Énergie Matériaux Télécommunications

QUANTIFICATION DU NETTOYAGE AU LASER INFRAROUGE PULSÉ FEMTOSECONDE SUR DES SURFACES À BASE DE CELLULOSE DANS LE DOMAINE DE LA CONSERVATION DU PATRIMOINE CULTUREL

Par

Canan Yağmur Boynukara

Thèse présentée pour l'obtention du grade de
Philosophiae Doctor, Ph.D.
en Sciences de l'énergie et des matériaux

Jury d'évaluation

Président du jury et examinateur interne	Aycan Yurtsever INRS- ÉMT, Varennes, Canada
Examinateur externe	Oliver Boine Frankenheim Technische Universität Darmstadt, Germany
Examinateur externe	Gianluca Sarri Queen's University Belfast, United Kingdom
Directeur de recherche	Patrizio Antici INRS-ÉMT, Varennes, Canada
	Mauro Migliorati Sapienza Università di Roma, Italy
Codirecteur de recherche	Andreas Peter Ruediger INRS- ÉMT, Varennes, Canada

SAPIENZA
UNIVERSITÀ DI ROMA

INRS
Institut national
de la recherche
scientifique

Quantification of femtosecond pulsed infrared laser cleaning on cellulose-based surfaces in cultural heritage conservation

Department of Physics
PhD in Accelerator Physics (XXXVI cycle)

Canan Yağmur Boynukara
ID number 1951032

Advisors
Prof. Mauro Migliorati
Prof. Patrizio Antici

Co-Advisor
Prof. Andreas Peter Ruediger

Academic Year 2024/2025

Thesis defended on 29 September 2025

In front of a Board of Examiners composed by:
Prof. Catia Milardi
Prof. Oliver Boine-Frankenheim
Dr. Gianluca Sarri

Quantification of femtosecond pulsed infrared laser cleaning on cellulose-based surfaces in cultural heritage conservation
PhD Thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: cananyagmur.boynukara@uniroma1.it, canan.boynukara@inrs.ca

In memory of all the women whose lives were taken by the tragedy of femicide...

Abstract

Femtosecond (fs) laser cleaning is emerging as a promising method for the conservation of cultural heritage materials, yet challenges remain in quantifying cleaning performance, particularly on porous and morphologically heterogeneous substrates.

In this doctoral thesis, conducted at La Sapienza University in Rome and the Institut National de la Recherche Scientifique (INRS) in Montréal, a quantitative and spectroscopy-anchored evaluation framework for near-infrared (NIR) fs-laser cleaning has been developed.

The first part introduces fundamental aspects of laser–material interactions, with particular emphasis on mechanisms underpinning selective cleaning, alongside a review of diagnostic techniques such as ATR-FTIR spectroscopy, colorimetry, scanning electron microscopy, and profilometry applied in heritage conservation.

The second part quantitatively investigates NIR fs-laser cleaning on cellulose-based mock-up samples replicating historical papers, which were artificially aged and contaminated with graphite and kaolinite as model pollutants. The cleaning performance was assessed by ATR-FTIR spectroscopy coupled with exponential saturation and decay models, demonstrating effective removal of surface contaminants within a defined safe fluence window without compromising substrate integrity.

The third part explores fs-laser cleaning of historically prepared mock-up leathers, emphasizing that microstructural features such as porosity influence cleaning outcomes. Optical microscopy-based analyses revealed heterogeneous pore size distributions, while colorimetric measurements indicated substantial reflectance variability, highlighting limitations of area-averaged evaluation methods for textured surfaces.

In the fourth part, an uncertainty-integrated modeling approach based on Monte Carlo simulations has been implemented to propagate experimental noise and parameter variability in cleaning performance predictions. This enables confidence-based evaluation of cleaning effectiveness under extrapolated conditions.

Overall, this thesis establishes a quantitative, microstructure-informed, and uncertainty-aware framework for NIR fs-laser cleaning evaluation. It advances heritage conservation science by linking morphological features to cleaning dynamics and proposing predictive contamination removal protocols with minimal substrate damage risk.

Keywords: Femtosecond Laser Cleaning; Cultural Heritage Conservation; NIR Absorption; ATR-FTIR Spectroscopy; Exponential Saturation and Decay Modelling; Historical Paper and Leather; Surface Morphology and Porosity; Selective Contaminant Removal; Uncertainty Quantification; Monte Carlo Simulation.

Résumé

Le nettoyage au laser femtoseconde (fs) apparaît comme une méthode prometteuse pour la conservation des matériaux du patrimoine culturel, mais des défis subsistent quant à la quantification des performances de nettoyage, en particulier sur des substrats poreux et morphologiquement hétérogènes.

Dans cette thèse de doctorat, menée à l'université La Sapienza de Rome et à l'Institut national de la recherche scientifique (INRS) de Montréal, un cadre d'évaluation quantitatif et fondé sur la spectroscopie a été développé pour le nettoyage au laser fs dans le proche infrarouge (NIR).

La première partie présente les aspects fondamentaux des interactions laser-matériau, en mettant particulièrement l'accent sur les mécanismes qui sous-tendent le nettoyage sélectif, ainsi qu'une revue des techniques de diagnostic telles que la spectroscopie ATR-FTIR, la colorimétrie, la microscopie électronique à balayage et la profilométrie appliquées à la conservation du patrimoine.

La deuxième partie examine de manière quantitative le nettoyage au laser NIR fs sur des échantillons de maquette à base de cellulose reproduisant des papiers historiques, qui ont été vieillis artificiellement et contaminés avec du graphite et de la kaolinite comme polluants modèles. Les performances de nettoyage ont été évaluées par spectroscopie ATR-FTIR couplée à des modèles de saturation et de décroissance exponentiels, démontrant l'élimination efficace des contaminants de surface dans une fenêtre de fluence sûre définie sans compromettre l'intégrité du substrat.

La troisième partie explore le nettoyage au laser fs de cuirs factices préparés historiquement, en soulignant que les caractéristiques microstructurales telles que la porosité influencent les résultats du nettoyage. Les analyses basées sur la microscopie optique ont révélé des distributions hétérogènes de la taille des pores, tandis que les mesures colorimétriques ont indiqué une variabilité substantielle de la réflectance, soulignant les limites des méthodes d'évaluation moyennées sur une zone pour les surfaces texturées.

Dans la quatrième partie, une approche de modélisation intégrant l'incertitude basée sur des simulations Monte Carlo a été mise en œuvre pour propager le bruit expérimental et la variabilité des paramètres dans les prévisions de performance de nettoyage. Cela permet une évaluation basée sur la confiance de l'efficacité du nettoyage dans des conditions extrapolées.

Dans l’ensemble, cette thèse établit un cadre quantitatif, fondé sur la microstructure et tenant compte de l’incertitude, pour l’évaluation du nettoyage au laser NIR fs. Elle fait progresser la science de la conservation du patrimoine en établissant un lien entre les caractéristiques morphologiques et la dynamique du nettoyage et en proposant des protocoles prédictifs d’élimination de la contamination avec un risque minimal de dommages au substrat.

Mots-clés: Nettoyage au laser femtoseconde; Conservation du patrimoine culturel; Absorption NIR; Spectroscopie ATR-FTIR; Saturation exponentielle et modélisation de la décroissance; Papier et cuir historiques; Morphologie et porosité de la surface; Élimination sélective des contaminants; Quantification de l’incertitude; Simulation de Monte Carlo.

Synthèse de la thèse en français

La conservation des objets du patrimoine culturel exige un équilibre délicat entre un nettoyage efficace et la préservation des matériaux d'origine. Parmi le large éventail de substrats du patrimoine, les matériaux organiques tels que le papier historique et le cuir posent des problèmes particuliers. Ces matériaux sont intrinsèquement fragiles et chimiquement sensibles, et présentent souvent des traitements de surface en couches et des microstructures complexes. Au fil du temps, ils accumulent diverses formes de contamination, de suie, de poussière, de résidus biologiques et de polluants liés à la manipulation qui obscurcissent les caractéristiques de la surface, entravent la lisibilité et accélèrent la dégradation. Le nettoyage de ces objets est essentiel à leur préservation, mais les méthodes conventionnelles sont souvent imprécises, invasives ou coûteuses.

Les méthodes traditionnelles de nettoyage mécanique ou chimique, y compris les brosses sèches, les scalpels, les tampons de solvant ou les systèmes de gel, présentent des limites importantes. Les méthodes mécaniques présentent un risque d'abrasion ou de rupture des fibres, tandis que les approches chimiques peuvent pénétrer dans les structures poreuses, entraînant un gonflement, une décoloration ou une altération chimique à long terme. Ces techniques sont également très dépendantes de l'opérateur, avec des résultats qui peuvent varier d'un restaurateur à l'autre, d'une institution à l'autre, voire même d'applications successives sur le même objet. Ce manque de contrôle et de reproductibilité crée un besoin pressant de méthodes précises et quantifiables.

Le nettoyage au laser offre une alternative intéressante, en particulier lorsqu'on utilise des lasers pulsés dont la fluence, la longueur d'onde et la durée d'impulsion sont contrôlées. Les systèmes laser femtoseconde sont particulièrement adaptés aux substrats délicats en raison de leur durée d'impulsion ultra-courte, qui minimise la diffusion de la chaleur et permet l'enlèvement de matière par des mécanismes non thermiques. Lorsque les paramètres sont soigneusement sélectionnés, les contaminants de surface peuvent être éliminés de manière sélective, laissant le substrat pratiquement intact. Toutefois, ces avantages ne sont pas garantis. L'état de l'art utilise des lasers à impulsions UV, dont l'absorption omniprésente d'une bande à l'autre réduit la sélectivité, mais dont la forte absorption garantit une faible profondeur de pénétration de l'ordre de quelques dizaines de nanomètres, de sorte que les lasers UV sont souvent considérés comme un scalpel optique. Malheureusement, dans les matériaux à base de cellulose, le rayonnement UV provoque des effets parasites tels que le jaunissement, ce qui limite l'utilisation de ces lasers à des fins de nettoyage. Ceci motive l'utilisation de lasers femtosecondes NIR dans cette thèse, qui éviteront les taches photo-induites tant que les fluences resteront en dessous d'un certain seuil. Par conséquent, il est essentiel de comprendre le comportement du seuil des interactions laser-matériau pour garantir l'efficacité du nettoyage et la sécurité des matériaux. Parallèlement, la profondeur de pénétration des lasers NIR est nettement supérieure à celle des lasers UV, et leur absorption est beaucoup plus sélective en fonction du matériau, de sorte que des études détaillées sur la dynamique et la sélectivité du nettoyage doivent être menées pour garantir l'utilisation de ces lasers dans le cadre de la conservation du patrimoine culturel.

Le défi est double : Le laser doit être utilisé en toute sécurité et le résultat du nettoyage doit également être évalué quantitativement. Dans le cas de substrats fragiles, qui excluent par nature un échantillonnage destructif et présentent des topographies de surface très irrégulières, l'inspection visuelle ou la colorimétrie conventionnelle peuvent s'avérer inadéquates. L'incertitude des mesures, la saturation du signal et la réflectance non uniforme peuvent masquer les progrès réels du nettoyage. Il est donc essentiel de disposer de techniques de diagnostic offrant une spécificité chimique et une sensibilité élevée.

Cette thèse répond à ces défis en intégrant l'analyse spectroscopique, l'évaluation morphologique et la prédiction basée sur un modèle pour évaluer la performance du nettoyage au laser femtoseconde sur les matériaux à base de cellulose et de protéines. L'approche va au-delà de l'évaluation qualitative et introduit des cadres quantitatifs pour décrire, interpréter et valider les résultats du nettoyage. En particulier, ce travail démontre comment les modèles exponentiels de saturation et de décroissance ajustés aux données ATR-FTIR peuvent servir de descripteurs fiables de l'élimination des contaminants et de la récupération du substrat. Il explore également l'influence de la porosité de la surface sur l'efficacité du nettoyage et étudie comment la modélisation de l'incertitude par des simulations de Monte Carlo peut améliorer la fiabilité des prédictions.

La recherche a été menée dans le cadre d'un programme doctoral en cotutelle entre l'Université Sapienza de Rome et l'Institut national de la recherche scientifique (INRS), au Canada. Des campagnes expérimentales, comprenant la spectroscopie ATR-FTIR, la profilométrie et le MEB, ont été menées dans les laboratoires partenaires à Rome, Trois-Rivières et Montréal. Les résultats contribuent au développement de protocoles plus contrôlés, reproductibles et interprétables pour le nettoyage au laser, spécifiquement conçus pour répondre aux exigences nuancées de la conservation du patrimoine culturel.

Le chapitre 1 de cette thèse établit les bases théoriques et méthodologiques nécessaires pour comprendre et évaluer le nettoyage par laser femtoseconde de substrats factices reproduisant des matériaux historiques. Il est divisé en trois parties principales : la physique des interactions laser ultra-court–matériau, les techniques de diagnostic utilisées pour analyser les résultats du nettoyage, et les outils statistiques appliqués pour interpréter et modéliser les résultats expérimentaux.

Les lasers femtoseconde émettent des impulsions de l'ordre de 10^{-15} secondes, ce qui permet de délivrer de l'énergie à un matériau plus rapidement que la diffusion thermique ne peut se produire. Cette précision temporelle permet des mécanismes d'ablation qui sont en grande partie non thermiques, ce qui minimise les dommages collatéraux dans les régions adjacentes. Pour les matériaux du patrimoine culturel, il s'agit d'un avantage essentiel. Cependant, l'utilisation de lasers femtoseconde présente également des risques potentiels, en particulier à des fluences élevées. Lorsque la densité d'énergie locale dépasse des seuils spécifiques, des effets non linéaires tels que l'absorption multiphotonique peuvent se produire, entraînant la rupture des liaisons sous la surface, le jaunissement ou la dégradation de composés organiques délicats. Il est donc essentiel de comprendre la fenêtre opérationnelle sûre de la fluence laser. Il s'agit à la fois du seuil inférieur, en dessous duquel les contaminants ne sont pas affectés, et du seuil supérieur, au-delà duquel des dommages irréversibles peuvent se produire.

Une approche diagnostique multi-méthodes a été adoptée pour évaluer les effets du nettoyage au laser. Chaque technique a fourni des informations complémentaires sur différents aspects du processus de nettoyage : la spectroscopie ATR-FTIR a été le principal outil d'évaluation des changements chimiques. En surveillant l'évolution des bandes vibratoires associées aux contaminants ou aux substrats, cette technique a fourni des preuves directes, au niveau moléculaire, de l'élimination ou de l'altération. Les zones de bandes intégrées ont été utilisées comme indicateurs quantitatifs de l'intensité du signal et ont été intégrées dans des cadres de modélisation non linéaires. La profilométrie au stylet a été utilisée avant le nettoyage au laser pour étudier comment différents types de contaminants de surface affectaient la topographie initiale du substrat, plutôt que pour quantifier les changements de surface après le nettoyage. Les paramètres de rugosité de surface ont été extraits pour évaluer si l'ablation s'est produite de manière uniforme et si des dommages microstructurels ont été introduits. La microscopie électronique à balayage (MEB) a permis de visualiser à haute résolution la morphologie de la surface, l'intégrité des fibres et les particules résiduelles. Elle a été particulièrement utile pour confirmer l'absence de perturbation mécanique dans les fibres de cellulose et pour visualiser l'adhésion des particules dans le cuir poreux. La colorimétrie offre une mesure plus traditionnelle et non invasive des résultats du nettoyage. Les différences de couleur (ΔE) ont été utilisées pour évaluer les changements visuels. Cependant, ses limites, telles que la sensibilité à l'angle, à la rugosité de la surface et à la distribution des chromophores, ont été reconnues et évaluées de manière critique, en particulier au chapitre 3.

Chacune de ces techniques a contribué à une compréhension globale de la manière dont les paramètres du laser ont influencé à la fois l'élimination des contaminants et la préservation du substrat.

Pour aller au-delà de l'interprétation qualitative, la thèse a incorporé des techniques statistiques et de modélisation pour quantifier la dynamique du nettoyage. Deux cadres principaux ont été utilisés :

- **Modélisation non linéaire** : Des équations de saturation exponentielle et de décroissance exponentielle ont été ajustées aux données spectrales. Le modèle de saturation décrit la récupération des signaux du substrat (par exemple, lorsque les contaminants sont éliminés), tandis que le modèle de décroissance décrit la perte des signaux de contamination. Ces modèles ont permis de définir des paramètres clés tels que la fluence de saturation, l'efficacité du nettoyage et le niveau de contamination initial.

- **Analyse de la variance (ANOVA)** : Elle a été appliquée pour évaluer si les différences de performance de nettoyage entre les types d'échantillons (par exemple, composition granulométrique, conditions de brunissage) étaient statistiquement significatives. Ceci était particulièrement important pour identifier l'impact de la préparation de l'échantillon sur les résultats du nettoyage.
- **Simulations de Monte Carlo** : Introduites au chapitre 4, elles ont été motivées par la nécessité de tenir compte de l'incertitude des mesures et de la sensibilité du modèle. Les concepts préliminaires ont été décrits au chapitre 1 afin de jeter les bases de leur mise en œuvre ultérieure.

Ensemble, ces outils théoriques, diagnostiques et statistiques ont constitué l'ossature analytique de la thèse. Ils ont permis le développement de modèles prédictifs, la comparaison de stratégies de traitement et la quantification de l'incertitude – autant d'éléments essentiels pour faire progresser la base scientifique du nettoyage au laser dans les contextes du patrimoine culturel.

Le chapitre 2 présente un cadre quantitatif pour l'évaluation des performances de nettoyage des lasers femtoseconde sur des substrats de papier historiquement représentatifs. L'approche combine la spectroscopie ATR-FTIR avec un ajustement de courbe non linéaire et une analyse statistique pour extraire les paramètres clés qui décrivent l'efficacité de l'élimination des contaminants et la préservation du substrat.

Trois types de matériaux d'encollage historiquement pertinents ont été sélectionnés : l'amidon de maïs, l'amidon de blé et l'œuf d'alun, ainsi que des papiers de contrôle non encollés. Chaque groupe a été artificiellement souillé à l'aide d'une poudre sèche à base de graphite ou d'un mélange de kaolinite et de graphite, choisis pour reproduire les contaminants à base de carbone et les contaminants minéraux courants que l'on trouve sur les artefacts papier historiques. Ces contaminants ont également été sélectionnés en raison de leurs caractéristiques d'absorption différentes dans la région du proche infrarouge (NIR), qui influencent directement leur interaction avec le rayonnement laser femtoseconde à 1030 nm et, par conséquent, l'efficacité du nettoyage.

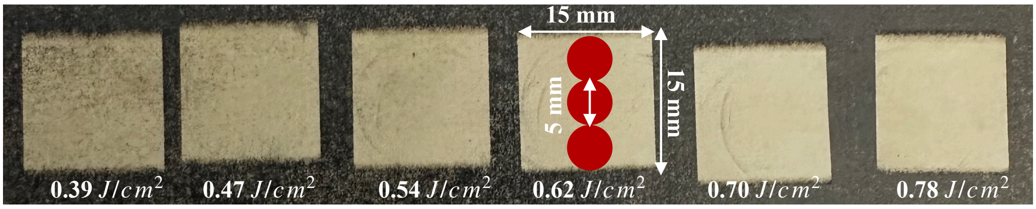


Figure 0.1. Régions nettoyées de papier artificiellement souillé et vieilli (encollage alum-œuf, contamination graphite+kaolinite). Les zones carrées ont été nettoyées avec un laser fs à 1030 nm. Les points rouges représentent les points de collecte des données pour les mesures ATR. Le succès progressif du nettoyage avec l'augmentation de la fluence du laser est visible, mais il est difficilement quantifiable.

Pour les échantillons non encollés, deux sous-groupes ont été préparés : l'un ayant subi un brunissage après salissure et l'autre non. Dans ces cas, le brunissage a été traité comme une variable optionnelle, utilisée pour étudier son influence sur le comportement des contaminants et la réponse au nettoyage. Comme ces papiers n'ont pas de couche d'encollage, le brunissage n'était pas une étape obligatoire de leur préparation. En revanche, les échantillons de tous formats ont été systématiquement brunis après l'application de l'encollage et avant la salissure, dans le cadre d'un protocole de finition historiquement exact. Cette conception a permis à l'étude de séparer les effets du compactage de la surface (brunissage) de ceux de l'encollage chimique, ce qui a aidé à clarifier la manière dont chaque facteur contribue aux interactions laser–contaminants.

Tous les échantillons ont été nettoyés à l'aide d'un système laser femtoseconde à 1030 nm dans le proche infrarouge, dans des conditions ambiantes. La fluence du laser a varié systématiquement entre 0.39 et 0.78 J/cm^2. Des spectres ATR-FTIR ont été recueillis après chaque étape de nettoyage afin de surveiller les changements chimiques, notamment en suivant l'évolution des bandes associées à la couche de contamination ou au matériau d'encollage sous-jacent. Comme le montre la figure 0.1, l'évolution spectrale à chaque niveau de fluence révèle l'élimination sélective des contaminants de surface et l'exposition progressive de la couche d'encollage sous-jacente.

Pour interpréter les changements spectraux ATR-FTIR, deux modèles non linéaires distincts ont été utilisés : le modèle de saturation exponentielle a été utilisé pour décrire la récupération des signaux du substrat au fur et à mesure de l'élimination des contaminants sous-jacents. Ce modèle rend compte de la manière dont le substrat devient de plus en plus visible avec l'augmentation de la fluence, mais finit par atteindre la saturation. Le modèle de décroissance exponentielle a été utilisé pour s'adapter aux signaux directement associés à la couche de contaminants, qui diminue avec l'augmentation de l'exposition au laser.

Ces modèles contiennent chacun quatre paramètres physiquement interprétables. Pour le modèle de saturation, les paramètres clés comprennent le niveau de signal maximal (représentant un nettoyage complet), le seuil de fluence à partir duquel le nettoyage commence, le taux de saturation (c'est-à-dire l'efficacité du nettoyage) et le signal de base avant le nettoyage. Pour le modèle de décroissance, les paramètres décrivent le niveau de contamination initial, le taux de décroissance du signal et le signal résiduel final.

En adaptant ces modèles aux données spectrales de chaque groupe d'échantillons, l'étude a défini un ensemble de mesures de performance qui vont au-delà de l'inspection visuelle :

- **Fluence de saturation** : Le niveau de fluence au-delà duquel de nouvelles augmentations de l'énergie du laser produisent une amélioration supplémentaire négligeable du signal.
- **Pourcentage de récupération du signal** : La mesure dans laquelle le signal après nettoyage se rapproche du maximum modélisé.
- **Seuil de fluence** : L'énergie laser minimale requise pour déclencher une réaction de nettoyage détectable.

Pour compléter la modélisation spectroscopique, des analyses par MEB et profilométrie au stylet ont été effectuées sur des échantillons sélectionnés avant le nettoyage au laser. L'imagerie MEB a été utilisée pour évaluer la morphologie de la surface et la distribution des contaminants avant toute exposition au laser, ce qui a permis d'identifier les schémas de dépôt des particules et les interactions au niveau des fibres. De même, des mesures de profilométrie ont été effectuées pour quantifier la rugosité de la surface et étudier comment les différents types de contaminants et les conditions de brunissage influencent les caractéristiques topographiques. Plutôt que d'évaluer les effets post-nettoyage, ces mesures ont servi à caractériser les conditions initiales et à élucider la relation entre la morphologie de la contamination et les interactions ultérieures entre le laser et le matériau.

L'analyse statistique utilisant l'ANOVA à deux voies a démontré que la composition de l'encollage et les conditions de brunissage affectaient de manière significative la réponse du nettoyage au laser. Il est important de noter que la mesure étudiée n'était pas l'efficacité en soi, mais la réactivité de chaque groupe d'échantillons, quantifiée par la pente des changements dépendant de la fluence dans les signaux ATR-FTIR normalisés. Les échantillons contaminés par le graphite présentaient systématiquement une réactivité plus élevée que ceux contaminés par le mélange graphite-kaolinite, ce qui reflète la plus forte absorption du graphite dans le proche infrarouge et son interaction plus favorable avec les impulsions femtosecondes. Parmi les cas de graphite, les échantillons non calibrés ont produit les courbes de réponse les plus raides, probablement en raison de l'absence d'une couche de calibrage rétentive, qui laisse les contaminants de surface plus exposés à l'interaction du laser.

L'influence du brunissage varie en fonction de l'encollage. Dans le cas des papiers non collés, le brunissage a réduit la réactivité, probablement en raison de la compaction de la surface qui limite l'accessibilité des contaminants. En revanche, les papiers à base d'amidon sont restés modérément réactifs même après le brunissage, ce qui est cohérent avec les observations du MEB et de la profilométrie qui montrent des films de surface plus lisses et plus uniformes. Ces résultats soulignent que non seulement la nature chimique de l'encollage, mais aussi l'état mécanique de la surface déterminent conjointement l'efficacité de l'interaction laser-contaminant.

