEE, 2016.

[7] Emma Chan, Marthe Aambø, Magdalena Godzimirska, Inger Grimstad, Tomasz Lojewski, and Irina C.A. Sandu. Light-induced Color Changes on "The Scream" Versions in the Munch Museum Collection. In *MUNCH2022: Understanding Munch and the Art at the turn of the Centuries - between the Museum and the Laboratory*, page 91, Oslo, March 2022.

[8] Peng Chen, Hongyong Yuan, and Xueming Shu. Forecasting crime using the arima model. In *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, volume 5, pages 627–630. IEEE, 2008.

[9] International Organization for Standardization. ISO 105-B02:2014 Textiles — Tests for colour fastness — Part B02: Colour fastness to artificial light, 2014. Retrieved September 10, 2022 from `https://www.iso.org/obp/ui/#iso:std:iso:105:-B02:ed-6:v1:en`.

[10] Instytut Fotonowy. Micro fading tester, 2022. Retrieved September 10, 2022 from `https://www.fotonowy.pl/products/micro-fading-tester/?lang=en`.

[11] Wayne A. Fuller. Introduction to statistical time series. *John Wiley, New York*, 1976.

[12] Inger Grimstad, Tomasz Lojewski, and Irina C.A. Sandu. Is the Interpretation and Application of Collected Microfading Data Straight Forward? In *MUNCH2022: Understanding Munch and the Art at the turn of the Centuries - between the Museum and the Laboratory*, page 103, Oslo, March 2022.

[13] Abed Haddad and Laura Neufeld. Realizing Sensations: Balancing the Exhibition and Preservation of Cézanne's Watercolors. 2022.

[14] James Heal. Blue wool standards, 2022. Retrieved September 10, 2022 from `https://www.jamesheal.com/essentials-blue-wool-standards-how-to-use`.

[15] E. Hendriks, A.W. Brokerhof, K. van den Meiracker, J. Bridgland, and AHM (FGw). Valuing van Gogh's colours. In *ICOM-CC 18th Triennial Conference Copenhagen*. International Council of Museums, 2017.

[16] Stanley Jere, Bornwell Kasense, and Bwalya Bupe Bwalya. Univariate Time-Series analysis of second-hand car importation in Zambia. *Open Journal of Statistics*, 7(4):718–730, 2017.

[17] Eric Kirchner, Ivo van der Lans, Frank Ligterink, Muriel Geldof, Luc Megens, Teio Meedendorp, Kathrin Pilz, and Ella Hendriks. Digitally reconstructing Van Gogh's Field with Irises near Arles part 3: Determining the original colors. *Color Research Application*, 43(3):311–327, 2018.

[18] Robert Nau. Statistical forecasting: notes on regression and time series analysis, 2022. Retrieved September 10, 2022 from `https://people.duke.edu/~rnau/411home.htm`.

[19] Dmitry Pavlyuk. Short-term traffic forecasting using multivariate autoregressive models. *Procedia Engineering*, 178:57–66, 2017.

[20] Abdur Rahman and Md Mahmudul Hasan. Modeling and forecasting of carbon dioxide emissions in Bangladesh using Autoregressive Integrated Moving Average (ARIMA) models. *Open Journal of Statistics*, 7(4):560–566, 2017.

[21] Gabriel Riutort-Mayol, Virgilio Gómez-Rubio, José Luis Lerma, and Julio M del Hoyo-Meléndez. Correlated functional models with derivative information for modeling microfading spectrometry data on rock art paintings. *Mathematics*, 8(12):2141, 2020.

[22] David Saunders and Jo Kirby. Light-induced colour changes in red and yellow lake pigments. *National Gallery Technical Bulletin*, 15(1):79–97, 1994.

[23] Gideon Schwarz. Estimating the dimension of a model. *The annals of statistics*, pages 461–464, 1978.

[24] Barbara G Tabachnick and Linda S Fidell. *Using multivariate statistics*, volume 7. Pearson Boston, MA, 2018.

[25] Jacob Thomas, Tina Grette Poulsson, and Thierry Ford. The changing face of Munch's scream. In *MUNCH2022: Understanding Munch and the Art at the turn of the Centuries - between the Museum and the Laboratory*, page 88, Oslo,

Mar 2022.

[26] van Gogh Museum. Revigo project (REassessing VIncent van GOgh). Retrieved September 10, 2022 from `https://www.vangoghmuseum.nl/en/about/knowledge-and-research/completed-research-projects/revigo`.

[27] Haonan Zhang and Niklas Rudholm. Modeling and forecasting regional GDP in Sweden using autoregressive models. *Dalama University*, 2013.

Author Biography

Irina-Mihaela Ciortan *is an early-stage researcher in Visual Computing Techniques for Cultural Heritage. Currently a PhD candidate at NTNU, Irina holds a joint MSc diploma in Spectral Sciences and Multimedia Technologies awarded by the 4-university Erasmus Mundus Consortium "Colour in Informatics and Media Technology" (2013) and a BSc in Informatics issued by the Faculty of Cybernetics, Statistics and Informatics in Bucharest, Romania (2011).*

Tina Grette Poulsson *is a paper conservator at the National Museum, Oslo since 2006; worked at the Museum Conservation Services, Cambridge, 2004-2006; holds an MA from the University of Oslo, Camberwell College of Arts, London and the previous National Gallery, Oslo; research interests: 19th Century drawing materials, colour changes in paper-based art works.*

Sony George *is currently Associate Professor at The Colourlab, NTNU since 2017. Before joining NTNU, he has worked as a researcher at Gjøvik University College, Norway. Sony obtained a PhD in Photonics from the Cochin University of Science and Technology, India in 2012. His research interests are in the field of colour imaging, spectral image processing, image quality, etc.*

Jon Yngve Hardeberg *received his MSc degree in signal processing from the Norwegian Institute of Technology in Trondheim, Norway in 1995, and his PhD from Ecole Nationale Supérieure des Télécommunications in Paris, France in 1999. After a short but extremely valuable industry career near Seattle, Washington, where he designed, implemented, and evaluated colour imaging system solutions for multifunction peripherals and other imaging devices and systems, he returned to academia and Norway in 2001. He is currently Professor of Colour Imaging at the Department of Computer Science at NTNU and member of the Colourlab.*