Il est à noter que toutes les mesures morphologiques, le MEB et la profilométrie ont été effectués avant l'exposition au laser. Leur but n'était pas d'évaluer directement les résultats du nettoyage, mais de contextualiser la façon dont les caractéristiques de la surface (par exemple, la porosité, la douceur) modulent le couplage du laser et l'accessibilité des contaminants. Avec les résultats de l'ANOVA, ces informations permettent de mieux comprendre les mécanismes physiques qui régissent la sélectivité du nettoyage au laser fs sur le papier historique.

Ces résultats soulignent l'importance de tenir compte des pratiques de fabrication historiques lors de la conception de stratégies de conservation. Le cadre de modélisation non linéaire développé ici fournit une méthode reproductible et quantitative pour comparer différentes conditions de nettoyage, ouvrant la voie à une prise de décision plus éclairée en matière de conservation du papier.

Le chapitre 3 explore l'application du nettoyage au laser femtoseconde à des substrats historiques en cuir, qui présentent des défis sensiblement différents de ceux des matériaux à base de cellulose en raison de leur morphologie de surface complexe, de leur porosité variable et de leur composition chimique hétérogène. Ce chapitre évalue de manière critique l'efficacité du nettoyage et les limites des outils de diagnostic couramment utilisés, en particulier la colorimétrie, pour saisir ces résultats.

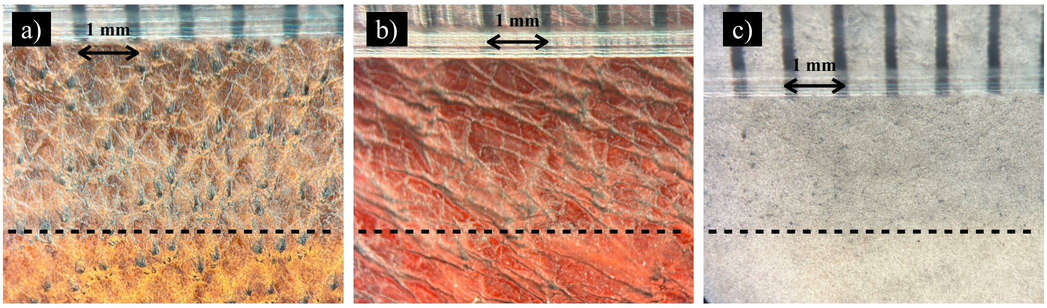


Figure 0.2. Images de surface des trois types d'échantillons après contamination et vieillissement, acquises à l'aide d'un stéréomicroscope Olympus à un grossissement de 2×. Les panneaux a), b) et c) correspondent respectivement au cuir de couleur havane, au cuir de couleur orange et au parchemin. Sur chaque image, la partie supérieure au-dessus de la ligne noire en pointillés représente la zone contaminée par le graphite, tandis que la partie inférieure correspond à la zone nettoyée au laser. La barre d'échelle représente 1 mm. Les couleurs perçues dans les images peuvent différer légèrement des échantillons réels en raison de l'éclairage et des paramètres d'imagerie.

Des échantillons de cuir de chèvre tanné et vieilli ont été sélectionnés pour leur pertinence historique et la diversité de leurs microstructures. Tous les échantillons présentaient une porosité variable due à la variabilité naturelle du matériau. Les expériences de nettoyage ont été réalisées à l'aide d'un système laser femtoseconde à 1030 nm dans le proche infrarouge au Laboratoire de photonique IR et THz de l'INRS, Canada. Les valeurs de fluence du laser ont à nouveau été modifiées dans une plage de sécurité contrôlée afin d'éviter d'endommager le substrat. Comme le montre la figure 0.2, les images microscopiques acquises après le nettoyage au laser mettent en évidence le contraste entre les régions contaminées par le graphite et les zones nettoyées au laser dans différents types de cuir.

Avant le nettoyage au laser, la microscopie optique et la quantification de la porosité basée sur ImageJ ont été utilisées pour analyser la distribution et la morphologie des pores sur chaque échantillon de cuir. Les résultats ont révélé des distributions hétérogènes de pores de petite et de grande taille sur toute la surface, entraînant des différences localisées dans l'absorption de la lumière et le comportement d'ablation.

Cette porosité a eu un impact direct sur l'interaction entre les impulsions laser et la surface du matériau. Les régions présentant une porosité plus dense retiennent plus fortement les contaminants, tandis que les zones plus lisses permettent une ablation plus uniforme. Ces résultats soulignent la nécessité d'évaluer les effets du nettoyage à l'aide de techniques sensibles aux variations spatiales.

Le succès du nettoyage a été évalué à l'aide de la colorimétrie de réflectance basée sur le (ΔE), une méthode couramment utilisée pour documenter les changements visuels. Les mesures ont été effectuées au département des sciences médico-légales de l'Université du Québec à Trois-Rivières. Malgré son utilisation répandue, la colorimétrie a révélé une variabilité significative des valeurs (ΔE), même entre des régions adjacentes d'un même échantillon.

L'utilisation d'un colorimètre à haute résolution a permis de remédier à plusieurs limitations de la colorimétrie conventionnelle. Contrairement aux méthodes d'imagerie à large surface, ce système a permis des mesures précises de la réflectance sur de petites régions localisées de la surface, souvent plus petites que le diamètre du faisceau laser, ce qui a permis une évaluation plus détaillée des changements au niveau de la surface. Cet avantage en termes de résolution spatiale a permis une évaluation plus précise de l'élimination des contaminants localisés et de l'uniformité de la surface. Cependant, il est important de noter que la mesure (ΔE) couramment utilisée, bien que numérique, ne varie pas linéairement avec les niveaux de contamination et ne peut donc pas être utilisée de manière fiable pour quantifier les performances de nettoyage. Des facteurs externes tels que l'angle d'illumination, l'orientation de l'échantillon et l'alignement du détecteur nécessitent toujours un contrôle strict pour garantir des résultats cohérents et interprétables.

Dans certains cas, les zones nettoyées présentaient des valeurs (ΔE) plus élevées que les témoins non nettoyés, bien que l'inspection visuelle et l'imagerie microscopique confirment l'élimination effective des contaminants. Cette contradiction apparente souligne une limite fondamentale de l'évaluation colorimétrique sur des surfaces non uniformes telles que le cuir. En raison de la microtopographie hétérogène du cuir tanné, qui comprend des pics, des vallées et des groupes de pores, l'interaction spatiale entre le faisceau laser et la surface varie considérablement d'une région à l'autre. Lorsque les mesures de couleur sont effectuées dans des zones élevées (pics de surface) où l'interaction laser-matière est plus directe et plus efficace, les spectres de réflectance reflètent avec précision l'état de nettoyage, ce qui se traduit par les réductions (ΔE) attendues. Cependant, lorsque la réflectance est mesurée à partir de zones en retrait ou irrégulières adjacentes à ces pics, telles que les vallées ou les bords des pores, les relevés de couleur peuvent ne pas représenter de manière fiable le véritable état de nettoyage. Ces zones peuvent encore présenter des effets de diffusion, d'ombrage ou de contamination résiduelle, conduisant à des valeurs (ΔE) élevées inattendues, même lorsque la surface a été nettoyée dans des conditions de fluence et d'impulsion identiques. De telles incohérences soulignent la nécessité de faire preuve de prudence lors de l'interprétation des données colorimétriques sur des substrats structurellement complexes et soutiennent l'utilisation de techniques analytiques complémentaires pour une évaluation robuste du nettoyage.

Ces résultats révèlent que les mesures de réflectance moyenne par zone peuvent masquer les différences locales de nettoyage, en particulier sur les substrats fibreux et poreux où l'interaction laser-matériau dépend fortement de la morphologie de la surface. Dans le cuir, les ombres localisées, les angles d'incidence variables et la contribution optique des bords ou des vallées des pores peuvent tous déformer les profils de réflectance, ce qui conduit à des valeurs (ΔE) inexactes ou exagérées. Par conséquent, la colorimétrie, bien que largement utilisée pour la documentation, n'a pas la spécificité spatiale et la pertinence physique requises pour évaluer quantitativement le nettoyage au laser femtoseconde sur des surfaces aussi complexes.

Pour surmonter ces limitations, les évaluations futures devraient intégrer des diagnostics chimiquement sélectifs et spatialement résolus. En outre, une meilleure compréhension de la manière dont les interactions optiques et thermiques varient sur les surfaces hétérogènes en cuir sera essentielle pour optimiser les paramètres du laser, en particulier la taille du spot et la fluence, afin d'équilibrer la sélectivité et la couverture de la surface. De telles stratégies multimodales sont vitales pour développer des protocoles robustes et fondés sur des preuves pour le nettoyage au laser dans la conservation du patrimoine culturel.

Le chapitre 4 s'appuie sur le cadre de modélisation spectroscopique présenté précédemment en intégrant la quantification de l'incertitude dans l'analyse du nettoyage par laser femtoseconde. Alors que les chapitres précédents se concentraient sur les ajustements déterministes des modèles exponentiels de saturation et de décroissance aux données expérimentales, ce chapitre souligne l'importance de la prise en compte de la variabilité des mesures et de la robustesse des ajustements pour générer des prédictions statistiquement significatives.

Dans le monde réel de la conservation, les décisions basées sur les prédictions des modèles doivent être fiables, en particulier lorsqu'il s'agit d'artefacts fragiles ou irremplaçables. Comme les données spectroscopiques contiennent intrinsèquement du bruit et de la variabilité expérimentale, les estimations ponctuelles des paramètres du modèle sont insuffisantes pour saisir toute la gamme des résultats possibles. C'est pourquoi un cadre probabiliste a été développé pour évaluer l'influence de l'incertitude des paramètres d'entrée sur les trajectoires de nettoyage prédites dans la gamme de fluence laser.

Avant d'effectuer des simulations d'incertitude, le modèle de saturation exponentielle a été réexaminé en mettant l'accent sur l'interprétation physique de ses paramètres. Les quatre paramètres d'ajustement a, b, c et d ont été examinés en termes de comportement mathématique et de pertinence physique :

- $\boldsymbol{a}$: représente le signal asymptotique maximal (niveau de saturation).
- $\boldsymbol{b}$: marque le seuil de fluence pour un nettoyage détectable.
- $\boldsymbol{c}$: contrôle la pente ou la sensibilité de la réponse à une fluence croissante.
- $\boldsymbol{d}$: tient compte du décalage du signal initial avant le nettoyage.

Des limites ont été imposées à ces paramètres pour garantir des ajustements physiquement significatifs. Par exemple, b a été limité entre 0 et la fluence la plus faible mesurée (0,39 J/cm^2), et c a été limité à environ 20% de la plage de fluence totale afin d'éviter des courbes plates physiquement peu plausibles. Ceci était particulièrement important car des valeurs élevées de c conduisaient à des extrapolations irréalistes et compromettaient la capacité du modèle à prédire avec précision les seuils de nettoyage.

Pour propager l'incertitude dans le modèle, la variabilité des paramètres a été introduite en perturbant les données expérimentales à l'aide d'un bruit gaussien de moyenne zéro, mis à l'échelle par l'écart type des mesures ATR-FTIR répétées. Pour chacune des 100 itérations de Monte Carlo, un nouvel ensemble de données synthétiques a été créé et réajusté à l'aide du modèle exponentiel de saturation. Ce processus a abouti à un ensemble de courbes ajustées, capturant à la fois la sensibilité des paramètres et le bruit expérimental.

Les intervalles de confiance ont été calculés ponctuellement à partir de cet ensemble, ce qui a permis d'obtenir une bande de confiance qui quantifie la fiabilité de la prédiction à chaque valeur de fluence.

Les simulations ont révélé que l'incertitude était la plus élevée dans la région de transition, où le signal évolue rapidement, et la plus faible dans les régimes asymptotiques, où le signal tend à se stabiliser. La forme de la bande de confiance, comme illustré dans la Figure 0.3, reflète clairement ce comportement physique:

- Des bandes étroites à faible fluence indiquent une réponse initiale stable et reproductible avant le début effectif du nettoyage.
- L'élargissement des bandes dans les plages de fluence intermédiaires met en évidence une forte sensibilité aux variations des paramètres du modèle.
- Enfin, l'aplatissement des régions supérieures suggère une convergence vers un comportement de saturation cohérent, indépendamment des perturbations aléatoires.

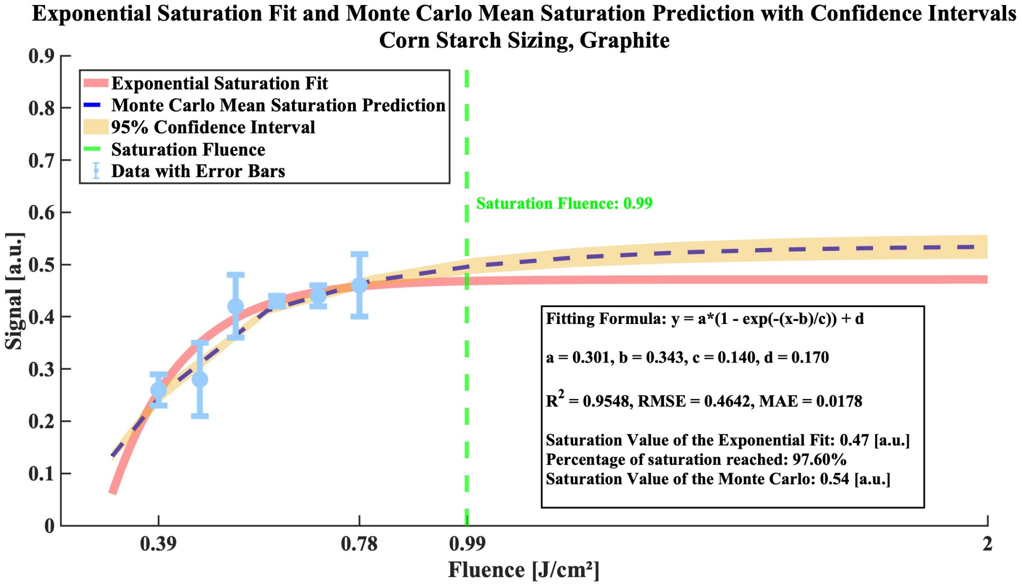


Figure 0.3. Ajustement de la saturation exponentielle et prédiction moyenne de Monte Carlo avec un paramètre c borné (Corn Starch Sizing, Graphite). L'ensemble Monte Carlo a utilisé une limite supérieure de $c = 0,4$, correspondant à 20% de la gamme de fluence totale (2 J/cm^2).

Il est important de noter que ce cadre a permis d'estimer les fluences seuils (minimum requis pour un nettoyage détectable) et les limites supérieures (au-delà desquelles une fluence supplémentaire n'apporte aucun avantage supplémentaire), toutes deux assorties d'intervalles de confiance. Ces résultats constituent des outils pratiques pour les restaurateurs qui souhaitent opérer dans des fenêtres de nettoyage sûres et efficaces.

Cette thèse présente un cadre intégré pour l'évaluation du nettoyage au laser femtoseconde de matériaux historiques, combinant l'analyse spectroscopique, la caractérisation morphologique, la modélisation non linéaire et la quantification de l'incertitude. En se concentrant sur les substrats en papier et en cuir, deux matériaux complexes et très sensibles fréquemment rencontrés dans la conservation du patrimoine, la recherche aborde les défis critiques de l'évaluation des résultats du nettoyage avec précision, reproductibilité et rigueur scientifique.

En utilisant la spectroscopie ATR-FTIR, des modèles exponentiels ont été développés pour quantifier l'élimination des contaminants et la récupération du substrat en fonction de la fluence du laser. En ajustant les courbes de saturation et de décroissance aux données spectroscopiques spécifiques à la bande, l'étude a démontré que l'efficacité et la sécurité du nettoyage peuvent être évaluées au-delà des critères subjectifs ou purement visuels. L'application de l'ANOVA à deux voies a en outre permis une comparaison statistique entre les différents traitements de surface, offrant un aperçu de la manière dont le calibrage et le brunissage influencent la rétention des contaminants et l'interaction avec le laser.

Les limites des méthodes d'évaluation traditionnelles, en particulier la colorimétrie, ont été examinées dans le contexte du nettoyage du cuir. Il a été démontré que la porosité, la rugosité de la surface et les effets de diffusion introduisent une variabilité significative dans les lectures colorimétriques, ce qui sape leur fiabilité pour les substrats complexes. En revanche, les signaux spectroscopiques fournissent des données plus cohérentes et chimiquement plus spécifiques pour suivre l'évolution du nettoyage.

Dans la phase finale, une approche de modélisation de l'incertitude basée sur Monte Carlo a été mise en œuvre pour simuler l'effet du bruit expérimental et de la variabilité des paramètres sur les prédictions du modèle. Les intervalles de confiance générés à partir d'ensembles de données synthétiques ont illustré les régions où la sensibilité des prédictions est la plus élevée et ont permis d'identifier des fenêtres de fluence opérationnelle statistiquement robustes.

En associant les approches déterministes et probabilistes, cette thèse propose une méthodologie scientifiquement fondée pour évaluer et optimiser le nettoyage au laser femtoseconde. Elle fournit non seulement un aperçu empirique de la façon dont différents matériaux réagissent à l'exposition au laser, mais aussi une stratégie de modélisation transférable adaptable à d'autres substrats, contaminants ou environnements de conservation.

À l'avenir, la recherche devrait étendre ce cadre pour inclure d'autres types de contamination, des surfaces multicouches et des formes de modèles plus avancées (par exemple, inférence bayésienne, modèles hybrides mécanistes-statistiques). Il existe également un fort potentiel d'intégration de modalités diagnostiques complémentaires telles que l'OCT, la spectroscopie Raman ou l'imagerie hyperspectrale pour améliorer encore l'évaluation du nettoyage en trois dimensions.

En fin de compte, ce travail contribue au développement d'une nouvelle norme dans la science de la conservation du patrimoine : une norme dans laquelle les procédures de nettoyage sont non seulement peu invasives et réversibles, mais aussi quantifiables, reproductibles et adaptées grâce à une vision diagnostique solide.

Contents

Synthèse de la thèse en français

List of Figures

List of Tables

List of Abbreviations

Introduction **1**

1 Fundamental Principles of Laser Cleaning, Diagnostic Techniques and Statistical Methods in Heritage Science **7**

1.1 Laser–Material Interaction Principles 7

1.2 Diagnostic Techniques for Surface and Material Analysis 10

1.2.1 ATR-FTIR Spectroscopy 10

1.2.2 Scanning Electron Microscopy 12

1.2.3 Profilometry 13

1.2.4 Colorimetry 14

1.3 Statistical Tools for Data Interpretation: ANOVA and Monte Carlo Simulations 16

2 Quantifying Selective Pulsed Laser Cleaning and Residual Contamination in Paper Artifacts by ATR Spectroscopy **19**

2.1 Introduction 20

2.2 Materials and methods . . . 23
2.2.1 Paper samples . . . 23
2.2.2 Femtosecond (fs) laser cleaning setup . . . 26
2.2.3 Scanning Electron Microscopy . . . 28
2.2.4 Stylus Profilometry . . . 28
2.2.5 Attenuated Total Reflectance (ATR) Spectroscopy . . . 28
2.2.6 Statistical Analysis by Analysis of Variance (ANOVA) . . . 28
2.2.7 Data Collection and Processing . . . 29
2.3 Results and Discussion . . . 35
2.3.1 Surface Morphology by Scanning Microscope . . . 35
2.3.2 Surface Roughness . . . 38
2.3.3 ANOVA-Based Assessment of Sizing, Contamination, and Burnishing Effects . . . 39
2.3.4 Quantifying Cleaning with Saturation Modeling . . . 44
2.4 Conclusion . . . 53

3 Challenges in Colorimetric Evaluation of Femtosecond Laser Cleaning on Historical Leather: The Role of Surface Porosity and Microstructure **55**

3.1 Introduction . . . 56
3.2 Materials and methods . . . 57
3.2.1 Leather samples . . . 57
3.2.2 Femtosecond Laser Cleaning Setup . . . 60
3.2.3 Optical Microscopy for Surface Morphology Analysis . . . 61
3.2.4 Color Measurement . . . 62
3.3 Results and Discussion . . . 64
3.3.1 Optical Microscopy-based porosity comparison . . . 64
3.3.2 Colorimetry and Reflectance Variability . . . 67

3.4 Conclusion 74

4 Integrating Uncertainty Distribution into Pulsed Laser Cleaning Models: A Case Study on Paper Artifacts 77

4.1 Introduction 79

4.2 Modeling and Uncertainty Quantification 82

4.2.1 Noise Modeling and Data Perturbation 86

4.2.2 Ensemble Prediction and Confidence Interval Construction . 86

4.3 Results and Discussion 87

4.3.1 Model Fit and Performance Metrics 87

4.3.2 Parameter Variability and Monte Carlo Results 89

4.4 Conclusion 95

Conclusion and perspective 97

Bibliography 99

Acknowledgments 107

List of Figures

0.1 Régions nettoyées de papier artificiellement souillé et vieilli (encollage alum-œuf, contamination graphite+kaolinite). Les zones carrées ont été nettoyées avec un laser fs à 1030 nm. Les points rouges représentent les points de collecte des données pour les mesures ATR. Le succès progressif du nettoyage avec l'augmentation de la fluence du laser est visible, mais il est difficilement quantifiable. . . . .

0.2 Images de surface des trois types d'échantillons après contamination et vieillissement, acquises à l'aide d'un stéréomicroscope Olympus à un grossissement de 2×. Les panneaux a), b) et c) correspondent respectivement au cuir de couleur havane, au cuir de couleur orange et au parchemin. Sur chaque image, la partie supérieure au-dessus de la ligne noire en pointillés représente la zone contaminée par le graphite, tandis que la partie inférieure correspond à la zone nettoyée au laser. La barre d'échelle représente 1 mm. Les couleurs perçues dans les images peuvent différer légèrement des échantillons réels en raison de l'éclairage et des paramètres d'imagerie. . . . .

0.3 Ajustement de la saturation exponentielle et prédiction moyenne de Monte Carlo avec un paramètre c borné (Corn Starch Sizing, Graphite). L'ensemble Monte Carlo a utilisé une limite supérieure de $c = 0,4$, correspondant à 20% de la gamme de fluence totale (2 J/cm^2).

1.1 Schematic processing comparison of long pulse duration (left), short pulse duration(right) [7]. 8

1.2 Comparative example for nanosecond laser pulses (left) versus femtosecond laser pulses (right) when drilling micro holes in stainless steel [7]. 9

1.3 Schematic representation of the ATR principle, showing the incoming IR beam, total internal reflection in the ATR crystal, and evanescent wave penetration into the sample. 11

1.4 Three-dimensional representation of the CIELAB color space, showing the lightness (L*), red-green (a*), and yellow-blue (b*) axes. Chroma is the radial distance from the origin, while hue is the angular position around the L* axis [32]. This perceptual model underlies quantitative color difference calculations in cultural heritage analysis. 15

2.1 Absorption coefficient spectra of a) graphite [60], and b) kaolinite [61]. 22

2.2 Scheme of sample preparation. 24

2.3 ATR-FTIR spectrum of the no-sizing, uncontaminated, and aged paper used in this study, showing no detectable bands for kaolinite. . 25

2.4 Femtosecond laser cleaning setup. 26

2.5 Cleaned regions of artificially soiled and aged paper (alum-egg sizing, graphite+kaolinite contamination). Square areas were cleaned with a fs laser at 1030 nm. Red spots represent the data collection points for ATR measurements. Progressive cleaning success with increasing laser fluence is visible; however, it is hardly quantifiable. 29

2.6 ATR-FTIR spectrum averaged from three measurements and baseline-corrected, showing the integrated absorbance peak of cellulose (shaded area) used for quantitative analysis. The 'Corrected Area' under the peak corresponds to this shaded region. The sample is alum–egg sized paper contaminated with graphite, aged, and cleaned at 0.47 J/cm^2. 31

2.7 ATR-FTIR spectrum averaged from three measurements and baseline-corrected, showing the integrated absorbance peak of kaolinite (shaded area) used for quantitative analysis. The 'Corrected Area' under the peak corresponds to this shaded region. The sample is alum–egg sized paper contaminated with graphite + kaolinite mixture, aged, and cleaned at 0.47 J/cm^2. 31

2.8 Scanning electron microscopy images of paper samples: a) no sizing, b) alum-egg sizing, c) corn starch sizing, and d) wheat starch sizing. Images were taken after aging but before contaminating and cleaning. 35

2.9 Scanning electron microscopy images of wheat starch sizing samples contaminated by a) graphite and b) graphite + kaolinite mixture. . 36

2.10 Scanning electron microscopy images of paper surfaces before contamination: a) unburnished and b) burnished sample. 37

2.11 Scanning microscope images comparing a) unburnished and b) burnished paper samples contaminated with graphite. 38

2.12 Estimated marginal means from two-way ANOVA for Subset I. Graphite and graphite + kaolinite contaminations are evaluated across three sizing types and a non-sized sample under fixed burnishing and fs-laser conditions. 41

2.13 Estimated Marginal Means (EMM) plot for Subset II, showing the interaction between burnishing and contamination types. 43

2.14 Exponential saturation model fit for alum-egg sizing with graphite contamination for fs laser cleaning. The green solid curve represents the fitted cleaning trajectory, and the vertical green dashed line marks the saturation fluence threshold ($1.12\,\mathrm{J/cm^2}$) determined from the model. Beyond this fluence value, the model extrapolates the cleaning response. However, prediction uncertainty increases as the curve extends beyond the range of measured data. 44

2.15 Exponential saturation model fit for a) wheat starch ve b) corn starch sizing with graphite contamination for fs laser cleaning. 46

2.16 Exponential saturation model fit for a) no sizing/burnishing ve b) no sizing/no burnishing with graphite contamination for fs laser cleaning. 47

2.17 Exponential decay model fit for corn starch sizing with graphite + kaolinite mixture contamination for fs laser cleaning. 48

2.18 Exponential decay model fit for a) wheat starch sizing and b) alum - egg sizing with graphite + kaolinite contamination for fs laser cleaning. 49

2.19 Exponential saturation model fit for a) no sizing/burnishing and b) no sizing/no burnishing with graphite + kaolinite contamination for fs laser cleaning. 50

3.1 Surface morphology of representative samples captured at 2× magnification using an Olympus stereo microscope. a) Orange colored goat leather, b) Havana colored goat leather, and c) parchment. A 1mm scale bar is shown in each image. The perceived colors in the images may slightly differ from the actual samples due to lighting and imaging parameters. 58

3.2 Cross-sectional optical images of leather samples dyed with a) Havana and b) Orange RODA NF dyes. Images were acquired using an Olympus stereo microscope at 2× magnification. Each scale division corresponds to 0.5 mm. The dyed layers are visible as distinct surface coatings. The perceived colors in the images may slightly differ from the actual samples due to lighting and imaging parameters. 59

3.3 Surface images of the three sample types after contamination and aging process, acquired using an Olympus stereo microscope at 2× magnification. Panels a), b), and c) correspond to Havana colored leather, Orange colored leather, and parchment, respectively. In each image, the upper region above the dashed black line shows the graphite-contaminated area, while the lower region corresponds to the laser-cleaned zone. The scale bar represents 1 mm. The perceived colors in the images may slightly differ from the actual samples due to lighting and imaging parameters. 60

3.4 Schematic of the femtosecond laser cleaning setup. 61

3.5 Optical microscopy images of Orange colored leather showing pore distribution at 10× magnification with a 100 μm scale bar. Images a) and b) were taken from different regions to assess spatial variability in surface porosity. The field of view for each image is approximately 1120 μm × 840 μm. 64

3.6 Optical microscopy images of Havana colored leather used for porosity analysis at 10× magnification with a 100 μm scale bar. Image a) shows the general surface morphology, while image b) illustrates a pore-focused view. The field of view for each image is approximately 1120 μm × 840 μm. 65

3.7 Optical microscopy images of parchment sample showing pore distribution at 10× magnification with a 100 μm scale bar. Images a) and b) were taken from different regions to assess spatial variability in surface porosity. The field of view for each image is approximately 1120 μm × 840 μm. 66

3.8 Representative pore size measurements from optical microscopy images: a) Orange colored leather, b) Havana leather. The field of view for each image is approximately 1120 μm × 840 μm. Approximate pore diameters are indicated in micrometers. 66

3.9 Grayscale-based surface plots generated from 8-bit converted optical microscopy images, showing pixel intensity distribution for a) Havana colored leather, b) Orange colored leather, and c) parchment samples. These plots visualize reflectance variation and surface texture but do not represent true topographical height measurements. 67

3.10 Color difference (ΔE^*) values measured on Orange colored leather samples contaminated with graphite after fs-laser cleaning, as a function of pulse energy for 50 and 100 laser shots. Error bars represent standard deviations from three replicate measurements. 68

3.11 Color difference (ΔE^*) values measured on Havana colored leather samples contaminated with graphite after fs-laser cleaning, as a function of pulse energy for 50 and 100 laser shots. Error bars represent standard deviations based on three replicate measurements. 70

3.12 Color difference (ΔE^*) values measured on Parchment sample contaminated with graphite after fs-laser cleaning, as a function of pulse energy for 50 and 100 laser shots. Error bars represent standard deviations from three replicate measurements. 71

4.1 Exponential saturation model fit for corn starch sizing with graphite contamination for fs laser cleaning. 87

4.2 Exponential decay fit for corn starch sizing contaminated with graphite and kaolinite. 89

4.3 Monte Carlo simulation results for the exponential saturation model applied to corn starch sized paper contaminated with graphite. The red curve represents the best-fit deterministic model, while the dashed blue curve corresponds to the ensemble mean prediction across 100 Monte Carlo realizations. The shaded yellow region denotes the 95% confidence interval (percentile-based), capturing both experimental noise and parameter uncertainty. Experimental data are shown as circular markers with error bars. The vertical green dashed line indicates the saturation fluence ($F_{\text{sat}} = 0.96\,\text{J/cm}^2$), and the horizontal red line marks the saturation value of the deterministic fit (0.47 a.u.). The widening of the confidence band at higher fluences reflects increased model uncertainty and reduced predictive stability beyond the measured data range. 90

4.4 Exponential saturation fit and Monte Carlo mean prediction with bounded c parameter (Corn Starch Sizing, Graphite). The Monte Carlo ensemble used an upper bound of $c = 0.4$, corresponding to 20% of the total fluence range (2 J/cm^2). 92

4.5 Monte Carlo simulation results for the exponential decay model applied to corn starch sizing paper contaminated with graphite. The red curve represents the best-fit deterministic model, while the dashed blue line indicates the ensemble mean from 100 simulations. The yellow band denotes the 95% confidence interval. 93

List of Tables

2.1 Chosen intervals for each paper treatment. 30

2.2 Average roughness values of the different paper treatments. 38

2.3 Two-way ANOVA results for Subset I (burnished samples). 40

2.4 Two-way ANOVA results for Subset II (non-sized samples with and without burnishing). 42

2.5 Fitted parameters for the exponential saturation model and saturation proximity. 51

2.6 Fitted parameters for the exponential decay model and saturation proximity. 52

3.1 Total color difference (ΔE^*) for fs-laser cleaned Orange colored leather sample at varying pulse energies and shot numbers. Values are reported as $\Delta E^* \pm$ propagated uncertainty ($\sigma_{\Delta E^*}$). 68

3.2 Total color difference (ΔE^*) for fs-laser cleaned Havana colored leather sample at varying pulse energies and shot numbers. Values are reported as $\Delta E^* \pm$ propagated uncertainty ($\sigma_{\Delta E^*}$). 69

3.3 Total color difference (ΔE^*) for fs-laser cleaned Parchment sample at varying pulse energies and shot numbers. Values are reported as $\Delta E^* \pm$ propagated uncertainty ($\sigma_{\Delta E^*}$). 70

3.4 CIE L*a*b* coordinates (L^*, a^*, b^*) are reported as mean ± standard deviation, while the total color difference (ΔE^*) is reported as value ± propagated uncertainty. All values correspond to 100-shot exposures. 73

List of Abbreviations

ANOVA	Analysis of Variance
ATR	Attenuated Total Reflectance
BSE	Backscattered Electrons
CIE L*a*b*	Commission Internationale de l'Éclairage L*a*b* color space
FTIR	Fourier Transform Infrared
IR	Infrared
IRE	Internal Reflection Element
MAE	Mean Absolute Error
Mid-IR	Mid-Infrared
MPA	Multi-Photon Absorption
OCT	Optical Coherence Tomography
Ra	Mean Roughness
RH	Relative Humidity
RMSE	Root Mean Square Error
Rz	Peak - Valley Height
SE	Secondary Electrons
SEM/EDS	Scanning Electron Microscopy / Energy-Dispersive X-ray Spectroscopy
TAPPI	Technical Association of the Pulp and Paper Industry
THz	Terahertz
UV/VIS/NIR	Ultraviolet/Visible/Near-Infrared
UQ	Uncertainty Quantification
Yb:Glass	Ytterbium-Doped Glass
Yb:KGW	Ytterbium-doped Potassium Gadolinium Tungstate

Introduction

The preservation of cultural heritage artifacts encompasses various materials, ranging from stone sculptures, metal objects, and wall paintings to glass, ceramics, textiles, and illuminated manuscripts. Each material class presents its own set of degradation pathways and cleaning challenges, shaped by intrinsic structure, environmental exposure, and historical usage. Among these, cellulose-based and proteinaceous substrates such as paper and leather are especially sensitive due to their organic composition and complex surface treatments. These objects often accumulate various types of contaminants over time, including soot, dust, grease, biological residues, and pollutants introduced by handling or display environments. Artifacts made from these materials often suffer from surface soiling, chemical alteration, and physical degradation, demanding conservation methods that are both effective and minimally invasive.

Traditional cleaning approaches include mechanical techniques such as dry brushing, erasing, scalpels, and chemical methods involving solvents or gels. While these methods are widely practiced in conservation, they pose risks. Mechanical tools can physically abrade fragile substrates, and chemical agents may diffuse into porous materials, leading to swelling, discoloration, or long-term chemical reactivity. Their effectiveness also depends highly on the operator's skill and may lack reproducibility across different conservators or institutions. In particular, the heterogeneous surfaces of paper and leather complicate the control of solvent penetration and mechanical force application, making precise and uniform cleaning outcomes difficult to achieve. These limitations underscore the need for more controllable, selective, and reproducible cleaning methods, especially when dealing with culturally significant artifacts where intervention must be minimal and well-documented.

Laser-based cleaning, when applied with controlled parameters, offers a compelling alternative to conventional chemical or mechanical methods. Unlike chemical solvents that can diffuse into porous substrates and alter their composition, or mechanical tools that risk abrasion, pulsed laser irradiation can target surface contaminants with high spatial selectivity. Among laser systems, femtosecond lasers provide superior precision due to their ultrashort pulse durations and low thermal diffusion. Single-photon or multiphoton absorption mechanisms may occur depending on the wavelength, fluence, and material properties, making parameter optimization critical to avoid undesired substrate alteration. Femtosecond laser systems, therefore, offer a promising solution for cleaning highly delicate surfaces such as historical paper or leather, where preservation of texture, integrity, and surface features is paramount. While ultraviolet (UV) femtosecond lasers may appear advantageous due to their lower penetration depth and high surface selectivity, their higher photon energy often increases the risk of photochemical side effects, such as bond cleavage, cross-linking, or yellowing, especially in sensitive organic substrates like paper or leather. Despite their theoretical precision, these drawbacks limit the practical use of UV systems for heritage cleaning. On the other hand, near-infrared (NIR) femtosecond lasers offer a distinct advantage due to their low linear absorption in most organic materials, which allows for controlled surface cleaning without inducing excessive thermal or photochemical stress. However, this benefit only holds when the laser is operated below the threshold for nonlinear effects such as multiphoton absorption. At higher fluences, NIR photons can combine through nonlinear interactions to effectively mimic the energy of UV radiation, reaching subsurface layers and potentially triggering undesired reactions such as yellowing or pigment degradation. Therefore, to avoid transferring UV-equivalent energy deeper into the material, especially in the presence of hidden chromophores (e.g., iron gall ink residues, degraded lignin) or sensitive compounds such as organic lake pigments or proteinaceous binders, NIR laser parameters must be carefully tuned to remain within the linear regime. This can provide more controlled ablation with reduced photochemical alteration, making NIR systems a safer and more versatile choice for cleaning applications in cultural heritage conservation.

Cleaning cellulose-based substrates such as paper poses unique challenges. These artifacts are not only fragile but also structurally heterogeneous. Sizing layers applied during manufacturing, often composed of starches or protein-based mixtures, modify the surface chemistry and absorption properties. In real-world conservation scenarios, cleaning often must be performed with extreme selectivity, protecting inked regions, preserving watermark structures, and ensuring that writing on the reverse side is unaffected. These conditions require a cleaning method capable of sub-millimeter control and non-destructive monitoring, making femtosecond laser systems particularly well-suited. Still, quantifying the degree of cleaning without compromising the artifact remains challenging, especially when visual indicators are insufficient. In such contexts, relying solely on visual inspection or colorimetry may be inadequate, underscoring the need for new approaches capable of providing chemically specific, quantifiable, and structurally meaningful insight into surface cleanliness.

Following the analysis of paper substrates, leather substrates introduce a different set of complications. Their microstructure is fibrous and porous, influenced by animal source, tanning methods, and historical aging. Femtosecond laser cleaning of leather must contend with surface unevenness and compositional variability, both of which affect light-material interaction. Additionally, the depth-dependent nature of porosity makes it difficult to predict or model ablation dynamics. In this context, the use of surface-sensitive techniques such as scanning electron microscopy (SEM) or colorimetric evaluation is common. However, they are often limited in precision and repeatability.

Although colorimetry is often used to quantify cleaning outcomes, its interpretation is uncertain. The $DeltaE$ - based color difference metric offers a numerical indicator of change. Still, it is highly sensitive to surface roughness, illumination angle, and microstructural scattering, while it is not linear in the concentration of the residual contamination. In particular, substrates like aged leather exhibit variable reflectance behavior across the visible spectrum, which can lead to misleading ΔE values even in the absence of real material changes. While colorimetry is technically quantitative, it is rarely accompanied by quantified measurement uncertainty or error analysis. Many studies report color shifts without indicating the instrument's repeatability or the influence of sample orientation, lighting geometry, or porosity. This lack of standardization undermines the reliability of colorimetry as a standalone diagnostic tool, especially on uneven or fibrous materials.

In contrast, spectroscopic methods such as ATR-FTIR provide chemically specific and quantifiable information about both the contaminants and the substrate. When coupled with modeling approaches, such data can be used not only to assess cleaning success but also to generate predictive frameworks. In particular, exponential saturation and decay models applied to vibrational bands allow for the estimation of saturation fluence and removal efficiency. However, even these models must account for data variability. Their predictive power remains limited without incorporating slope extraction error, instrumental noise, or parameter correlation.

To overcome these challenges, this thesis introduces a combined modeling and simulation framework to describe and evaluate femtosecond laser cleaning on paper and leather substrates. This framework is designed to bridge the gap between high-precision laser intervention and quantifiable diagnostic assessment, addressing the current lack of standardized tools for evaluating cleaning performance. The proposed approach aims to deliver actionable insight into material response under laser irradiation by integrating experimental spectroscopy, nonlinear fitting, and uncertainty propagation. Ultimately, the goal is to provide a practical decision-making tool for conservators, enabling the selection of optimal cleaning parameters while minimizing the risk of damage to historically significant artifacts.

This thesis is a cotutelle doctoral project between Sapienza University of Rome and the Institut National de la Recherche Scientifique, Canada (INRS). It is structured in four main chapters, each addressing a distinct aspect of femtosecond laser cleaning and its quantitative evaluation on cultural heritage substrates.

Chapter 1 introduces the fundamental principles of femtosecond laser–material interactions and reviews key diagnostic techniques used for surface and material analysis in heritage science. These include ATR-FTIR spectroscopy for chemical characterization, scanning electron microscopy (SEM) for microstructural imaging, profilometry for quantitative surface topography measurements, and colorimetry for objectively evaluating visual changes. Together, these techniques form the analytical basis for assessing cleaning performance, interpreting laser effects, and guiding safe conservation treatments. Together, these techniques form the analytical basis for assessing cleaning performance, interpreting laser effects, and guiding safe conservation treatments.

Chapter 2 presents a spectroscopically guided, model-based methodology for quantifying the removal of artificial surface contaminants from historically representative paper samples. Three sizing types: corn starch, wheat starch, and alum–egg were selected to reflect common historical surface treatments, each influencing laser-material interaction in unique ways. Cleaning experiments were conducted using a 1030 nm near-infrared (NIR) femtosecond laser system across a controlled fluence range. Surface changes were tracked using ATR-FTIR spectroscopy, with all measurements performed at the *Surface Modification and Characterization (SMAC) Group* laboratories of the Roma Tre University. Selected vibrational bands served as molecular indicators for either the contamination layer or the sizing material. These spectral evolutions were fitted with exponential saturation and decay models, enabling the estimation of key parameters such as saturation fluence and removal efficiency. To assess the statistical significance of sizing effects, two-way ANOVA was applied to model outcomes. Additionally, scanning electron microscopy (SEM) and stylus profilometry were employed to evaluate post-cleaning morphology and surface roughness, providing complementary insight into mechanical and optical surface responses.

Chapter 3 investigates the challenges of using colorimetry to evaluate femtosecond laser cleaning on historical leather, a material that, unlike paper, presents complex porosity, irregular topography, and heterogeneous surface chemistry. In this study, all leather samples were prepared in the laboratory to replicate the properties of aged historical leather, including tanning methods and surface degradation. All cleaning experiments were carried out at the *IR and THz Photonics Laboratory* at Institut National de la Recherche Scientifique (INRS), using a Yb-based amplified femtosecond laser operating at 1030nm, under carefully controlled fluence conditions. Tanned and aged leather samples exhibiting varying porosity were subjected to laser treatment. Optical microscopy was employed to examine and compare surface porosity across different leather samples prior to cleaning, using a Zeiss Axio Scope A1 microscope and subsequent ImageJ analysis. The cleaning outcomes were then analyzed using ΔE-based reflectance colorimetry, with all colorimetric measurements performed at the Forensic Science Department of the University of Quebec at Trois-Rivières. Although colorimetry provides numerical outputs, the results revealed considerable variability in ΔE values, largely influenced by surface roughness, scattering effects, and measurement geometry. This chapter highlights the limitations of color metrics in isolation and emphasizes the need for structurally and chemically specific diagnostics when dealing with fibrous or optically complex substrates.

Chapter 4 expands the modeling approach introduced in Chapter 2 by integrating quantified measurement uncertainty into the analysis of cleaning performance. A Monte Carlo simulation framework was developed to propagate spectral variability and slope extraction noise into the nonlinear fitting process. One hundred slope values were generated by perturbing experimental data within measured uncertainty bounds, and each realization was fitted using the same nonlinear model to form an ensemble-based prediction of the cleaning trajectory. This procedure yielded confidence intervals around predicted cleaning trajectories and revealed how prediction uncertainty varies across the fluence spectrum. The findings demonstrate that uncertainty peaks within the transition regions, where spectral signals change most steeply, and diminishes toward asymptotic regimes. These insights not only enhance the interpretability of cleaning models but also provide conservators with more reliable guidance when selecting safe and effective laser parameters.

Chapter 1

Fundamental Principles of Laser Cleaning, Diagnostic Techniques and Statistical Methods in Heritage Science

1.1 Laser–Material Interaction Principles

The interaction between pulsed laser radiation and material surfaces is a key area in modern materials processing technologies. Among these, ultrashort-pulse lasers such as femtosecond sources have gained increasing attention due to their capacity to ablate material with exceptional precision while minimizing thermal damage to the surrounding matrix [1]. These lasers are employed in various applications, including microfabrication, biomedical treatments, and fundamental scientific research, due to their ability to concentrate energy into extremely short time intervals [2].

Comparative studies between femtosecond and nanosecond laser pulses have shown that ultrashort pulses induce ablation primarily through nonthermal mechanisms. In contrast, longer pulses tend to deposit energy more diffusively as heat into the material [3]. Because of their ultrashort pulse durations and high peak powers, femtosecond lasers can trigger multiphoton absorption processes, allowing energy deposition through nonlinear optical interactions rather than purely thermal effects. This difference leads to highly localized energy deposition, which avoids collateral damage and allows for precise surface removal.

The effectiveness of femtosecond laser ablation depends on various parameters such as pulse duration, fluence, repetition rate, and material properties. However, one of the distinguishing advantages of femtosecond interaction is the suppression of heat-affected zones, making it ideal for delicate substrates commonly encountered in conservation and precision engineering [1, 2]. Accurate material processing requires understanding these interactions, as femtosecond lasers not only achieve minimal heat-affected zones but also offer high spatial resolution critical for fragile substrates, with efficiency strongly dependent on material characteristics and laser parameters such as fluence, pulse duration, and ambient conditions [1, 3, 4]. These parameters determine how energy is absorbed and distributed within the substrate.

Ultrashort pulse lasers, particularly femtosecond, but also picosecond systems, are widely used in domains requiring fine control over energy delivery, such as microfabrication or conservation of delicate substrates. Their high peak powers, short pulse widths, and ability to confine energy temporally and spatially make them ideal tools for damage-sensitive applications [5].

Recent studies have highlighted the substantial advantages of ultrashort pulsed lasers over their longer pulsed counterparts [1]. In particular, femtosecond lasers have gained considerable popularity in ablation applications due to their ability to process a wide range of materials while minimizing thermal damage. Unlike nanosecond lasers, such as Nd: YAG systems with pulse durations around 6 ns or 30 ns, which allow energy to dissipate into the lattice as heat, femtosecond lasers concentrate their energy within a timescale shorter than thermal diffusion, effectively channeling it into the kinetic energy of the ablated plume [6]. This mechanism leads to cleaner ablation with minimal collateral effects.

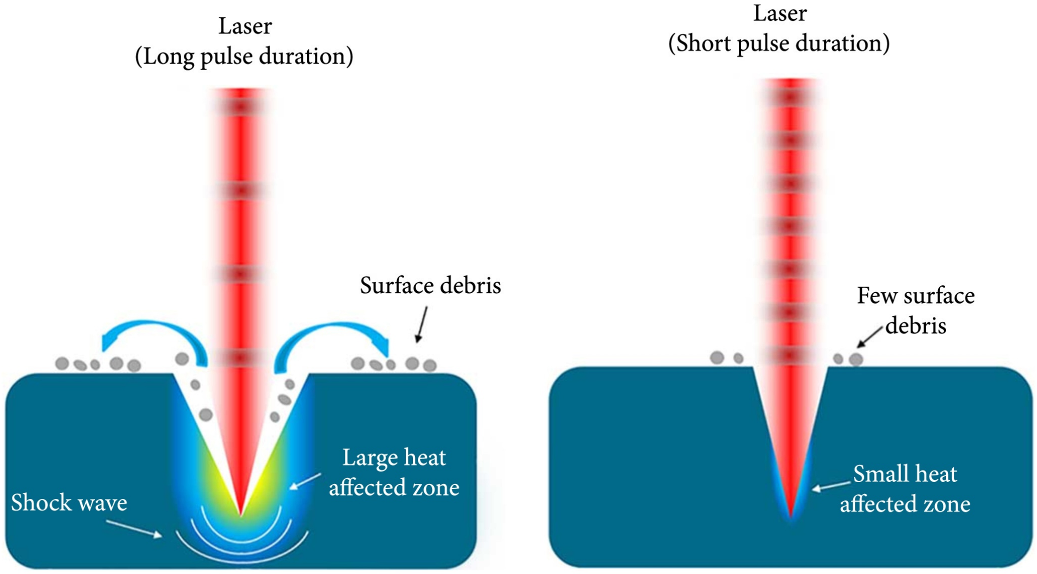


Figure 1.1. Schematic processing comparison of long pulse duration (left), short pulse duration(right) [7].

The schematic presented in Figure 1.1 [7] illustrates a general comparison between short and long pulse durations in laser-material interactions. Although the figure itself labels these regimes simply as “short pulse duration” and “long pulse duration,” such terminology is often associated with specific time ranges in the literature. For instance, microsecond pulses are generally considered long, while femtosecond pulses represent the ultrashort regime. The distinct behavior of these regimes, particularly the minimized thermal diffusion associated with femtosecond pulses, is extensively discussed in prior work comparing microsecond and femtosecond interactions in medical device processing applications [8]. Similarly, studies have also demonstrated the thermal advantages of femtosecond lasers over nanosecond lasers in materials processing contexts, reinforcing the classification of nanosecond pulses within the long pulse regime for comparative purposes [9]. This comparison highlights how the minimized thermal diffusion achieved by femtosecond pulses underpins their suitability for precise cleaning tasks in cultural heritage conservation.

Because many materials used in cutting, drilling, or precision ablation have characteristic response times in the picosecond regime, femtosecond pulses ensure efficient energy deposition with minimal thermal diffusion, enabling athermal or near-athermal ablation that preserves material integrity [10].

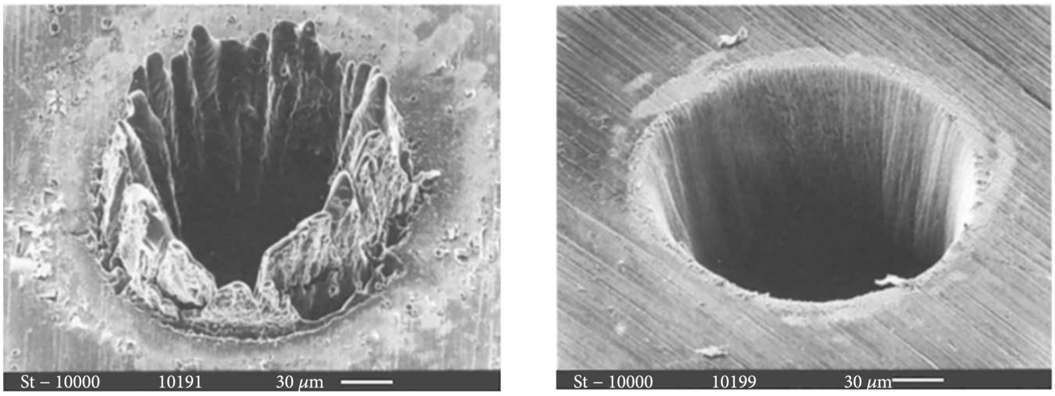


Figure 1.2. Comparative example for nanosecond laser pulses (left) versus femtosecond laser pulses (right) when drilling micro holes in stainless steel [7].

A representative comparison of femtosecond and nanosecond laser ablation is illustrated in Figure 1.2. These SEM images, as reported by Lin and Hong [7], show micro-holes fabricated on a 100 μm steel foil using (left) a 780 nm nanosecond laser (3.3 ns, 0.5 J/cm^2) and (right) a 780 nm femtosecond laser (200 fs, 0.5 J/cm^2). This comparison illustrates the superior precision and reduced thermal damage achieved with femtosecond laser ablation, supporting their application in the safe cleaning of delicate heritage materials.

Femtosecond lasers are widely regarded as the optimal choice for minimizing thermal damage during micromachining. Achieving such precise results depends heavily on the optimization of the optical setup, particularly in managing pulse broadening effects. A well-designed system allows for effective control of dispersion, ensuring that the ultrashort pulses maintain their integrity upon reaching the target [5].

Another fundamental aspect of laser–material interaction is the wavelength-dependent absorption behavior of the irradiated medium. The absorption spectrum of a material governs how incident laser energy is deposited across its depth, influencing not only the efficiency of ablation but also the extent of thermal or photochemical effects. Materials typically exhibit distinct absorption characteristics across ultraviolet (UV), visible, and near-infrared (NIR) regions, which in turn shape their response to laser irradiation. In general, shorter wavelengths, such as those in the UV range, are absorbed more strongly by organic and many inorganic materials, resulting in very shallow penetration depths. While this promotes high spatial selectivity and precise surface interaction, the high photon energy associated with UV light increases the likelihood of photochemical side effects, including bond cleavage, cross-linking, and molecular degradation. These effects may induce unwanted yellowing, particularly in chemically sensitive substrates [11, 12].

On the other hand, NIR femtosecond lasers offer a unique advantage in heritage cleaning due to their low linear absorption in many organic and dielectric materials. When carefully operated below the threshold for multiphoton absorption, they enable controlled surface interactions without triggering deep or irreversible modifications. This selectivity is particularly valuable when contaminants strongly absorb NIR radiation, while the underlying artwork, comprising cellulose fibers, natural dyes like indigo or cochineal, or sizing materials such as alum-egg mixtures, remains largely transparent. In such cases, efficient contaminant removal is achievable without compromising the visual or chemical integrity of the artifact.

To conclude, the unique properties of femtosecond lasers, namely their ultrashort pulse durations, high peak intensities, and ability to confine energy spatially and temporally, make them highly suitable for delicate material processing. These features are particularly advantageous in the field of cultural heritage conservation, where historical substrates such as paper, parchment, or leather are extremely sensitive to heat and mechanical stress. Femtosecond lasers allow for precise surface cleaning with minimal collateral damage, making them an ideal choice for safely removing contamination from fragile cultural artifacts without compromising their integrity.

1.2 Diagnostic Techniques for Surface and Material Analysis

1.2.1 ATR-FTIR Spectroscopy

Fourier Transform Infrared (FTIR) spectroscopy is a vibrational spectroscopic technique that provides valuable molecular-level information based on the interaction between infrared radiation and matter. Molecules absorb IR radiation at specific wavenumbers corresponding to their characteristic vibrational modes, which are influenced by bond strengths and atomic masses. The resulting spectrum acts as a molecular fingerprint, allowing for qualitative and quantitative analysis of various substances [13, 14].

In an IR-active vibrational mode, energy absorption is possible only if the vibration induces a change in the dipole moment of the molecule. Thus, symmetric vibrations in nonpolar molecules may remain undetected by IR spectroscopy [15, 16]. Vibrations in IR spectroscopy are broadly categorized as stretching or bending motions. Stretching involves changes in bond length, while bending refers to changes in bond angle. Bending can further be classified into scissoring, rocking, wagging, and twisting deformations, each providing unique spectral features [17].

Among the operational modes of FTIR, such as transmission and reflection, Attenuated Total Reflection (ATR) is particularly advantageous in heritage science due to its minimal sample preparation and high surface sensitivity. It is especially well-suited for the surface analysis of solid or layered samples. ATR relies on the principle of total internal reflection, wherein an infrared beam is directed into a high-refractive-index internal reflection element (IRE) such as diamond, germanium, or zinc selenide (ZnSe). As the beam reflects within the IRE, it generates an evanescent wave at the interface, which penetrates only a short distance into the sample [13, 14].

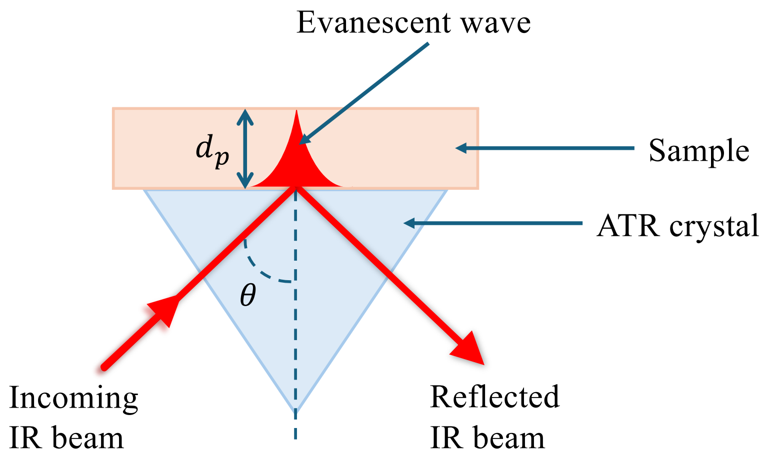


Figure 1.3. Schematic representation of the ATR principle, showing the incoming IR beam, total internal reflection in the ATR crystal, and evanescent wave penetration into the sample.

Figure 1.3 illustrates this principle, showing how the infrared beam undergoes total internal reflection within the ATR crystal and produces an evanescent field that interacts with the sample surface. The penetration depth d_p of this evanescent wave is given by equation 1.1, which shows that it depends on the wavelength, refractive indices of the crystal and sample, and the angle of incidence:

$$d_p = \frac{\lambda}{2\pi \left(n_c^2 \sin^2\theta - n_s^2\right)^{1/2}} \tag{1.1}$$

where λ is the wavelength of the incident IR radiation, n_c and n_s are the refractive indices of the ATR crystal and sample, respectively, and θ is the angle of incidence [15]. Given that all expressions in the denominator of equation 1.1 (such as the refractive indices and the angle-dependent term) are typically of the order of 1, the resulting penetration depth d_p is expected to be on the same order of magnitude as the incident wavelength. Such shallow penetration makes ATR particularly effective for surface-sensitive diagnostics in heritage science, providing valuable information on coatings, contamination, and degradation layers without interference from subsurface features.

ATR-FTIR offers several advantages in cultural heritage research: it requires minimal sample preparation, is non-invasive, and allows for repeatable surface characterization. Its applicability spans a range of substrates, including paper, parchment, textiles, varnishes, and paint layers [17, 16]. Furthermore, modern ATR-FTIR spectroscopic imaging not only records spectra but also integrates spatial resolution capabilities, enabling two-dimensional chemical mapping at the micrometer scale, which reveals the distribution of components such as pigments, binders, or degradation products across the sample surface. This is especially useful for investigating heterogeneous and and the surface layers of multilayered objects such as manuscripts, paintings, or laminated materials [18].

By capturing both spectral and spatial information, ATR-FTIR spectroscopy enables in situ chemical diagnostics of priceless artifacts without physical removal or alteration. Overall, it bridges fundamental molecular analysis with practical conservation needs, ensuring both analytical depth and preservation of cultural heritage materials. This combination of analytical power and non-invasive accessibility has rendered it an essential tool in heritage science.

1.2.2 Scanning Electron Microscopy

Scanning Electron Microscopy (SEM) is a powerful imaging technique that utilizes a focused beam of high-energy electrons to produce high-resolution images of material surfaces. Unlike optical microscopy, which relies on visible light, SEM exploits the much shorter wavelengths of electrons to achieve magnifications up to the order of one million times, with resolution in the nanometer range [19]. The primary signals generated during SEM analysis are secondary electrons (SE), which reveal surface morphology, and backscattered electrons (BSE), which provide compositional contrast based on atomic number. These electrons are emitted as the focused beam raster scans across the specimen in a vacuum chamber, and they are detected to form a detailed topographic image [20].

A typical SEM setup includes an electron gun, electromagnetic lenses for beam focusing, scanning coils for raster movement, and detectors for SE and BSE signals. The interaction volume from which signals are generated is generally limited to the top few microns of the sample, making SEM highly surface-sensitive [19]. One of its key advantages is the large depth of field compared to light microscopy, which allows for sharply focused imaging of uneven surfaces. Moreover, when SEM is coupled with Energy Dispersive X-ray Spectroscopy (EDS), which detects characteristic X-rays emitted from the sample to determine elemental composition, it offers a dual capability of morphological and compositional analysis [21].

In cultural heritage science, SEM has proven indispensable due to its ability to investigate materials at micro- and nano-scales with relatively simple sample preparation in most cases, though requirements may vary with object dimensions and the goal of analysis. Paint cross-sections, textile fibers, and archaeological glass fragments are routinely analyzed to understand fabrication methods, degradation processes, and to guide conservation strategies. For example, Schreiner et al. [21] demonstrated how SEM-EDS analysis can differentiate pigment layers and detect alteration products in historical paintings. Similarly, Kontogeorgopoulou et al. [22] reviewed SEM applications in architectural heritage, highlighting its use in evaluating stone, mortar, and ceramic microstructures.

Textiles and organic artifacts are another significant application area. Lukesová et al. [23] emphasize the role of SEM in distinguishing plant fiber types in ancient fabrics, while Burattini and Falcieri [24] illustrate its utility in identifying fungal decay in archaeological wood. In these studies, SEM enables conservators to examine structural details and assess the effectiveness of treatment methods. As Botti et al. [25] further show, SEM imaging, combined with spectroscopic techniques, is critical in diagnosing degradation in historical paper artifacts.

Overall, SEM remains a cornerstone technique in cultural heritage diagnostics, providing unmatched morphological insights and, when coupled with EDS, elemental information. Its adaptability to diverse materials and compatibility with other analytical methods make it a valuable complement to other analytical methods in the conservation scientist's toolkit.

1.2.3 Profilometry

Profilometry is a surface metrology technique encompassing both contact (stylus) and non-contact (optical) methods, enabling the measurement of height variations (surface topography) across an object. This allows for the quantification of roughness and texture, which plays a vital role in cultural heritage diagnostics by allowing conservators to monitor degradation, evaluate restoration effects, and study surface features at micro-scale resolution.

Traditional stylus profilometers utilize a contact-based mechanical probe to trace the surface and produce a 2D cross-sectional profile. In contrast, optical profilometers use laser or white light to acquire non-contact 3D surface maps, which is essential for fragile or valuable artifacts. Key roughness parameters such as Ra (mean roughness) and Rz (peak-valley height) are standardized descriptors for assessing the surface condition [26, 27].

Among optical methods, confocal microscopy and white-light interferometry offer high-resolution height measurements. These techniques are capable of detecting sub-micrometer topographic changes, which is critical when monitoring slow weathering processes or subtle treatment effects. Structured-light and fringe projection systems provide fast, full-field scans, although often at slightly reduced resolution [28, 29].

In cultural heritage applications, profilometry has been used for documenting artist brushstrokes, studying craquelure patterns, analyzing wear on ancient coins, and monitoring stone erosion. Mironova et al.[30] demonstrated the use of optical profilometry to capture roughness profiles of painting brushstrokes, enabling better digital preservation and technique attribution. Mazzocato and Daffara [31] applied conoscopic holography for micro-profilometry of sculptural elements by combining height and intensity signals acquired from the same surface. To overcome the typical alignment challenges in profilometry, particularly when dealing with heterogeneous materials or curved surfaces, they employed a spatial registration strategy that aligned intensity and height datasets pointwise. This approach enabled more accurate 3D reconstructions and allowed the correlation of micrometric topography with visual features such as pigments or tool marks.

Profilometric analysis has also proven useful in evaluating conservation treatments. Surface roughness measurements before and after cleaning operations on stone or metal can reveal whether treatments are non-invasive and preserve the original microstructure [28]. By providing quantitative, repeatable surface data, profilometry serves as a critical link between material characterization and conservation decision-making in heritage science.

1.2.4 Colorimetry

Colorimetry provides a non-invasive and objective method for quantifying color as perceived by the human eye, making it a vital diagnostic tool in the field of cultural heritage science. At the core of colorimetry lies the CIE 1931 XYZ system, which remains the reference standard. However, in cultural heritage science, the CIELAB color space is commonly employed due to its perceptual uniformity and intuitive interpretation, where color is represented through lightness (L), red-green (a), and yellow-blue (b) coordinates (as illustrated in Figure 1.4[32]). This standardized color model allows precise numerical comparison of color measurements taken at different times or under different conditions, forming the basis for evaluating degradation, treatment effects, or pigment distributions in historical artifacts [33, 34].

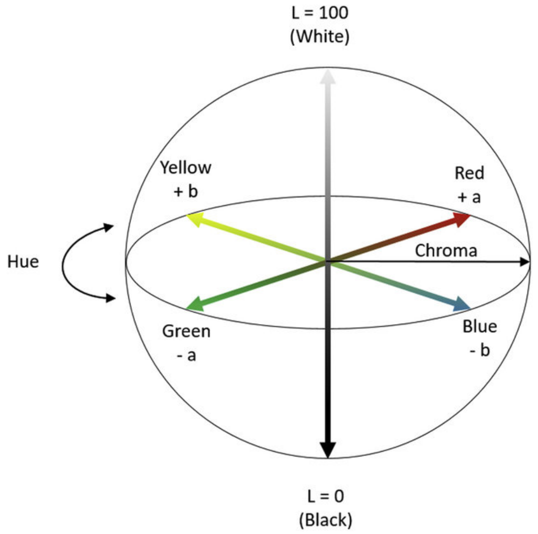


Figure 1.4. Three-dimensional representation of the CIELAB color space, showing the lightness (L*), red-green (a*), and yellow-blue (b*) axes. Chroma is the radial distance from the origin, while hue is the angular position around the L* axis [32]. This perceptual model underlies quantitative color difference calculations in cultural heritage analysis.

The key metric derived from CIELAB measurements is the total color difference, ΔE^*, calculated as the Euclidean distance between two points in the Lab space given in equation 1.2 below:

$$\Delta E^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \tag{1.2}$$

This metric enables quantification of even subtle changes in surface appearance, such as those caused by aging, environmental exposure, or conservation interventions [33]. In practical conservation scenarios, ΔE^* values are interpreted using perceptibility thresholds. According to Mokrzycki and Tatol [33], $\Delta E^* < 1$ is typically imperceptible, 1–2 may be detected only by experienced observers, while $\Delta E^* > 3.5$ constitutes a clear and visible difference. These thresholds provide a practical framework for conservators to assess whether a color change is visually significant, guiding treatment acceptability criteria.

Applications of colorimetry in heritage science are diverse and have a significant impact. One of the most common uses is monitoring the effects of environmental stressors or artificial aging. For example, Cabello Briones et al. used colorimetric analysis to evaluate the impact of dust deposition and cleaning methods on Roman mosaics. By measuring color coordinates before and after treatment, they showed that wet cleaning recovered color more effectively than dry methods, with lower ΔE^* values indicating better visual restoration [34]. Similarly, colorimetry has been used to assess the visual integrity of surface coatings or consolidants on stone and mural paintings, ensuring that protective treatments do not induce perceptible alterations [33].

Moreover, the selection of measurement conditions, specifically the illuminant and standard observer, plays a critical role in ensuring consistent and meaningful results. Most studies adopt the D65 illuminant (representing average daylight) and the 10° standard observer to simulate typical viewing conditions for museum audiences [35, 36]. These standards ensure that color measurements reflect realistic visual experiences and facilitate comparability across instruments and studies. Spectral imaging and colorimetric calibration have also been employed in 3D photogrammetry and virtual documentation workflows to maintain color fidelity in digital reconstructions of cultural objects [37].

Colorimetry bridges subjective human color perception and quantitative scientific measurement. Its integration into heritage diagnostics empowers conservators to document visual changes with precision, optimize treatment approaches, and ensure the aesthetic preservation of cultural artifacts. However, ΔE^* is not linearly correlated with the concentration of dyes or absorbers, including surface contaminants. This nonlinearity limits its sensitivity, particularly in detecting thin or sparsely distributed residues following laser cleaning. Therefore, laser cleaning requires complementary evaluation metrics that provide a more direct physical correlation to the contaminant layer, such as the intensity of characteristic absorption bands obtained from vibrational spectroscopy. These band intensities enable a quantifiable estimation of the contaminant thickness and chemical composition.

Moreover, when the contaminant produces little or no change in the visible reflectance of the surface, as in the case of kaolinite, colorimetric methods may fail to detect its presence since little optical contrast is generated. Under such circumstances, complementary spectroscopic methods become essential. ATR-FTIR spectroscopy, for instance, can probe the vibrational modes of specific molecular groups irrespective of their visibility in the optical range, offering a more reliable means to detect and quantify residual contamination.

1.3 Statistical Tools for Data Interpretation: ANOVA and Monte Carlo Simulations

Analysis of Variance (ANOVA) is a statistical method to test whether the differences observed between the means of three or more groups are greater than would be expected from random variation alone, in both single-factor (one-way) and multifactor (two-way or higher) designs, particularly when multiple factors or experimental conditions are involved. In other words, it evaluates whether the variation between group means reflects a systematic effect of the factors under study rather than chance fluctuations. It decomposes the total observed variance into components attributable to different sources, allowing researchers to identify which variables significantly affect a given outcome[38]. This is especially important in cultural heritage science, where materials and conditions are inherently heterogeneous, requiring robust statistical tools to account for such variability. For example, ANOVA has been used to evaluate differences in microbial deterioration activity on photographic artworks [39], as well as to analyze how environmental factors like temperature and humidity impact the preservation of earthen heritage sites [40]. Additionally, factorial ANOVA can be employed to examine potential interaction effects between independent variables, such as cleaning method and laser fluence, which is critical when evaluating complex treatments in conservation science. Repeated measures ANOVA is also useful when analyzing data collected from the same samples over time, for example, to assess the durability of cleaning treatments under accelerated ageing conditions.

While ANOVA provides a powerful statistical evaluation of experimental data, computational modeling approaches such as Monte Carlo simulations offer complementary insights by predicting physical interactions and outcomes based on probabilistic frameworks. Together, these tools enable a more comprehensive understanding of material behavior and treatment effects in cultural heritage conservation.

Monte Carlo simulation is a computational method that uses repeated random sampling to model and analyze complex systems. It has become increasingly valuable in cultural heritage science, where destructive testing is often not an option and many variables remain uncertain. By generating large numbers of possible scenarios, Monte Carlo simulations provide a statistical framework to assess risks and predict system behavior under uncertainty. The technique allows researchers to simulate a range of possible outcomes for a given process, offering insight into the underlying variability of materials, treatments, or structural behavior, as demonstrated by Bottaini et al., who used Monte Carlo simulations to model and correct for compositional and surface-related uncertainties in nondestructive ED-XRF analysis of copper-based archaeological artifacts [41]. Its strength lies in the ability to incorporate uncertainty directly into the modeling framework, making it particularly suited for heritage contexts where experimental control is limited, such as when only a single fragile artifact can be analyzed and systematic replication is not feasible.

One major application of Monte Carlo methods is in uncertainty quantification (UQ). For instance, Monte Carlo simulations have been used to estimate the variance of predicted degradation patterns in heritage collections exposed to fluctuating environments, or to analyze how environmental and topographic factors influence site selection and preservation conditions, as demonstrated in the study of the Ming Yansui Great Wall heritage, where Monte Carlo modeling was used to assess site preferences relative to terrain features [42]. In such studies, agent-based models or spatial simulations are combined with stochastic sampling to explore how chemical aging, accidental events, or geographic variables jointly influence heritage conditions over time. This approach enables conservators to prepare for both typical and extreme degradation scenarios. By repeating simulations with varied input parameters, researchers can predict not only average behavior but also the likelihood of rare but damaging outcomes, improving preventive conservation strategies.

In the field of radiation-based treatment and analysis, Monte Carlo simulations provide a non-invasive means to evaluate how ionizing radiation penetrates cultural artifacts. Morgunov et al. employed the GEANT4 Monte Carlo toolkit to simulate X-ray dose distribution inside wooden artworks exposed to accelerator-based sterilization sources [43]. Their simulations showed good agreement with measurements, allowing precise predictions of radiation effectiveness and safety for conservation. Similarly, in material analysis, Monte Carlo modeling of SEM-EDS setups helps optimize measurements for ultrathin fragments, essential for accurate mineral or pigment characterization without damaging the object [44].

Monte Carlo methods are also used in structural safety assessments of historical buildings. When material properties or loading conditions (e.g., during an earthquake) are uncertain, hundreds or thousands of simulations with randomly sampled inputs can provide fragility curves and vulnerability functions that describe the probability of damage at different stress levels. For example, Barbat and Moya used Monte Carlo simulations to generate damage probability matrices and fragility curves for unreinforced masonry buildings in Barcelona, incorporating variability in building typologies and seismic behaviour to support urban seismic risk assessments [45].

Such probabilistic assessments are invaluable for heritage risk management, as they enable prioritization of reinforcement interventions based on quantifiable damage likelihoods.

Chapter 2

Quantifying Selective Pulsed Laser Cleaning and Residual Contamination in Paper Artifacts by ATR Spectroscopy

Foreword

This chapter is based on the content of the following article; as a result, some passages may appear verbatim from this publication:

Quantifying Selective Pulsed Laser Cleaning and Residual Contamination in Paper Artifacts by ATR Spectroscopy, submitted in 2025.

C. Y. Boynukara, V. Graziani, M. Uguryol, G. Mavili, L. Tortora, A. Ruediger, and P. Antici,

This work was led by C. Y. Boynukara, who performed the laser cleaning experiments, conducted the data analysis, and coordinated the overall research. V. Graziani and L. Tortora provided guidance on ATR-FTIR spectroscopy and assisted with data acquisition. M. Uguryol and G. Mavili supervised the preparation of the cellulose-based mock-up papers reproducing historical Turkish-Ottoman material and supported the experimental strategy. A. Ruediger contributed to the development of the conceptual framework and advised on modeling and interpretation. The manuscript was written by C. Y. Boynukara, with feedback and editorial input from M. Uguryol, A. Ruediger, and P. Antici.

Abstract

Pulsed laser cleaning has emerged as a promising method for addressing conservation challenges in cultural heritage, particularly for organic substrates such as paper, parchment, or leather, which are significantly more sensitive to degradation than inorganic materials like corroded metals. In such contexts, the cleaning process itself poses risks of damage, particularly when contaminants are deeply embedded. While pulsed laser cleaning offers the potential for partially automated workflows that reduce both time and cost, it remains in an early stage of development. Although certain fundamental parameters, such as wavelength and pulse duration, have been explored, only a few have become widely established or standardized in conservation practice. However, further parametric investigation is needed to define the damage thresholds that vary across substrate types. This study proposes a quantitative framework to evaluate femtosecond laser cleaning for laboratory-prepared replicas of cellulose-based historical papers by incorporating two model contaminants, carbon and mineral-based particles, and by systematically examining the effects of surface burnishing. Using ATR-FTIR spectroscopy in combination with exponential saturation and decay models, cleaning efficiency and substrate preservation were assessed within a fluence range of 0.39–0.78 J/cm^2. The method proved effective in removing difficult residues like kaolinite with minimal remaining contamination. Furthermore, non-burnished samples responded more effectively to laser cleaning, as burnishing was found to embed contaminants more deeply. These results highlight the potential of femtosecond laser cleaning as a selective, controlled, and non-invasive technique and offer a basis for developing targeted cleaning protocols tailored to the sensitivity and condition of historical artifacts.

2.1 Introduction

Cleaning is a critical process in conservation that must be carefully performed to preserve the authenticity of cultural heritage objects [46]. Artworks and artifacts face degradation from various biological, physical, and chemical factors. For instance, biological damage is caused by fungi, bacteria, and insects, while physical degradation stems from temperature and humidity fluctuations. Chemical deterioration can result from a variety of factors, including pollutants, light exposure, and acidic inks, among others. Vulnerable objects such as historical documents and paper-based artworks are therefore at particularly high risk. Traditional conservation methods, including washing, alkalization, and mechanical cleaning, have been developed to counteract such degradation. While washing and mechanical cleaning are true cleaning approaches, alkalization is more properly a stabilization treatment aimed at neutralizing acidity [47, 48]. While effective in many cases, each method carries limitations. Washing can harm inks and must be tested before application, and choosing the right solvent is crucial for stain removal [47]. When water fails to remove acidity, basic salt solutions like magnesium bicarbonate or calcium hydroxide are used for alkalization. Alternatively, non-aqueous suspensions of these salts can be sprayed onto acidic paper. While deacidification can be beneficial, it may cause color changes due to pH-sensitive components in the paper [48]. Mechanical cleaning methods, such as using soft brushes, sponges, vinyl erasers, and gentle vacuuming, are effective and minimally invasive. Tools like scalpels, aspirators, and specialized vacuum cleaners help remove insect residues and mold, while freezing can kill pests [48].

Across all approaches, the core challenge remains the same: removing surface contaminants without damaging the underlying material. Successful cleaning strikes a balance, sufficient to restore appearance and stability, yet gentle enough to preserve original features. This balance is often maintained through the expertise of trained conservators.

Lasers offer a powerful alternative to traditional cleaning methods and were first introduced in art conservation in 1971 during a restoration campaign in Venice [49]. Over the past fifty years, laser cleaning has become widely accepted for various materials such as metals, stone, ceramics, textiles, and paintings [50]. Due to the time-consuming and costly nature of manual cleaning and the increasing demand for restoration, laser cleaning has gained traction for its precision and partial automation. While expert oversight is still valuable, time and cost constraints favor an automated approach.

The development of non-destructive cleaning and diagnostic techniques is vital for the effective treatment of cultural objects. Fast (ns) and ultra-fast (fs) pulsed lasers have become more commercially accessible and thus gained attention [51]. While initially focused on stone, ultra-fast laser cleaning has now expanded to metals, textiles, paintings, and paper [52]. Cleaning paper artifacts is especially difficult due to the need to preserve the delicate organic fiber structure [53]. UV lasers, while offering excellent depth control due to their shallow penetration, can cause irreversible yellowing of cellulose.

Laser cleaning, as a non-contact method, offers control over parameters such as wavelength, pulse length, and polarization, making it a superior alternative to traditional wet or dry mechanical methods [54, 55]. Its advantages include speed, cost-efficiency, and enhanced safety [56]. Lasers are categorized based on wavelength and pulse duration. Laser wavelengths range from infrared (IR) to ultraviolet (UV), while pulse durations span from nanoseconds to femtoseconds. The wavelength controls the penetration depth into the surface, ranging from nanometers for UV to millimeters for IR, and offers selective targeting of specific absorbers. Pulse length determines thermal dissipation; for longer pulses, the heat spreads, causing moderate heating over a larger area.

In contrast, femtosecond pulses deposit energy in a very short time, leading to immediate surface evaporation. Femtosecond infrared lasers are often preferred for cleaning paper-based materials due to their ability to minimize both thermal and photochemical damage, particularly when operated with moderate fluence and stable scanning parameters. Moderate fluence reduces the probability of multiphoton absorption, which can lead to uncontrolled ablation or structural degradation, while stable scanning prevents local energy accumulation that might otherwise cause localized overheating or damage. The combination of wavelength and pulse length, as well as the staining issues caused by UV lasers on cellulose, has pushed femtosecond

IR lasers to the forefront of cultural heritage cleaning. Another critical factor is laser fluence, since excessively high fluences, especially at UV wavelengths, can cause discoloration or chemical degradation due to the high single-photon energy. The laser fluence, or energy per area, is particularly important in pulsed laser systems. Excessively high fluences, especially at ultraviolet (UV) wavelengths, the high photon energy enables direct bond-breaking, which increases the risk of photochemical degradation in sensitive materials such as cellulose. By contrast, near-infrared (NIR) femtosecond lasers offer greater selectivity due to their lower linear absorption in most organic substrates. When applied below the threshold for nonlinear effects, they allow for effective contaminant removal while minimizing interaction with the underlying layers. This is particularly important in heritage conservation, where subsurface chromophores, such as natural dyes or metal-based inks, may be unintentionally altered. Ensuring that the applied fluence stays below the multiphoton absorption regime is therefore critical for maintaining both material integrity and cleaning precision.

It has been shown that high laser fluences or UV wavelengths can negatively impact paper by causing discoloration or altering chemical and mechanical properties due to their high photon energy [57, 58, 59]. This study aims to determine the optimal fluence for a femtosecond laser system to clean model pollutants without damaging cellulose or sizing. Two model pollutants were used: graphite and kaolinite, which represent two contrasting behaviors in the infrared region. Graphite was selected as a model contaminant because its optical properties resemble those of soot and other carbonaceous deposits commonly found on cultural heritage objects due to combustion sources such as candles, oil lamps, or air pollution. Like these deposits, graphite exhibits strong broadband absorption across the UV–visible–NIR spectrum, as shown in Figure 2.1a.

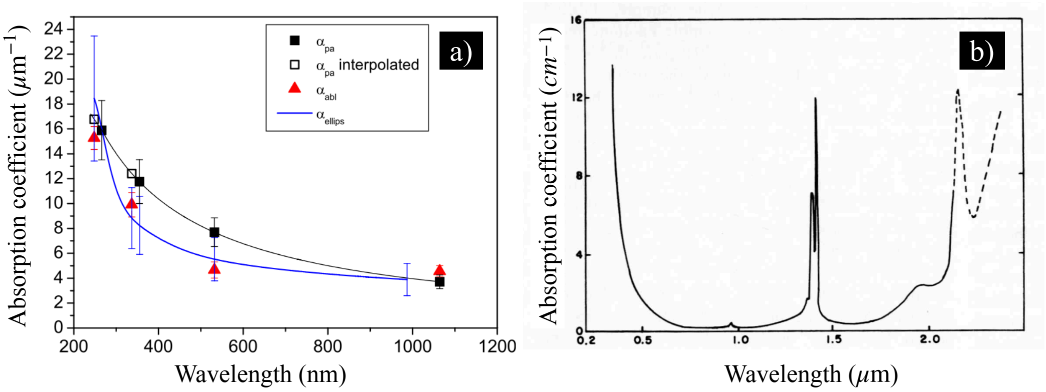


Figure 2.1. Absorption coefficient spectra of a) graphite [60], and b) kaolinite [61].

In contrast, kaolinite, a representative clay mineral found in dust and environmental deposits, displays significantly lower absorption in the visible and NIR regions, as seen in Figure 2.1b.

As illustrated in Figure 2.1 [60, 61], graphite exhibits a high absorption coefficient of approximately $4\,\mu\text{m}^{-1}$ at 1030 nm, resulting in efficient surface energy deposition. In contrast, kaolinite displays extremely low absorption, below $1\,\text{cm}^{-1}$ or $10^{-4}\,\mu\text{m}^{-1}$, indicating that 1030 nm radiation largely penetrates through it without significant interaction. This vast difference in absorption, exceeding four orders of magnitude, helps explain why carbon-based contaminants are more responsive to laser ablation, while low-absorbing mineral particulates such as kaolinite pose a greater challenge. Their high optical transparency in the near-infrared range allows the radiation to penetrate without significant interaction, reducing energy deposition and cleaning efficiency. We therefore chose kaolinite as a challenging pollutant.

Effective cleaning requires non-destructive methods to measure the impact of laser pulses on the surface, ensuring the technique has enough penetration to monitor both contaminants and the underlying surface for signs of recovery or damage. While colorimetric measurements (e.g., ΔE) provide a useful visual metric for perceptual changes, they may fail to detect residual contamination if the color change is negligible or absent. For instance, $\Delta E = 0$ does not necessarily imply that all pollutants have been removed. Therefore, complementary techniques with molecular specificity are required to evaluate the cleaning outcome more reliably. Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectroscopy and imaging enable detailed surface analysis, considering the necessary penetration depth [62, 63].

This study utilizes ATR-FTIR spectroscopy in the mid-IR range (4000-400 cm^{-1}) to monitor the laser cleaning of artificially contaminated paper samples and evaluate the quality of the underlying cellulose. Infrared spectroscopy identifies specific vibrational signatures from functional molecular groups, allowing for a quantitative analysis of selective cleaning and residual contamination.

The current research focuses on paper samples modeled after Turkish-Ottoman manuscripts, which feature a unique surface treatment called “ahar” (a sizing applied to the paper’s surface). Ahar enhances the paper’s mechanical strength and sheen, preventing ink spread and allowing for corrections. Sizing, commonly used in cellulose-based materials (manuscripts, books, paintings), fills microscopic gaps between fibers, improving ink and color flow by replacing capillary action with a more sponge-like behavior. Therefore, sizing and its optical properties must be considered during laser cleaning. Optimizing laser parameters is crucial for cleaning Ottoman paper artifacts. Comparative models that analyze how contaminated papers respond to laser cleaning are foundational for developing effective techniques. All samples for this research were created using the most common sizings in Turkish-Ottoman archives, including alum-egg, corn starch, and wheat starch.

2.2 Materials and methods

2.2.1 Paper samples

As mentioned earlier, laser cleaning works by absorbing the laser pulse energy and converting it into heat on the surface. This process is largely influenced by the absorption characteristics of the molecular species present on the surface in significant amounts. The more information gathered about the surface, the more precisely the cleaning process can be tailored. Knowledge about any prior chemical or mechanical treatments on the artifact, which can be obtained through historical records (such as conservation reports), additional scientific analysis (e.g., spectroscopy), or material testing, should be considered before performing laser cleaning.

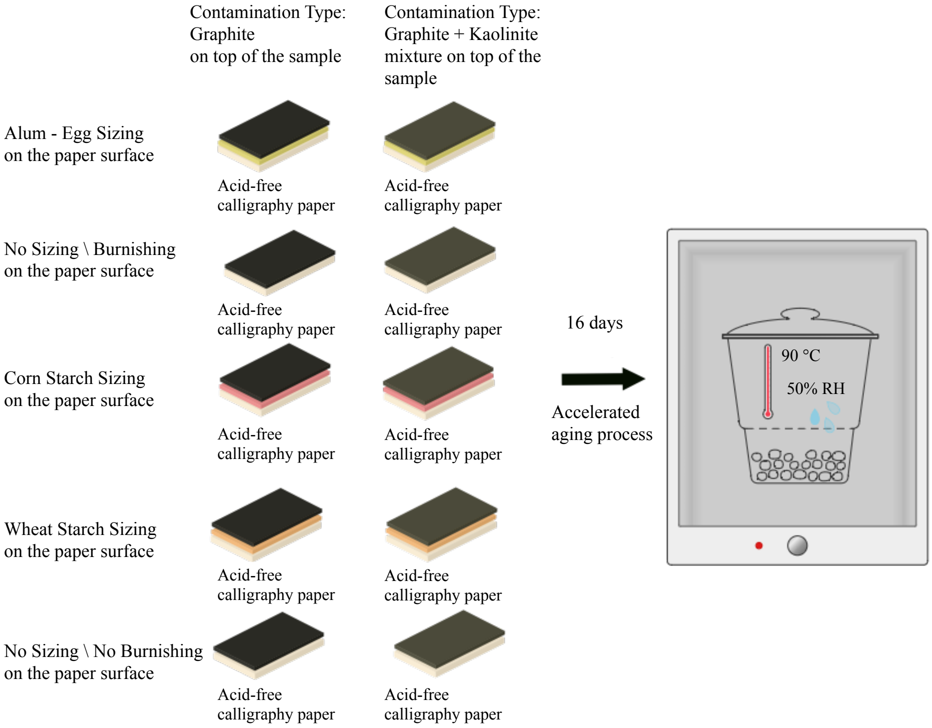


Figure 2.2. Scheme of sample preparation.

The variables considered in preparing the samples for this study include paper sizing, burnishing, and the type of contamination. Alum-egg sizing, corn starch sizing, and wheat starch sizing were chosen for the sample preparation, as these are among the most common sizing techniques used in Turkish-Ottoman manuscripts, representing traditional sizing methods. Samples without any sizing were also prepared. It is important to note that when sizing is applied, burnishing is always performed because sizing is traditionally applied wet with a brush, followed by burnishing after it dries. Therefore, all samples with sizing were also burnished. Historically, the burnishing process, traditionally done by rubbing a smooth, rounded tool (burnisher) with pressure on a paper placed on a hard surface, is used to create a more uniform, smooth surface and provide better conditions for ink transfer. After the samples were sized, they were burnished using a smooth glass burnisher, chosen specifically for its ability to press the paper's surface evenly. Samples without sizing were also prepared without burnishing to examine any potential differences from this treatment.

Two contamination types were selected: graphite and a graphite-kaolinite mixture. Graphite powder (AF spezial, Kropfmühl, D90 $\leq$ 25 μm) represents unburned carbon particles (soot and char) produced by the combustion of fossil fuels or from burning candles and oil lamps, which were relevant during historical periods or after fires. Kaolinite, a clay mineral, represents dust from the weathering and erosion of natural rocks or man-made materials containing silica, such as concrete or cement. These particles can originate from building materials or environmental sources, like gardens or unpaved roads around storage areas. Although kaolinite is occasionally used as a filler in paper manufacturing [64, 65], preliminary ATR-FTIR analysis confirmed that the acid-free calligraphy papers used in this study do not exhibit any of the characteristic kaolinite bands. As shown in Figure 2.3, the ATR-FTIR spectrum of the no sizing, uncontaminated, and aged paper reveals no absorption features corresponding to kaolinite, specifically the OH stretching near $3650\,\mathrm{cm}^{-1}$ and the deformation bands in the $900–750\,\mathrm{cm}^{-1}$ range [66, 67]. This confirms that any kaolinite observed in the post-cleaning analyses originates solely from the applied contamination.

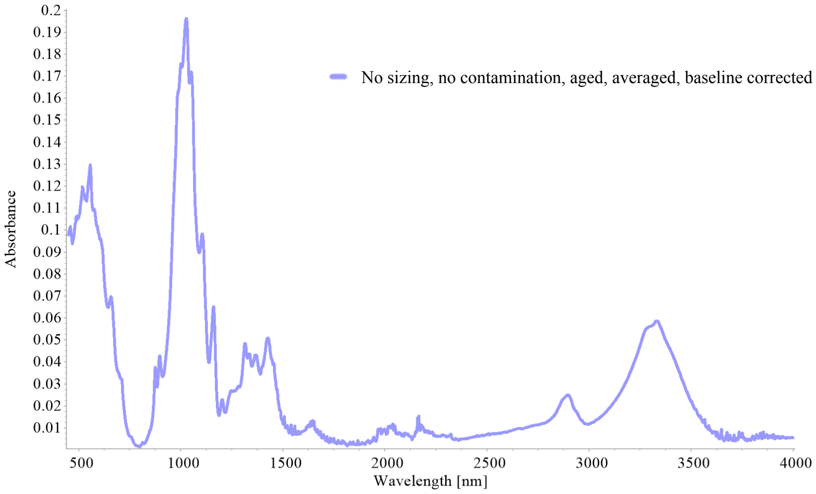


Figure 2.3. ATR-FTIR spectrum of the no-sizing, uncontaminated, and aged paper used in this study, showing no detectable bands for kaolinite.

In our experiment, these contaminants were manually applied to the paper samples with a brush, simulating environmental contamination in a controlled and repeatable way to study laser cleaning efficiency. All contaminated samples were burnished to improve the adhesion of the particles. Additionally, we prepared a sample without sizing that was burnished after contamination to allow deeper penetration of the pollutants into the sample.

Fig. 2.2 illustrates the sample preparation combinations. A detailed description of the sample preparation process is available in [68]. Lastly, all samples were aged under the environmental conditions specified by the Technical Association of the Paper and Pulp Industry (TAPPI) standard test method T 544 cm-08 (at 90°C, 50% relative humidity, and for 16 days). The aging process helps distribute the sizing and enhances the penetration of pollutants into the surface's voids.

2.2.2 Femtosecond (fs) laser cleaning setup

A Yb:Glass chirped-pulse amplification system with a center wavelength of 1030 nm (Amplitude Systems, s-Pulse) was employed for cleaning experiments, delivering pulses with a duration of 550 femtoseconds at a 1 kHz repetition rate. The maximum pulse energy of the system is approximately 250 μJ. The paper samples were mounted on an XYZ motorized stage, and the laser beam was focused onto the surface using a cylindrical lens with a 25 mm focal length, as illustrated in Figure 2.4. The samples were translated under the focused beam along the short axis of the cylindrical lens at a constant speed of 1 mm/s, with a scanning pitch of 0.5 mm. These parameters were kept constant throughout the femtosecond laser cleaning experiments. The number of laser pulses delivered per spot was controlled by the defined scanning speed and pitch. Neither incubation effects nor cumulative pulse interactions were investigated in this study.

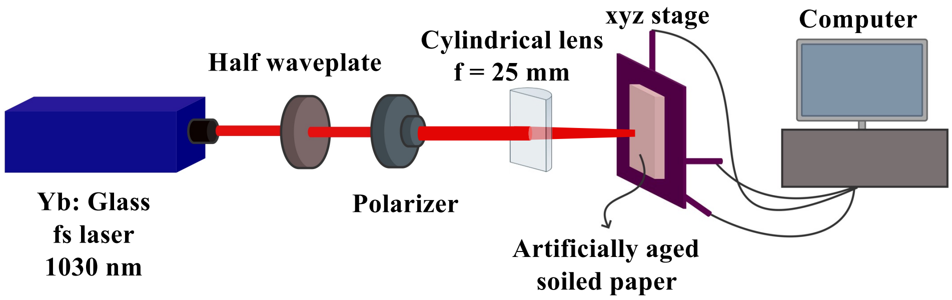


Figure 2.4. Femtosecond laser cleaning setup.

At a repetition rate of 1 kHz, the laser emits one pulse every 1 millisecond (1 ms = 0.001 s). However, each pulse lasts only 550 femtoseconds. To put this in perspective, the laser is active for just 550 femtoseconds, then rests for the remaining $9.9999999945 \times 10^{11}$ femtoseconds before the next pulse. This means the laser is "on" for only about 1 part in 1.8×10^9 of the total time.

This extremely short interaction, followed by a relatively long rest period, has a significant advantage: the material has plenty of time to cool down before the next pulse arrives. This avoids heat build-up, called thermal accumulation, which is one of the leading causes of unwanted damage in conventional laser processing. As a result, the cleaning is highly controlled and non-thermal, allowing safe treatment of delicate heritage materials like paper and historical coatings. This precise temporal control is a defining feature of femtosecond laser cleaning systems and one of the main reasons they are so effective for sensitive surfaces.

Six different fluence levels between 0.39 and 0.78 J/cm^2 were selected to evaluate cleaning performance across a range of conditions. These values were obtained by adjusting the laser output power while maintaining a constant beam footprint on the surface of artificially aged and contaminated paper samples. The fluence selection was based on preliminary experimental trials rather than direct references from the literature. The average laser power was varied between 100 and 200 mW in increments of 20 mW. Corresponding fluence values were calculated based on the directly observed beam area, measured from the line-focused footprint at the sample plane (beam height × beam waist). The beam waist here refers to the diameter at which the fluence drops to $1/\ e^2$ of its maximum value.

The selected fluence range, from 0.39 to 0.78 J/cm^2, was empirically established to achieve effective cleaning while avoiding damage to the paper substrate. Preliminary tests showed that laser powers above 200 mW led to observable alterations in the fiber structure of the paper under optical microscopy. Therefore, 200 mW was set as the upper operational limit. The corresponding maximum fluence of 0.78 J/cm^2 was close to the threshold where minor substrate degradation was first observed, allowing evaluation of the cleaning performance at near-critical intensity. Conversely, the lower fluence limit of 0.39 J/cm^2 represented the minimum energy required to produce discernible cleaning effects.

Although femtosecond lasers are capable of inducing nonlinear interactions such as multiphoton absorption (MPA), this mechanism was ruled out under our experimental conditions. The applied fluence range (0.39–0.78 J/cm^2) produced a peak intensity of approximately 2.84×10^{12} W/cm^2 at the beam waist. However, due to the strong attenuation by the contamination and sizing layers, the effective peak intensity reaching the cellulose substrate was significantly reduced.

For example, in the worst-case scenario (maximum fluence with graphite contamination and alum–egg sizing), the graphite layer absorbed approximately 97% of the incident energy, and the alum–egg sizing transmitted 95% of the remainder. The resulting peak intensity at the substrate surface was estimated to be around 8.08×10^{10} W/cm^2, well below the commonly cited threshold of 1×10^{11} W/cm^2 for two-photon absorption. Even in the case of a 50:50 graphite–kaolinite contamination layer (transmittance $\approx$ 51.5%), the final intensity at the cellulose surface reached only 1.39×10^{12} W/cm^2. Despite this higher intensity, neither damage to the substrate nor complete contaminant removal was observed, indicating that the interaction regime remained dominated by linear absorption.

These calculations indicate that the peak intensities transmitted through the contamination and sizing layers remained below or only slightly above the commonly reported thresholds for nonlinear processes such as two-photon absorption (typically cited as 1×10^{11} to 1×10^{13} W/cm^2). Nevertheless, no morphological damage or photochemical degradation was observed. This is further supported by colorimetric analysis from our previous study [68] on the same sample set, where all Δb^* values remained negative. For instance, Δb^* ranged between -4 and -8 for sized samples and from -2 to -11 for unsized ones, even at maximum fluence levels. Since negative Δb^* indicates a shift toward blue and away from yellowing, these results confirm that no photochemical degradation occurred, thus supporting the conclusion that the laser–material interaction was governed primarily by linear or near-threshold mechanisms rather than intense nonlinear effects.

Furthermore, cellulose has a wide band gap ($\sim$5.5 eV), and even two-photon absorption at 1030 nm (1.2 eV/photon) yields only $\sim$2.4 eV, which is insufficient to trigger electronic transitions. Reaching the conduction band would require at least a four-photon process, whose probability scales with the fourth power of intensity. Since the applied fluences are orders of magnitude below this requirement, nonlinear excitation can be confidently excluded. This conclusion is further supported by the absence of nonlinear damage features in both microscop images and ATR-FTIR analyses.

A total of six cleaned areas, each measuring 15 × 15 mm^2, were generated on the paper surface. These were produced using the cylindrical lens focused along the x-axis, which enabled sharply defined square cleaning patterns during scanning.

2.2.3 Scanning Electron Microscopy

A Tescan Vega3 LMH Scanning Electron Microscope was employed to examine the surface morphology of the artificially aged and contaminated samples. For these analyses, an acceleration voltage of 10 kV was used for samples covered with graphite and kaolinite, with the working distance set to approximately 17 mm.

2.2.4 Stylus Profilometry

Surface roughness measurements of the artificially prepared samples were conducted using a KLA Tencor P-7 Stylus Profilometer. The scan length was set at 1000 μm, with a scan speed of 50 μm/second and a sampling rate of 200 Hz. To minimize potential damage to the surface, the applied force was kept at 1.5 mN, which is at the lower limit of stylus tracking forces.

2.2.5 Attenuated Total Reflectance (ATR) Spectroscopy

ATR spectroscopy was performed using a Thermo Fisher Nicolet™ iS50 FTIR Spectrometer equipped with an Attenuated Total Reflectance mode. The system features a diamond crystal with a 45° incident IR beam and a single internal reflection, ensuring consistent contact between the diamond and the sample surface. The active area measured is only a few mm^2. Absorbance spectra were recorded in the 4000–400 cm^{-1} range, covering the mid-infrared and "fingerprint" regions, with a spectral resolution of 4 cm^{-1}. Each spectrum was averaged from 64 scans. The background spectrum was taken in the air and used for numerical corrections.

2.2.6 Statistical Analysis by Analysis of Variance (ANOVA)

To assess whether the observed differences in cleaning behavior across experimental conditions were statistically meaningful, Analysis of Variance (ANOVA) was employed. The objective was first to determine whether cleaning performance was correlated with material-related parameters, such as the type of sizing, the nature of contamination, and the presence or absence of burnishing. If such correlations were present, the extent to which each factor contributed to the overall variance in cleaning efficiency was then quantified.

This analysis was considered essential due to the complex interactions between laser irradiation and the physical-chemical properties of historical paper surfaces. While visual and spectroscopic trends provide qualitative insight into the cleaning process, they do not always constitute conclusive evidence. By using ANOVA, a structured evaluation of individual factors and their interactions was made possible. The statistical metrics obtained, including significance values and effect sizes (η^2), enabled a more rigorous comparison across treatment conditions and provided a quantitative foundation for optimizing laser cleaning strategies.

2.2.7 Data Collection and Processing

Surface Roughness Analysis: Surface roughness measurements using the profilometer were conducted in triplicate for each contamination type across different sample surfaces. The data were averaged to produce a representative assessment of surface contamination and its effects.

Scanning Electron Microscopy: For scanning electron microscopy, measurements were taken at a 45° angle with a magnification of around 5000x. To avoid sample charging, all cleaned and aged samples, including those contaminated with graphite and kaolinite, were coated with a 4 nm layer of Pt/Pd alloy using a sputter coater before imaging.

ATR-FTIR Measurements and Peak Selection: ATR spectroscopy was conducted at three points (upper, middle, and lower areas, each 5 mm apart) within the cleaned zones of each sample (marked by red dots in Figure 2.5). Repeated measurements enabled averaging and error estimation. Error bars in the figures reflect uncertainties derived from these three data points. ATR corrections (assuming a refractive index of 1.50 for the sample) and baseline corrections were applied to all spectra collected.

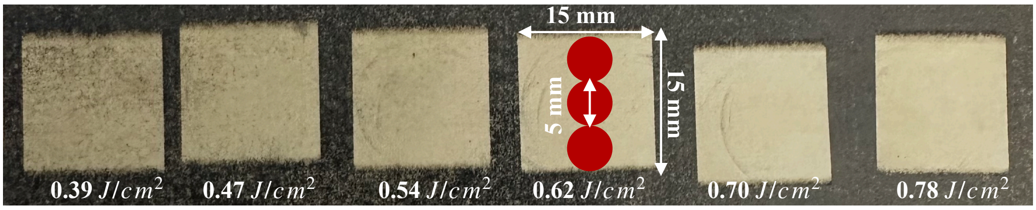


Figure 2.5. Cleaned regions of artificially soiled and aged paper (alum-egg sizing, graphite+kaolinite contamination). Square areas were cleaned with a fs laser at 1030 nm. Red spots represent the data collection points for ATR measurements. Progressive cleaning success with increasing laser fluence is visible; however, it is hardly quantifiable.

The analysis focused on spectral regions where significant changes were observed due to the cleaning process compared to the reference spectra. These selected regions correspond to peaks associated either with the contamination (such as kaolinite, when present) or the substrate (cellulose, corn starch, and wheat starch in the case of graphite-only contamination). The identified peaks reflect specific features in the corrected spectra, corresponding to the characteristic absorption bands of the materials in the samples. It was expected that the intensity of the absorption bands related to contaminants (as discussed below) would decrease while the intensity of the substrate would increase with further cleaning. Contamination intensity refers to the strength of the spectral signal from contaminant materials like kaolinite or graphite on the paper samples. In contrast, substrate intensity indicates the signal strength from the underlying materials, such as cellulose or the starches used for sizing. The overlapping spectral features of cellulose, polysaccharide-based sizings, and kaolinite were carefully accounted for when selecting these peaks (regions). Table 2.1 outlines the peak intervals for each paper treatment case.

Table 2.1. Chosen intervals for each paper treatment.

sizing	**contamination type**	**intervals** $(\mathbf{cm}^{-1})$	
alum – egg	graphite	$\sim 850 - 1200$	cellulose
alum – egg	graphite + kaolinite	$\sim 500 - 550$	kaolinite
no sizing	graphite	$\sim 800 - 1200$	cellulose
no sizing	graphite + kaolinite	$\sim 500 - 550$	kaolinite
corn starch	graphite	$\sim 500 - 550$	corn starch
corn starch	graphite + kaolinite	$\sim 3600 - 3650$	kaolinite
wheat starch	graphite	$\sim 500 - 550$	wheat starch
wheat starch	graphite + kaolinite	$\sim 3600 - 3650$	kaolinite

The area under each peak was calculated through numerical integration over the ranges specified in Table 2.1 , which is directly proportional to the concentration of the respective chemical bonds.

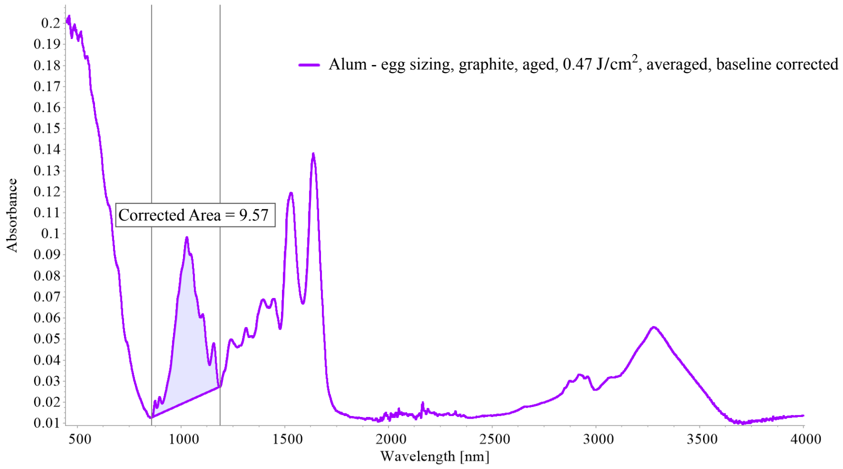


Figure 2.6. ATR-FTIR spectrum averaged from three measurements and baseline-corrected, showing the integrated absorbance peak of cellulose (shaded area) used for quantitative analysis. The 'Corrected Area' under the peak corresponds to this shaded region. The sample is alum–egg sized paper contaminated with graphite, aged, and cleaned at 0.47 J/cm^2.

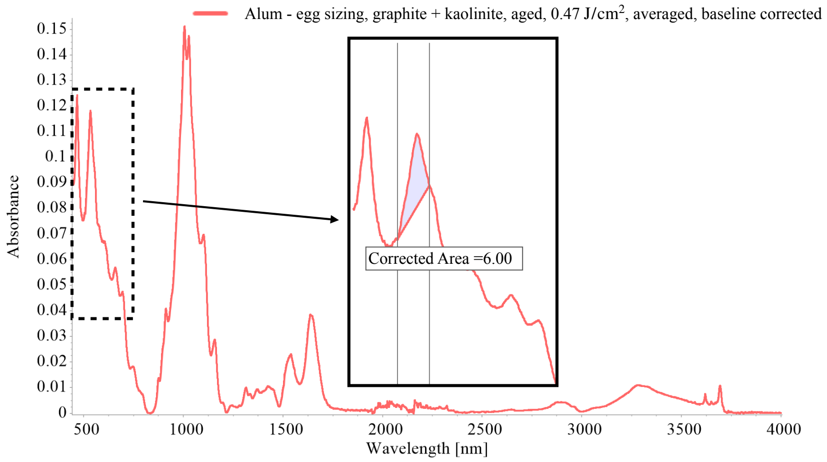


Figure 2.7. ATR-FTIR spectrum averaged from three measurements and baseline-corrected, showing the integrated absorbance peak of kaolinite (shaded area) used for quantitative analysis. The 'Corrected Area' under the peak corresponds to this shaded region. The sample is alum–egg sized paper contaminated with graphite + kaolinite mixture, aged, and cleaned at 0.47 J/cm^2.

Figure 2.6 and 2.7 illustrate examples of this process, showing the numerically integrated peaks for cellulose and kaolinite, respectively. The shaded areas represent the spectral regions over which integration was performed.

After determining the area under the peaks from repeated measurements, the standard deviation was calculated to evaluate the variability in the data, followed by the calculation of the Standard Error of the Mean to estimate the precision of the mean values. The corresponding graphs displayed these Standard Error of the Mean values as error bars.

To evaluate the quality of laser cleaning and quantify its effectiveness, we analyzed the area under the peaks in ATR spectra, which varies with laser fluence. As fluence increases, the area corresponding to the substrate concentration is expected to rise, then approach a plateau, and eventually begin to decrease. This plateau does not reflect detector saturation due to instrumental settings, but rather the physical limit at which additional contaminant removal no longer increases the detectable substrate signal. Beyond this point, further increases in fluence cause the laser to affect the substrate itself, leading to signal reduction and material degradation. Therefore, in practice, fluences should be limited to the range where the substrate signal approaches its plateau, beyond which additional cleaning does not offer benefits and may harm the substrate.

Data Normalization and Variability Sources for ANOVA: The cleaning data, derived from ATR-FTIR peak area measurements across six laser fluence levels (0.39 to 0.78 J/cm^2), were subject to several potential sources of variability that could affect the interpretation of laser cleaning efficiency. Three primary sources were identified:

1. **Variations across cleaned regions**, such as differences in contamination distribution, sizing thickness, and burnishing pressure between square and circular laser-treated areas.
2. **Manual inconsistencies** in the laser cleaning process, including variations in beam focus, laser stability, and sample positioning.
3. **Differences in peak selection** between substrate and contaminant signals exhibit opposite trends (increase vs. decrease) as cleaning progresses.

A preprocessing strategy was applied to minimize the influence of these factors. First, the integrated ATR-FTIR peak area at each fluence was normalized to the corresponding uncleaned reference value, correcting for baseline differences caused by preparation inconsistencies (points 1 and 3). Second, a **linear fit** was performed on each normalized dataset (fluence vs. absorbance), and the absolute value of the slope was extracted as a metric for cleaning responsiveness. Using the absolute value eliminated the sign-related bias introduced by inverse trends in substrate and contaminant signals (point 3).

However, it should be noted that using a linear fit is not physically rigorous, since it assumes an unlimited improvement of the spectral signal with increasing fluence. This is unrealistic in most material systems, where a saturation point is eventually reached. Moreover, a linear slope does not indicate whether cleaning was successful or complete; it only reflects how the sample responds to increased fluence. In that sense, the slope acts as a comparative index of responsiveness across different sample types, rather than a definitive indicator of cleaning efficiency. This limitation is later addressed through a more physically grounded saturation model, which captures the asymptotic behavior of contaminant removal and allows for quantifying both maximum achievable cleaning and the fluence level at which it occurs.

To account for each linear fit's uncertainty, the slope's standard error was calculated and incorporated into the dataset. Each slope value was expanded into a triplet:

- slope,
- slope $+\sigma$,
- slope $-\sigma$,

where σ is the standard error. This expansion allowed within-group variance to be estimated during statistical analysis and addressed variability linked to measurement noise or manual operation (point 2).

Due to the limited sample size and constraints in a fully factorial design, balanced data subsets were created for ANOVA evaluation. These subsets isolated the effects of specific independent variables, such as sizing type, contamination composition, and burnishing. Prior to conducting ANOVA, each subset was tested for normality using the Shapiro-Wilk test and for homogeneity of variances using Levene's test to validate the assumptions of the model.

Nonlinear Modeling with Saturation and Decay Fits: The goal was to model the laser fluence at which substrate concentration saturates and to assess how closely the experimental data matched this model. To quantify cleaning efficiency, we employed an exponential saturation model, represented by equation (2.1):

$$A(x) = a \cdot \left(1 - \exp\left(-\frac{x-b}{c}\right)\right) + d \tag{2.1}$$

Where A(x) is the area under the selected peak (represented as 'corrected area' on the Figures 2.6) at laser fluence x, a represents the maximum possible area, b corresponds to the minimum fluence required to initiate a detectable cleaning effect, it is associated with the ablation threshold, c is the saturation constant, and d is the baseline offset.

This model assumes that the area approaches a maximum value a beyond which additional fluence does not significantly enhance cleaning. The fluence offset b defines the minimum fluence required to initiate cleaning. To clarify the physical meaning of the parameter a in the exponential saturation model, examining the fluence value at which no visible cleaning effect occurs, i.e., when the signal is essentially zero, is helpful. Starting from the model equation (2.1) and solving for $A(x) = 0$, one obtains the following expression:

$$x = b - c \cdot \ln\left(1 + \frac{d}{a}\right), \tag{2.2}$$

which defines the minimum fluence x required to observe a nonzero cleaning signal. This threshold depends on the saturation amplitude a, the steepness constant c, and the baseline offset d. In cases where d < a, the correction term becomes negligible and the threshold simplifies to $x \approx b$. Thus, the parameter b effectively captures the onset of the cleaning process and provides a practical indicator of the minimum fluence required to initiate contaminant removal. While the saturation constant c reflects how quickly saturation is reached. The baseline offset d accounts for any initial signal before cleaning. We performed non-linear least squares fitting using MATLAB, with initial parameter estimates based on preliminary observations. The fitting process adjusted the parameters iteratively to minimize the difference between the experimental data and the model, and the fit quality was assessed using metrics like the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). The fitted parameters offered insights into the cleaning efficiency and highlighted the optimal fluence range for effective results.

To quantify residual contamination, we used an exponential decay model, represented by equation (2.3):

$$A(x) = a \cdot \exp\left(-\frac{x - b}{c}\right) + d \tag{2.3}$$

Where A(x) represents the area under the selected peak (represented as 'corrected area' on the Figures 2.7) at laser fluence x, a is the initial area indicating contamination levels, b is the fluence offset (which is also positive), c is the decay constant, and d is the baseline offset.

This model assumes that as fluence increases, the area corresponding to contaminant concentration decreases but never fully reaches zero, as some residual contamination is expected to remain in surface voids. This behavior reflects a physical limitation in laser cleaning. With each successive laser pulse, a portion of the contaminants may become embedded deeper into the substrate, particularly into fibrous microstructures, making complete removal increasingly difficult. Attempting to remove such deeply entrapped residues would require higher fluences or prolonged exposure, which risks irreversible damage to the material's microstructure. We applied the same non-linear least squares fitting method in MATLAB, using initial parameter estimates. The fitting process minimized the difference between the observed data and the model, with fit quality assessed using the same metrics as the saturation model. The fitted parameters provided estimates of residual contamination, helping evaluate the effectiveness of different laser cleaning conditions. Both models offered valuable insights into the cleaning process and enabled us to quantify and compare the efficacy of various laser cleaning conditions.

2.3 Results and Discussion

2.3.1 Surface Morphology by Scanning Microscope

Corn and wheat starch sizings are fine, and white powders are commonly dispersed in water to create non-Newtonian suspensions or pastes, which are then applied to paper surfaces. Once dried, these sizings form thin, translucent films that improve the paper's strength, smoothness, and resistance to ink penetration. In contrast, alum-egg sizing, a mixture of aluminum potassium sulfate (alum) and egg white (albumen), has a slightly viscous or gel-like consistency.

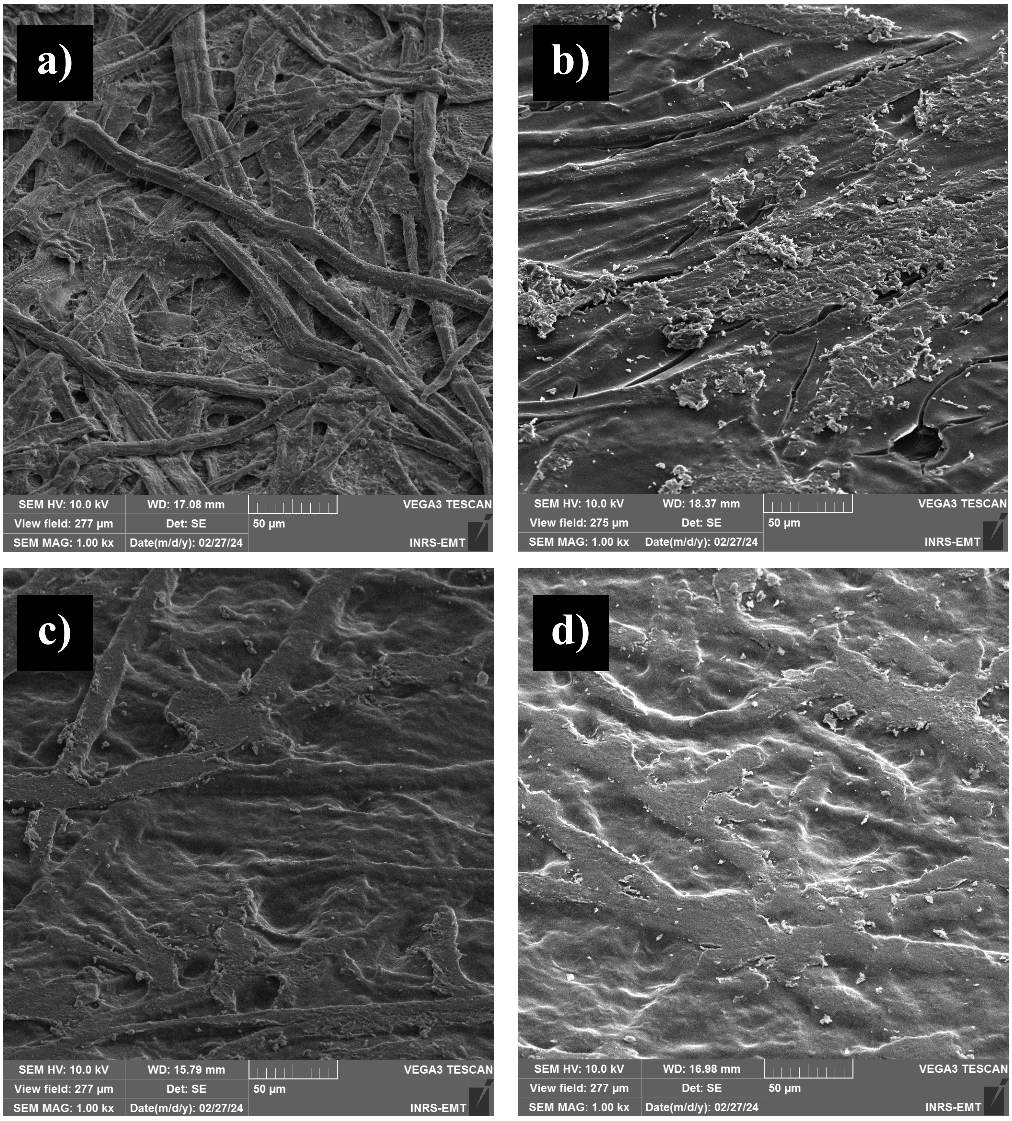


Figure 2.8. Scanning electron microscopy images of paper samples: a) no sizing, b) alum-egg sizing, c) corn starch sizing, and d) wheat starch sizing. Images were taken after aging but before contaminating and cleaning.

The scanning electron microscopy images in Figure 2.8 reveal that samples a) without sizing show more irregularities and gaps between the paper fibers compared to those treated with b) alum-egg sizing, c) corn starch sizing, and d) wheat starch sizing. Due to its gel-like consistency, alum-egg sizing penetrates and adheres to the paper fibers more effectively than other sizings, resulting in a smoother surface texture. The scanning electron microscopy images clearly show a reduction in surface irregularities and an improvement in smoothness for paper treated with sizing.

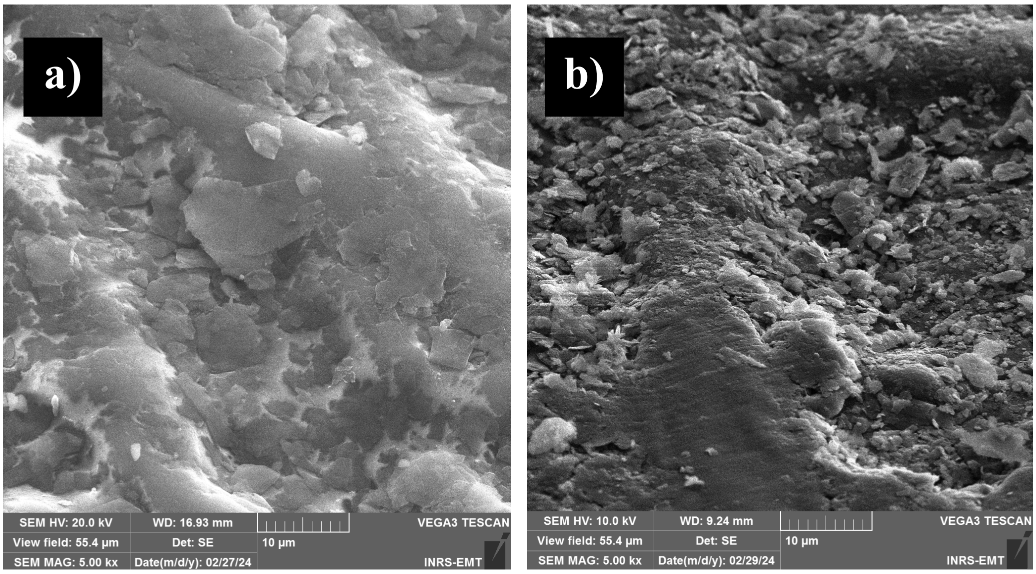


Figure 2.9. Scanning electron microscopy images of wheat starch sizing samples contaminated by a) graphite and b) graphite + kaolinite mixture.

Figure 2.9 contains scanning electron microscopy images of wheat starch-sized samples contaminated with a) graphite and b) graphite + kaolinite mixture. These images show that the graphite + kaolinite mixture forms a more uniform and continuous layer between the fibers compared to graphite alone, indicating better adhesion and more complete coverage, effectively filling the gaps between the paper fibers. This characteristic was consistently observed across all samples.

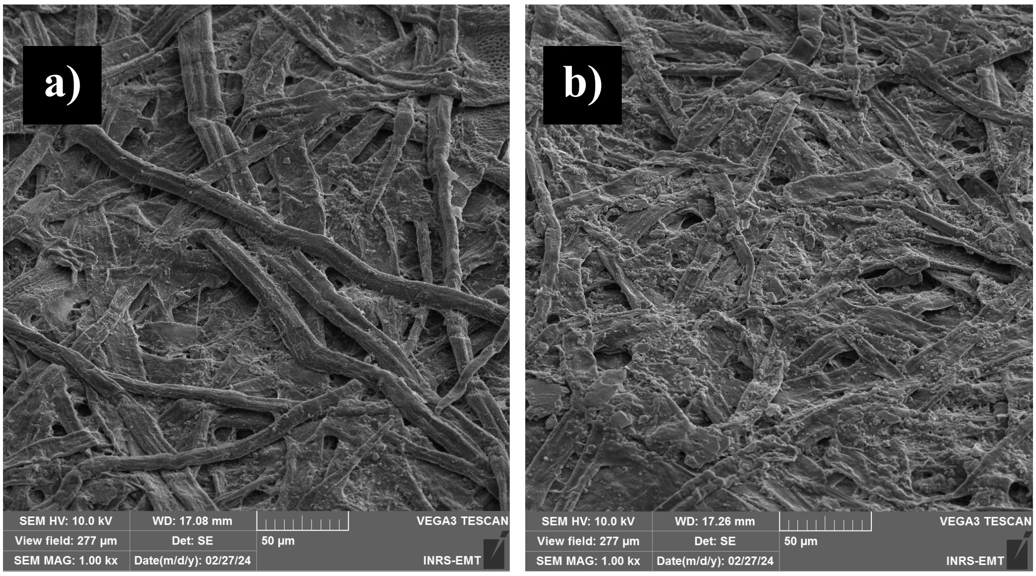


Figure 2.10. Scanning electron microscopy images of paper surfaces before contamination: a) unburnished and b) burnished sample.

Figure 2.10 presents scanning electron microscopy (SEM) images comparing a) unburnished and b) burnished paper surfaces prior to contamination. The burnishing process significantly alters the fiber morphology, visibly compacting the fibrous network and reducing the surface porosity. This densification not only affects the visual appearance but also plays a critical role in contaminant interaction. By compressing the surface structure, burnishing breaks down particulate contaminants and pushes them deeper into the inter-fiber voids.

This behavior is especially evident when considering the graphite particles applied in this study. According to manufacturer specifications (AF spezial, Kropfmühl), the graphite powder has a D90 $\leq 25\ \mu$m, meaning 90% of the particles are smaller than 25 μm. However, as shown in the SEM images in Figure 2.9, many graphite particles on burnished samples appear significantly smaller than 10 μm, as evident from the embedded scale bars. This apparent discrepancy is not a cause for concern; rather, it reflects the mechanical disintegration of the particles during the burnishing process. The applied mechanical pressure effectively crushes and spreads the particles, creating a thin, uniform contamination layer that penetrates the compacted fiber network. This behavior has direct implications for laser cleaning, as it influences both the adhesion of the pollutants and the laser fluence required to remove them from the confined microstructure.

To further support this observation, Figure 2.11 compares the morphology of graphite contamination on a) unburnished and b) burnished paper surfaces. The burnished surface clearly exhibits more compacted graphite particles, many of which appear broken into sub-10 μm fragments and embedded more deeply into the fiber matrix. This visual evidence corroborates the mechanical crushing and redistribution described previously, reinforcing the idea that burnishing significantly alters the

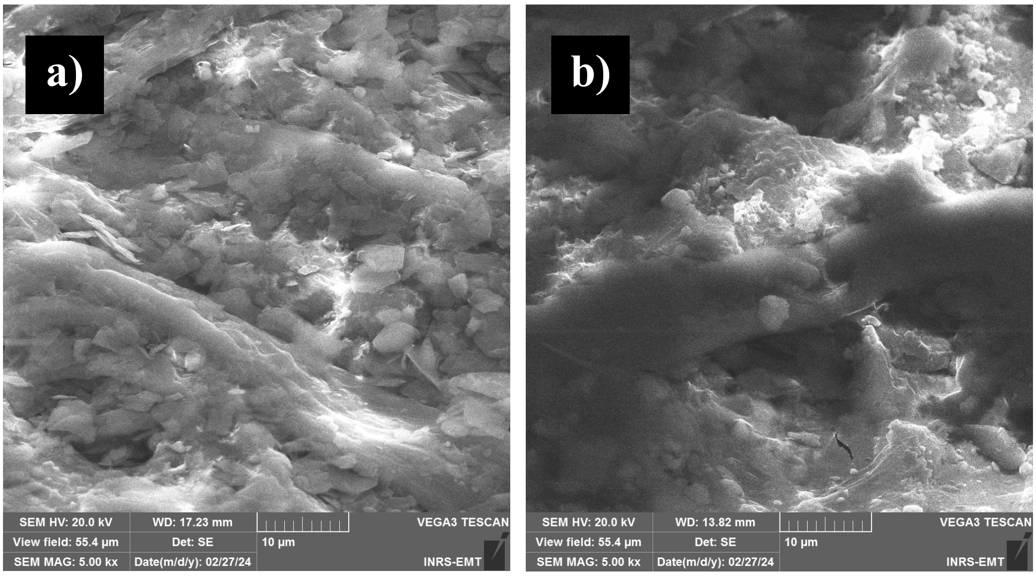


Figure 2.11. Scanning microscope images comparing a) unburnished and b) burnished paper samples contaminated with graphite.

contaminant layer structure and its interaction with laser energy.

This behavior has direct implications for laser cleaning, as it influences both the adhesion of the pollutants and the laser fluence required to remove them from the confined microstructure.

2.3.2 Surface Roughness

The profilometer data in Table 2.2 shows an overall trend of increasing surface roughness (within the margin of error) among samples contaminated with graphite compared to those contaminated with the graphite + kaolinite mixture, with the exception of the alum-egg-sized sample. This particular sample exhibits higher surface roughness in the graphite + kaolinite contamination case, highlighting a distinct behavior in comparison to the other samples.

Table 2.2. Average roughness values of the different paper treatments.

samples	graphite	graphite + kaolinite
no sizing / no burnishing	$1.9 \pm 0.6\,\mu$m	$1.7 \pm 0.3\,\mu$m
no sizing / burnishing	$1.9 \pm 0.9\,\mu$m	$1.1 \pm 0.5\,\mu$m
alum – egg sizing	$0.6 \pm 0.2\,\mu$m	$1.2 \pm 0.6\,\mu$m
wheat starch sizing	$1.6 \pm 0.8\,\mu$m	$1.5 \pm 0.6\,\mu$m
corn starch sizing	$1.9 \pm 1.0\,\mu$m	$1.6 \pm 0.3\,\mu$m

The higher surface roughness values in samples contaminated with graphite indicate a coarser texture that hinders the infiltration of contaminants into the paper's pores.

On the other hand, samples contaminated with the graphite + kaolinite mixture show lower roughness values, implying smoother surfaces and better contaminant penetration into the paper's pores. These findings suggest that the presence of kaolinite with graphite may facilitate deeper penetration of contaminants into the paper substrate. This deeper infiltration significantly affects the effectiveness of cleaning methods, including laser cleaning, as it becomes more difficult to remove contaminants embedded within the substrate due to reduced laser interaction. Scanning electron microscopy images in Figure 2.9 support these conclusions, showing that the samples contaminated with the graphite + kaolinite mixture have reduced surface irregularities and smoother surfaces compared to those contaminated with graphite alone.

A comparison between the no sizing/no burnishing sample and the no sizing/burnishing sample reveals that the no sizing/no burnishing sample exhibits the highest surface roughness, regardless of the type of contamination. The absence of sizing, which typically fills surface pores, and the lack of burnishing, which compacts the surface, contribute to this outcome. Table 2.2 outlines surface roughness by sizing type. Corn starch sizing consistently results in the highest roughness across all contamination types, followed by wheat starch and alum-egg sizing. While roughness measurements provide insight into the effects of burnishing and different sizings, the influence of the presence or absence of sizing or the contamination type on surface roughness remains unclear. These observations highlight the complex relationship between sizing, contaminant behavior, and surface roughness, emphasizing the challenges and effectiveness of laser cleaning in removing contaminants from paper substrates.

The surface morphology of these samples after femtosecond laser cleaning has been demonstrated in a previous study, [68], where microscopy images confirmed that no damage occurred under the laser parameters used (1030 nm, 300 fs, 1 kHz, 100–200 mW). As the same samples were used in this study, the findings regarding surface preservation remain valid.

2.3.3 ANOVA-Based Assessment of Sizing, Contamination, and Burnishing Effects

Subset I: In Subset I, only burnished paper samples, cleaned using femtosecond laser pulses, are analyzed to isolate the effects of sizing and contamination under consistent surface conditions. In Table 2.3, a two-way ANOVA revealed statistically significant main effects for both sizing type and contamination composition ($p < 0.001$ for both), with very large effect sizes ($\eta_p^2 = 0.95$ and 0.97, respectively). A significant interaction effect between the two variables was also observed ($p < 0.001$, $\eta_p^2 = 0.91$), indicating that the impact of contamination on cleaning efficacy depends on the sizing applied to the paper.

Table 2.3. Two-way ANOVA results for Subset I (burnished samples).

Variables	**Levels**	
Sizing Type	No Sizing, Alum–Egg Sizing, Corn Starch Sizing, Wheat Starch Sizing	
Contamination Type	Graphite, Graphite + Kaolinite	
Factor Effect	**Significance**	**Effect Size (η_p^2)**
Main Effects:		
Sizing	<0.001	0.95
Contamination	<0.001	0.97
Two-Factor Interaction:		
Sizing × Contamination	<0.001	0.91

These statistical results are well supported by surface-level observations. Scanning electron microscopy (SEM) images showed that corn and wheat starch sizingsform smooth, uniform surface films that reduce porosity and fiber exposure. This morphology likely promotes more consistent contaminant distribution, which in turn may influence how the surface interacts with the laser during cleaning. In contrast, non-sized papers exhibited a more porous and irregular morphology, as confirmed by both SEM and profilometry, which recorded the highest surface roughness values among all samples. This elevated roughness arises from the intrinsic structure of the fibrous network, where the absence of a sizing layer leaves wider inter-fiber spaces. These voids facilitate deeper penetration of particulate contaminants, especially in the case of the graphite + kaolinite mixture, which tends to embed within the surface architecture rather than remaining as a uniform external coating.

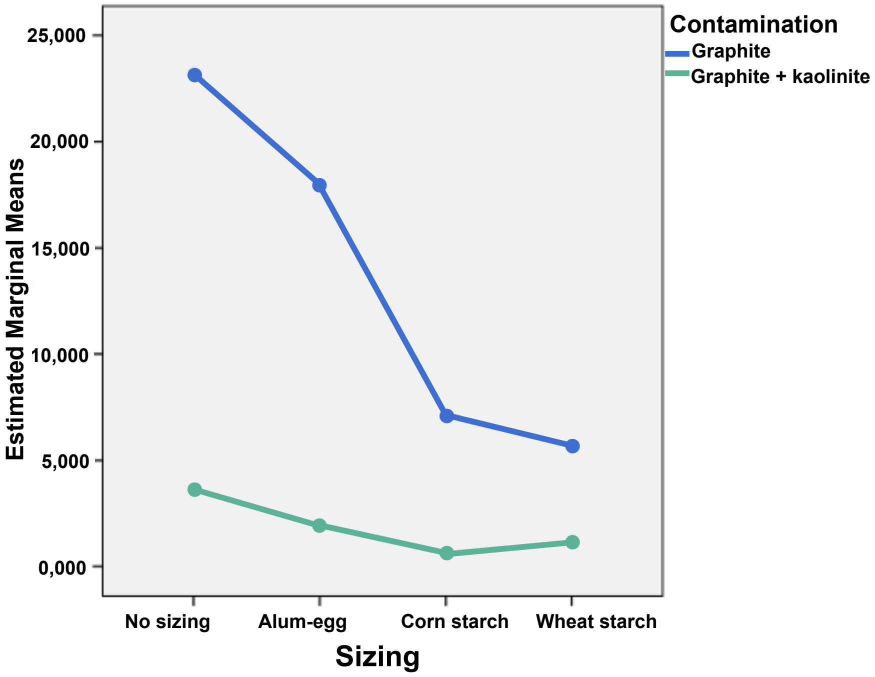


Figure 2.12. Estimated marginal means from two-way ANOVA for Subset I. Graphite and graphite + kaolinite contaminations are evaluated across three sizing types and a non-sized sample under fixed burnishing and fs-laser conditions.

The Estimated Marginal Means (EMM) plot in Figure 2.12 illustrates how different paper treatments statistically respond to increasing laser fluence, as modeled by the ANOVA framework. Higher EMM values reflect stronger predicted responsiveness to fluence variation, meaning that the cleaning outcome changes more sharply with increasing energy input. It is important to emphasize that these values indicate the amplitude of the modeled response rather than absolute contamination levels or cleaning success.

Graphite-contaminated samples consistently exhibit higher EMM values compared to those contaminated with graphite + kaolinite, reflecting their stronger predicted responsiveness to increasing laser fluence. This can primarily be attributed to graphite's higher optical absorbance, which allows for more efficient coupling with laser energy and leads to more pronounced material interaction. Among the graphite-contaminated groups, samples without sizing display the highest EMM values overall. In the absence of a sizing layer, there is no retentive or modulating barrier on the surface, which means that contaminants remain more exposed and unconfined. As a result, each laser pulse induces a sharper and more immediate change in the cleaning outcome. This leads to a steeper fluence–response trajectory, as captured by the statistical model. Starch-sized samples contaminated with graphite also demonstrate relatively high EMM values, although to a lesser extent. This is likely due to the presence of the sizing layer, which acts as a retentive and modulating medium, enhancing contaminant adhesion to the surface and thereby dampening the per-pulse effectiveness of laser exposure.

In contrast, samples contaminated with the graphite + kaolinite mixture show uniformly low EMM values across all sizing types. This reduced statistical response suggests that the kaolinite component, due to its lower optical absorbance and possibly deeper embedding, leads to limited interaction with laser energy, regardless of surface treatment.

Taken together, the ANOVA results, visual data, and morphological insights converge to show that both the type of sizing and the contaminant composition, and their interaction, play a critical role in determining the success of femtosecond laser cleaning on historical paper substrates.

Subset II: In Subset II, only non-sized paper samples are analyzed to isolate the effect of burnishing on laser cleaning performance. In Table 2.4, a two-way ANOVA revealed that both burnishing and contamination type significantly influenced cleaning efficiency ($p < 0.001$), with large effect sizes ($\eta_p^2 = 0.91$ and 0.98, respectively). Additionally, a strong interaction between these two factors was observed ($p < 0.001$, $\eta_p^2 = 0.79$), indicating that the influence of contamination depends on whether or not the surface was burnished prior to cleaning.

Table 2.4. Two-way ANOVA results for Subset II (non-sized samples with and without burnishing).

Variables	**Levels**	
Burnishing	Yes, No	
Contamination Type	Graphite, Graphite + Kaolinite	
Pulse Duration	fs laser	
Factor Effect	**Significance**	**Effect Size (η_p^2)**
Main Effects:		
Burnishing	<0.001	0.91
Contamination	<0.001	0.98
Two-Factor Interaction:		
Burnishing × Contamination	<0.001	0.79

These results are visually supported by the Estimated Marginal Means (EMM) plot in Figure 2.13, which shows that graphite-contaminated samples consistently exhibit higher cleaning responsiveness than samples treated with the graphite + kaolinite mixture. Burnished samples show significantly greater responsiveness than their unburnished counterparts, suggesting that burnishing enhances the sensitivity of the cleaning outcome to laser fluence.

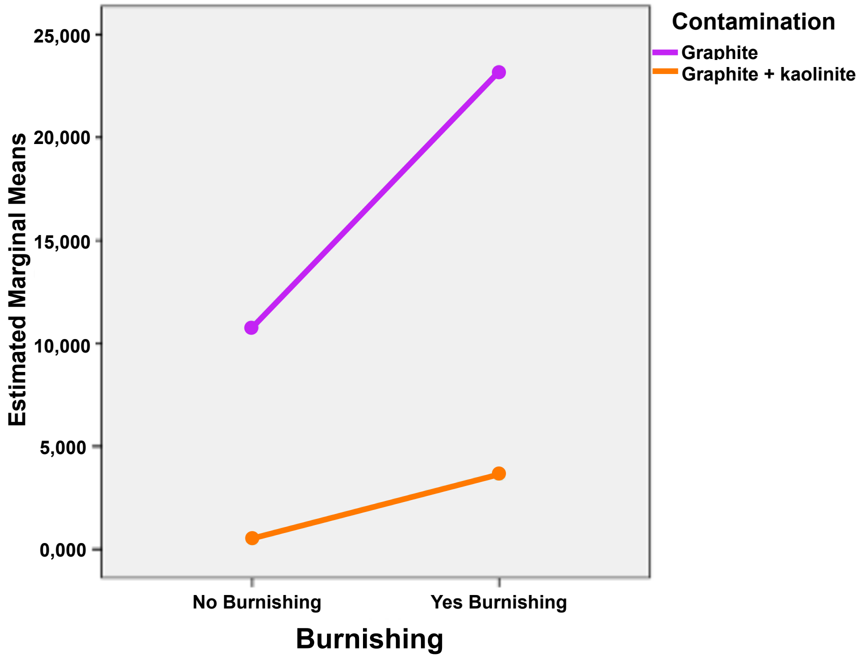


Figure 2.13. Estimated Marginal Means (EMM) plot for Subset II, showing the interaction between burnishing and contamination types.

Morphological observations obtained through scanning electron microscopy (SEM) provide a physical context for interpreting the increased responsiveness observed in burnished samples. SEM images (Figure 2.10) reveal that burnished surfaces are smoother and more compact, with densely packed fiber networks, while unburnished samples exhibit looser, more irregular structures with open fiber arrangements. This compaction resulting from burnishing enhances the uniformity and optical density of the surface, which in turn facilitates more effective laser-material interaction. Such conditions are favorable for energy absorption and ablation efficiency, especially under increasing fluence. These morphological characteristics likely contribute to the higher Estimated Marginal Means observed in burnished samples, reflecting greater sensitivity of the cleaning outcome to fluence variation.

Quantitative surface profilometry further supports these morphological findings. Burnished samples exhibited consistently lower roughness values, confirming a smoother and more uniform surface topography. This reduction in roughness aligns with the enhanced laser-material interaction suggested by SEM analysis. Smoother surfaces enable more consistent and direct energy coupling, which amplifies the cleaning response under varying fluence levels. In contrast, unburnished samples with higher surface roughness and open fiber structures may lead to more diffuse laser interactions, reducing the overall responsiveness captured in the EMM analysis.

Together, these results demonstrate that surface morphology and roughness modulate not only contaminant retention but also the system's sensitivity to laser parameters.

2.3.4 Quantifying Cleaning with Saturation Modeling

The following section presents the experimental results using a saturation model that defines the relationship between laser fluence and the recovered substrate signal or residual contamination. Although various mathematical models could be considered for fitting, we justify the use of our model based on the nature of the interaction between the laser and surface absorbers. As more material is removed, the interaction between the laser and the remaining absorbers diminishes. Since the total absorption coefficient is linearly related to the concentration of absorbers still present on the surface, we expect exponential relationships, which form the foundation of our numerical analysis in the upcoming section.

Figure 2.14 shows the fitting curve for the exponential saturation model applied to alum-egg sizing with graphite contamination during femtosecond laser cleaning. The results demonstrate that as fluence increases, the area under the peak also increases, eventually reaching a saturation point where further increases in fluence have little to no impact on the area.

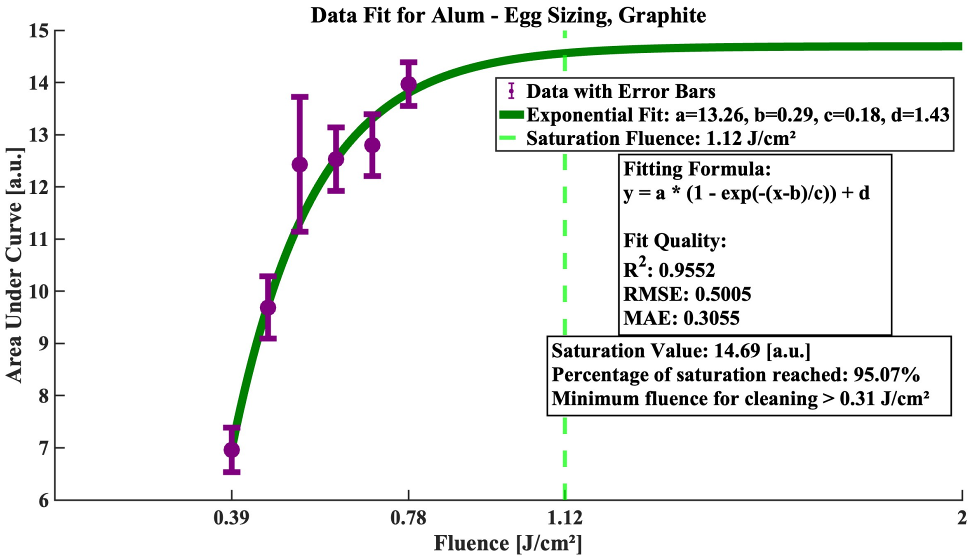


Figure 2.14. Exponential saturation model fit for alum-egg sizing with graphite contamination for fs laser cleaning. The green solid curve represents the fitted cleaning trajectory, and the vertical green dashed line marks the saturation fluence threshold (1.12 J/cm^2) determined from the model. Beyond this fluence value, the model extrapolates the cleaning response. However, prediction uncertainty increases as the curve extends beyond the range of measured data.

The graph illustrates how the signal corresponding to alum-egg sizing, which represents the relative concentration of the substrate in the ATR spectra, rises with increasing fluence until it reaches a saturation value of 14.69 a.u. at a saturation fluence of 1.12 J/cm^2. The fit quality metrics (R^2 = 0.9552, RMSE = 0.5005, MAE = 0.3055) reflect a high degree of accuracy between the model and the experimental data. An R^2 value close to 1 signifies an excellent fit, while the low RMSE (Root

Mean Square Error) and MAE (Mean Absolute Error) demonstrate minimal differences between the predicted and actual values. The model achieved 95.07% of the maximum possible area, underscoring its effectiveness in forecasting cleaning efficiency. In addition, we extrapolated to determine the fluence below which no visible cleaning effect occurs. This was done by calculating the fluence at which the area under the fitted curve approaches zero. During this process, constraints were applied to ensure that the fluence (x) remained real, non-negative, and below 0.39 J/cm^2. For this configuration, the threshold fluence was determined to be 0.31 J/cm^2. At fluence values below this, no cleaning effect should be expected.

In addition to the exponential saturation model fit for alum-egg sizing with graphite contamination, model fits for wheat starch sizing and corn starch sizing are shown in Figure 2.15 as well as for no sizing/burnishing and no sizing/no burnishing samples in Figure 2.16. Each figure illustrates the relationship between laser fluence and cleaning efficiency for different the sizings and the same contamination conditions, enabling direct comparisons.

Table 2.5 summarizes the fitted parameters for all samples, highlighting key metrics such as saturation fluence and ‘saturation reached %’ values.

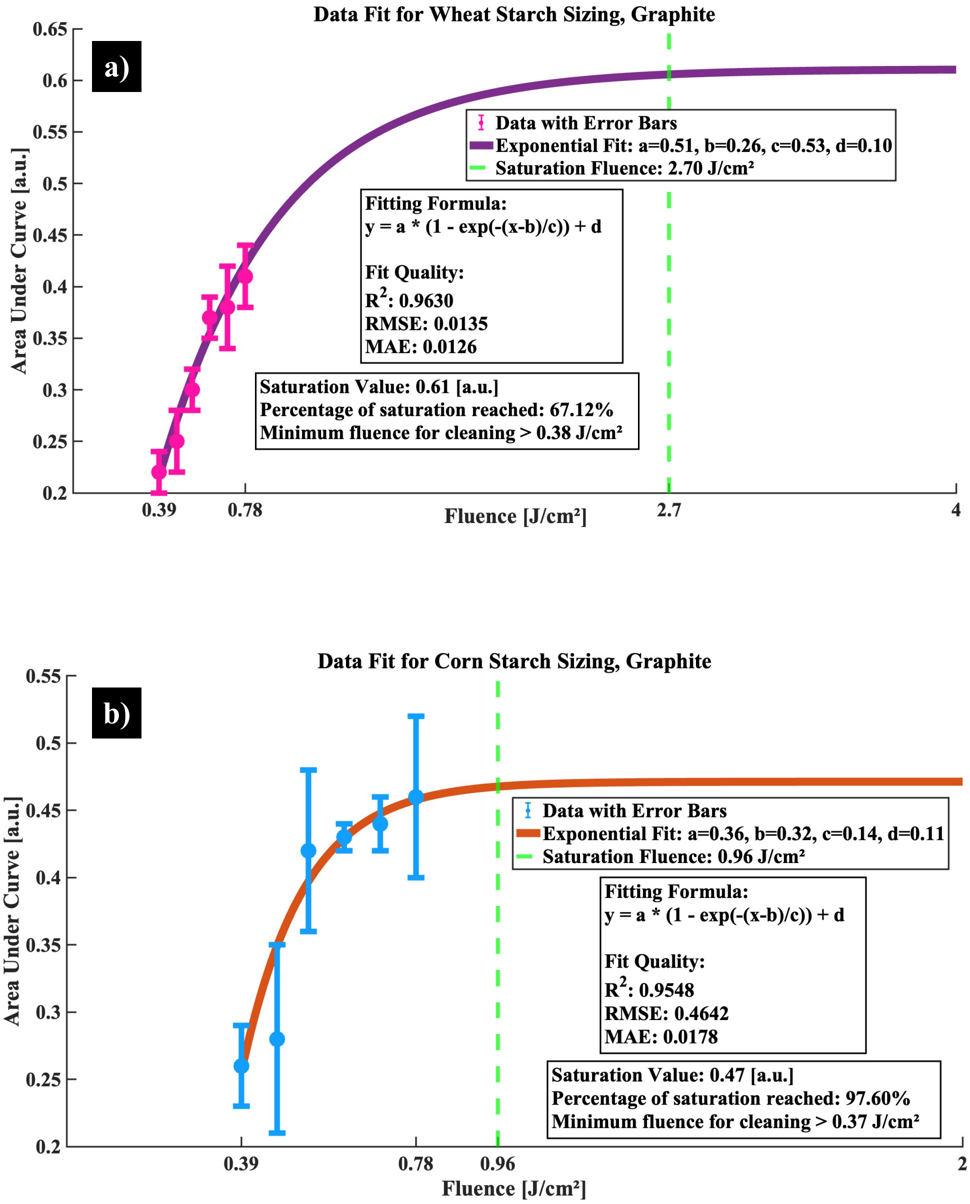


Figure 2.15. Exponential saturation model fit for a) wheat starch ve b) corn starch sizing with graphite contamination for fs laser cleaning.

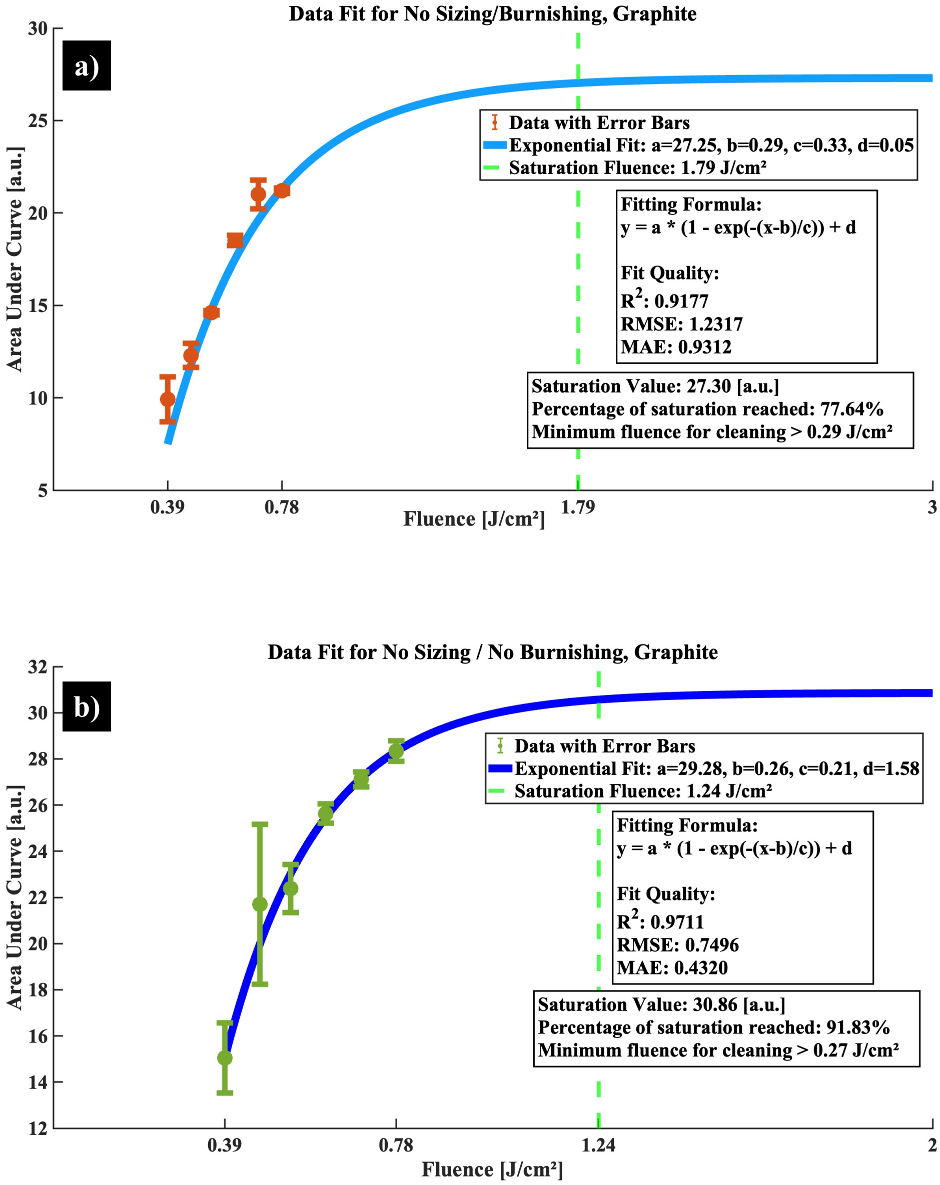


Figure 2.16. Exponential saturation model fit for a) no sizing/burnishing ve b) no sizing/no burnishing with graphite contamination for fs laser cleaning.

The results from the exponential decay model show that the area under the peak, representing contaminant concentration, decreased as fluence increased but never fully reached zero, indicating the presence of residual contamination. Figure 2.17 presents the fitting curve for the exponential decay model applied to corn starch sizing with graphite + kaolinite mixture contamination during femtosecond laser cleaning.

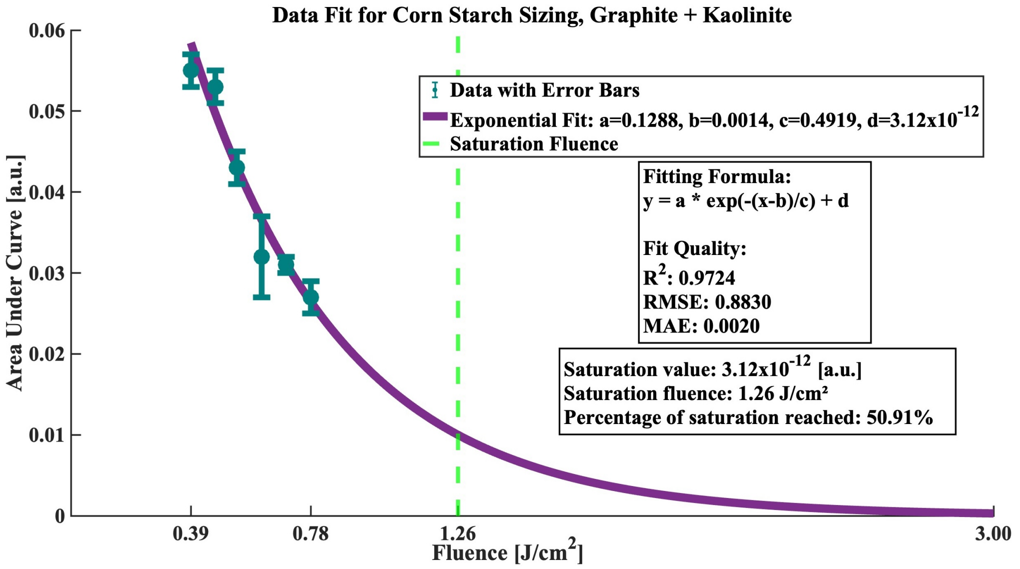


Figure 2.17. Exponential decay model fit for corn starch sizing with graphite + kaolinite mixture contamination for fs laser cleaning.

Figure 2.17 illustrates how the area under the curve decreases with increasing fluence, eventually approaching a saturation value of 3.12×10^{-12} a.u. at a saturation fluence of 1.26 J/cm^2. The fit quality metrics ($R^2 = 0.9724$, RMSE = 0.8830, MAE = 0.0020) indicate a strong correlation between the model and the experimental data, accurately capturing the decay process. However, the experimental data only reached 50.91% of the model's predicted saturation value. This suggests that complete removal of the contamination is unlikely unless extremely high laser fluences, beyond the cellulose destruction threshold, or multiple pulsing are employed. It's important to note that, with a laser repetition rate of 1 kHz, equating to 1000 pulses per second, multiple consecutive pulses could achieve a greater level of cleaning without significantly affecting the overall process time.

In addition to the exponential decay model fit for corn starch sizing with graphite + kaolinite contamination, the model fits for wheat starch sizing and alum egg sizing are shown in Figure 2.18 as well as for no sizing/burnishing and no sizing/no burnishing samples in Figure 2.19. Each figure illustrates the relationship between the laser fluence and cleaning efficiency for the different sizings and the same contamination conditions, enabling direct comparisons.

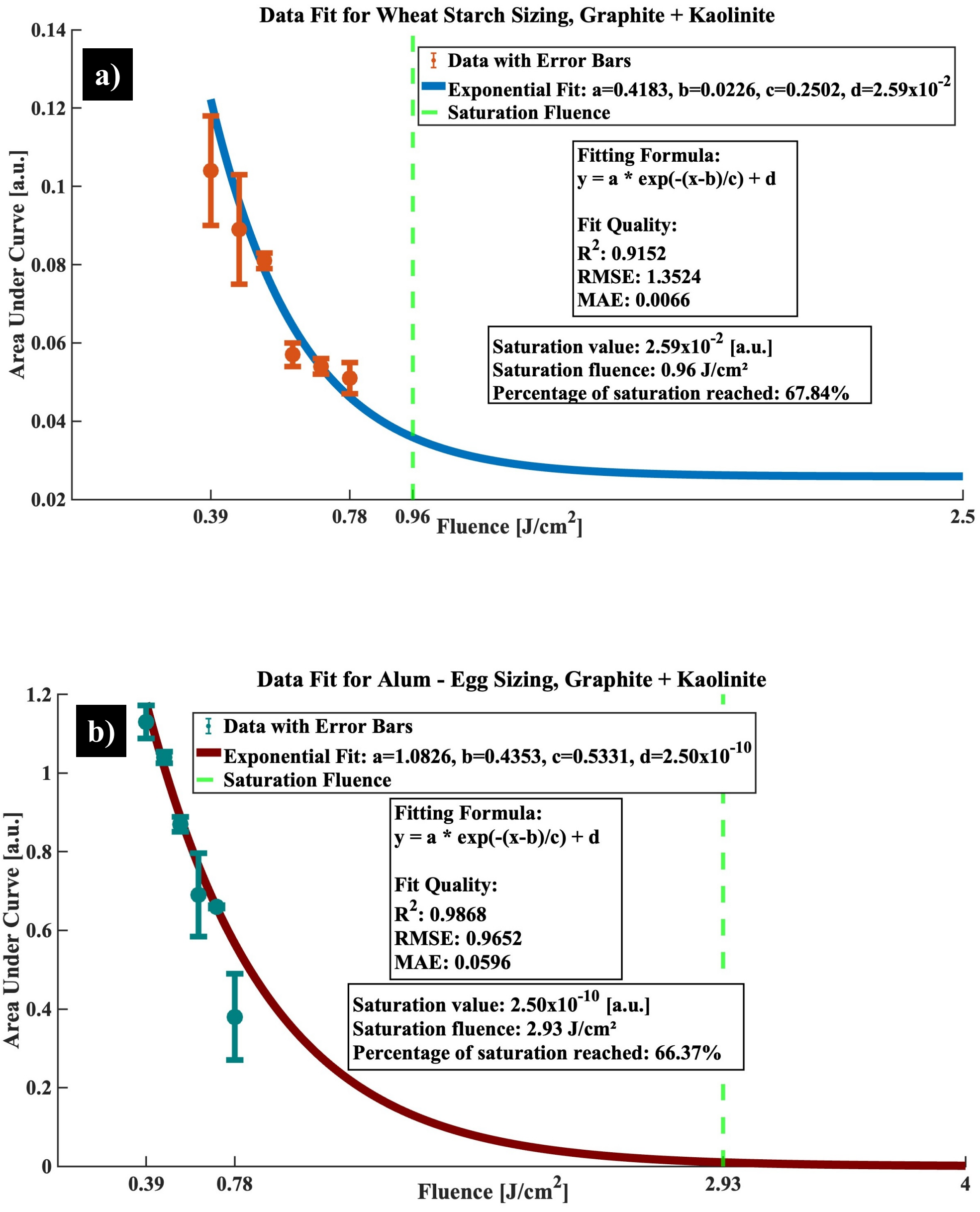
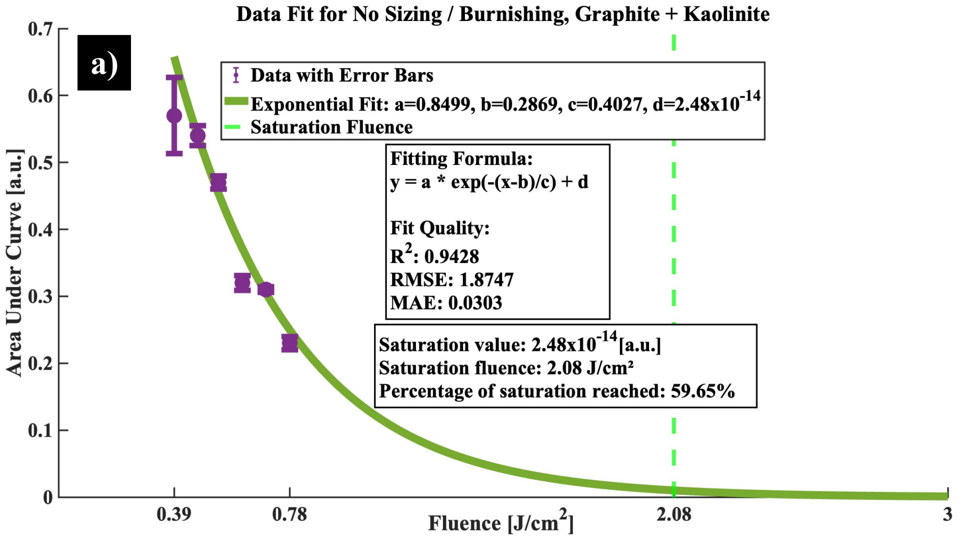


Figure 2.18. Exponential decay model fit for a) wheat starch sizing and b) alum - egg sizing with graphite + kaolinite contamination for fs laser cleaning.



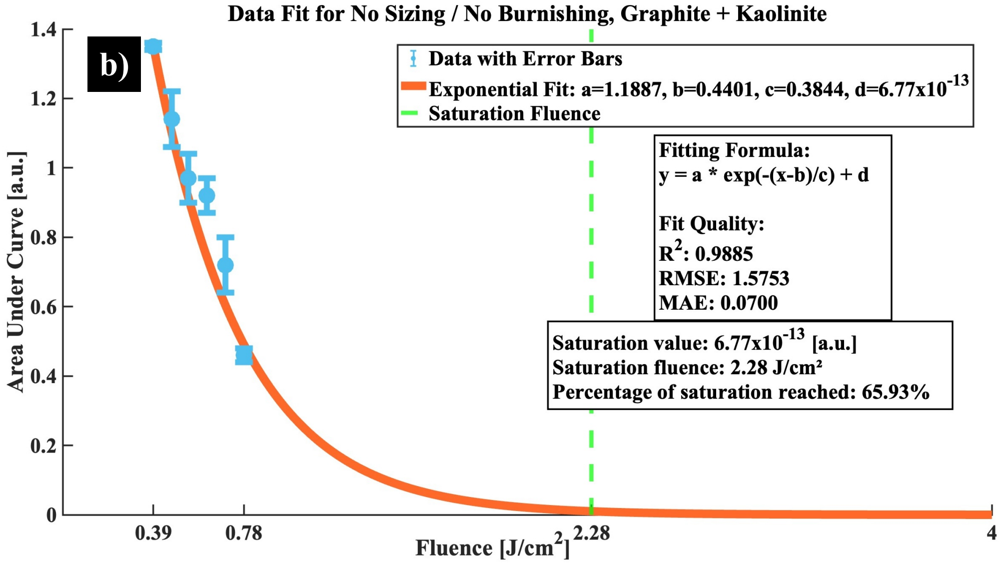


Figure 2.19. Exponential saturation model fit for a) no sizing/burnishing and b) no sizing/no burnishing with graphite + kaolinite contamination for fs laser cleaning.

The parameters in Table 2.5 and Table 2.6, such as “saturation fluence” and “saturation reached %,” are essential for understanding laser cleaning efficiency. “Saturation fluence” indicates the fluence level where the cleaning effect reaches its maximum, while the “saturation reached %” reflects how closely the experimental results approach this maximum, offering a quantitative measure of selective contaminant removal. Higher “saturation reached %” values indicate a more effective cleaning process, aligning with the concept of “quantitative selective cleaning,” where the goal is to maximize substrate exposure while minimizing potential damage.

Table 2.5. Fitted parameters for the exponential saturation model and saturation proximity.

Sample	**Saturation reached (%)**	**Saturation fluence [J/cm²]**	**Minimum fluence for cleaning [J/cm²]**	**R²**	**RMSE**	**MAE**
Wheat starch	67.12	2.70	> 0.38	0.9630	0.0135	0.0126
Alum - egg	95.07	1.12	> 0.31	0.9552	0.5005	0.3055
Corn starch	97.60	0.96	> 0.37	0.9548	0.4642	0.0178
No sizing/Burnishing	77.64	1.79	> 0.29	0.9177	1.2317	0.9312
No sizing/No burnishing	91.83	1.24	> 0.27	0.9711	0.7496	0.4320

The exponential saturation and exponential decay models were applied to evaluate how the specially prepared paper samples responded to laser treatment. According to the results presented in Table 2.5, the highest saturation levels were observed in the corn starch and alum–egg sized papers, reaching 97.60% and 95.07% respectively. The no sizing/no burnishing also achieves a high saturation level of 91.83%, while the wheat starch samples exhibited a more moderate saturation proximity of 67.12%.

The differences observed between burnished and unburnished samples provide critical insight into how mechanical surface treatment affects laser cleaning efficiency. No sizing/burnishing sample reached a lower saturation level of 77.64% compared to the no sizing/no burnishing sample, which reached a 91.83% saturation level. This discrepancy can be attributed to the mechanical compaction introduced by burnishing, as previously demonstrated in SEM analysis (see Figure 2.10). Burnishing compresses the fibrous network and reduces surface porosity, leading to tighter lodging of particulate contaminants within the inter-fiber voids. As a result, contaminants become more embedded and confined, which diminishes their exposure to the laser beam and reduces per-pulse removal efficiency. In contrast, unburnished samples present a more open fiber structure, where contaminants remain more loosely bound and thus more easily detached by laser pulses. This distinction emphasizes that, even in the absence of sizing, the physical morphology of the surface, altered by mechanical treatment, plays a decisive role in determining the outcome of laser cleaning.

The minimum fluence required to initiate visible cleaning varies among the different surface treatments, reflecting how each configuration interacts with laser energy. Among the sized samples, alum–egg requires the lowest energy input for cleaning to begin, with a threshold slightly above 0.31 J/cm^2. The corn starch-sized paper followed with a minimum fluence above 0.37 J/cm^2, while the wheat starch-sized sample requires the highest threshold among the three, exceeding 0.38 J/cm^2. This trend suggests that alum–egg sizing offers less resistance to contaminant removal, potentially due to its composition or interaction with the laser pulse, whereas starch-based sizing retains contaminants more firmly at the surface, delaying the onset of cleaning. This order of minimum fluence thresholds among the sized papers is further supported by the ANOVA Subset I results (Figure 2.12), where responsiveness to fluence is modeled for each condition. The alum–egg sized sample, which required the lowest fluence to initiate cleaning, also displayed the steepest modeled response among sized papers. Conversely, the wheat starch-sized sample, which exhibited the highest minimum fluence threshold, showed a comparatively dampened responsiveness. This coherence between physical thresholds and statistical responsiveness reinforces the interpretation that the ease of contaminant removal is strongly modulated by the nature of the sizing layer.

For the no sizing samples, the effect of burnishing significantly altered the laser response. The minimum fluence required to induce cleaning in the unburnished sample is just above 0.27 J/cm^2, whereas the burnished version required a higher energy input, exceeding 0.29 J/cm^2. This increase supports the notion that mechanical burnishing not only compacts the surface but also pushes contaminants deeper into the fiber network, making them more difficult to dislodge. Consequently, a higher laser fluence is needed to achieve the same initial level of cleaning on burnished, unsized paper.

Table 2.6. Fitted parameters for the exponential decay model and saturation proximity.

Sample	**Saturation reached (%)**	**Saturation fluence [J/cm^2]**	**R^2**	**RMSE**	**MAE**
Corn starch sizing	50.91	1.26	0.9724	0.8830	0.0020
Wheat starch sizing	67.84	0.96	0.9152	1.3524	0.0066
Alum - egg sizing	66.37	2.93	0.9868	0.9652	0.0596
No sizing/Burnishing	59.65	2.08	0.9427	1.8776	0.0303
No sizing/No burnishing	65.93	2.28	0.9885	1.5753	0.0700

According to the results presented in Table 2.6, the highest saturation removal is obtained for the wheat starch and alum–egg sized papers, reaching 67.84% and 66.37% respectively. The no sizing/no burnishing sample follows closely with 65.93%, while the no sizing/burnishing sample exhibits a lower removal level of 59.65%. The lowest saturation reached is recorded for the corn starch sizing paper, at only 50.91%, suggesting a greater resistance to contaminant removal under the applied conditions.

As with the saturation model, the role of mechanical surface treatment becomes apparent when comparing the no sizing/no burnishing and no sizing/burnishing samples. The burnished sample exhibits a reduced removal percentage (59.65%) compared to its unburnished counterpart (65.93%). This pattern aligns with the physical interpretation that burnishing compacts the surface fibers and reduces porosity, making contaminants more embedded and less susceptible to laser-induced detachment.

The fitted saturation fluence values provide additional insight into the behavior of each surface configuration under laser irradiation. Among the sized papers, wheat starch requires the lowest fluence to reach its cleaning limit (0.96 J/cm^2), followed by corn starch (1.26 J/cm^2), while alum–egg sizing demands the highest energy input (2.93 J/cm^2). Despite its high final removal percentage, the alum–egg layer appears to offer more resistance during cleaning of graphite + kaolinite mixture, which may be attributed to the cohesive nature of its composition or stronger contaminant adhesion for the kaolinite.

When examining no sizing samples, both variants required higher fluence values (2.08 and 2.28 J/cm^2) compared to the sized groups. This reinforces the notion that the absence of a sizing layer allows contaminants to penetrate more deeply into the fibrous structure, necessitating stronger laser input for effective removal. Between these two, the unburnished version reaches a slightly higher saturation level, supporting the interpretation that mechanical burnishing restricts contaminant exposure by tightening the fiber network.

Overall, the exponential decay model illustrates how both the chemical composition of the surface sizing and the physical structure influence laser cleaning efficiency. While some sizings facilitate contaminant removal at lower fluence levels, others create more resistant interfaces. Similarly, mechanical surface treatments such as burnishing significantly affect how readily the laser can access and remove contaminants embedded within the substrate.

The quantitative analysis presented here allows for multiple conclusions based on the experimental data. First, it enables the extraction of key values, such as the maximum substrate signal that can be recovered (asymptotic saturation) or the remaining contaminant concentration, both expressed relative to initial pre-treatment levels. It also allows us to determine the minimum fluence required to achieve noticeable cleaning effects. Furthermore, knowing the saturation value enables the assessment of how much of the ideal cleaning has been achieved based on the highest experimental fluence, and whether further cleaning, given the material's destruction threshold, would be safe. Overall, this approach offers a set of practical experimental parameters for professionals to optimize laser cleaning.

2.4 Conclusion

This study demonstrates that both the chemical composition of the sizing layer and the mechanical treatment of the surface critically influence the outcomes of femtosecond laser cleaning. The dual-model approach, combining exponential saturation and exponential decay analyses, enabled a deeper evaluation of contaminant removal dynamics from historical paper substrates.

The exponential saturation model reflects the emergence of the underlying substrate signal as contaminants are progressively removed. Within this framework, saturation levels varied across different sizing types and surface conditions, revealing how the accessibility of the paper substrate influences the perceived cleaning progression. In parallel, the exponential decay model characterizes the diminishing presence of the contaminant signal during laser exposure, offering insight into how effectively the overlying material is ablated. While each model is based on a different physical signal origin, their combined use allows for a complementary understanding of surface behavior during laser cleaning.

In terms of energy input, the minimum fluence required to initiate visible cleaning ranged from 0.27 to 0.38 J/cm^2 across the tested configurations. Although some fitted saturation fluence values extended beyond 2 J/cm^2, experimental protocols were designed to remain within safe operating margins for mock-up paper substrates. In such cases, increasing the number of laser pulses at moderate fluence levels remains a viable strategy for improving contaminant removal while minimizing potential damage. Since femtosecond lasers can operate at high repetition rates, this approach does not significantly extend the overall cleaning time.

The effect of mechanical surface treatment was also clearly observed. Non-burnished papers consistently responded more effectively to laser cleaning than their burnished counterparts, both in terms of saturation levels and contaminant removal efficiency. This behavior is attributed to the surface compaction caused by burnishing, which reduces porosity and fiber openness, thereby limiting contaminant exposure to the laser and reducing per-pulse cleaning efficacy.

While fluence remains a critical parameter in determining cleaning outcomes, it must be considered alongside other factors such as the number of laser pulses, the type and structure of the sizing, the nature of the contamination, and the physical morphology of the surface. Preliminary testing is therefore essential to define safe and efficient cleaning protocols tailored to the specific characteristics of each paper artifact. These insights contribute to the refinement of laser cleaning strategies, paving the way toward safer, more selective, and material-sensitive conservation practices in heritage science.

Chapter 3

Challenges in Colorimetric Evaluation of Femtosecond Laser Cleaning on Historical Leather: The Role of Surface Porosity and Microstructure

This chapter is based on the content of the following article; as a result, some passages may appear verbatim from this publication:

Challenges in Colorimetric Evaluation of Femtosecond Laser Cleaning on Historical Leather: Influence of Surface Porosity and Microstructure, submitted, 2025.

C. Y. Boynukara, M. Welsch, L. Caron, M. Uguryol, G. Mavili, L. Razzari, C. Muehlethaler, A. Ruediger, and P. Antici

This work was led by C. Y. Boynukara, who conceptualized the study, developed the experimental design, performed the laser cleaning procedures, and conducted the data analysis. M. Uguryol and G. Mavili contributed to the preparation of the leather samples and surface contamination protocols. M. Welsch assisted in the execution of the cleaning experiments. L. Caron collaborated with C. Y. Boynukara on the colorimetric measurements and data acquisition. L. Razzari, C. Muehlethaler, A. Ruediger, and P. Antici reviewed the manuscript and provided scientific and editorial feedback throughout the study.

Abstract

This study presents a critical evaluation of femtosecond (fs) laser cleaning applied to historical leather, addressing the persistent challenge of quantifying cleaning efficacy on porous and morphologically heterogeneous surfaces. Vegetable-tanned goat leather samples, artificially aged and contaminated with graphite powder, were treated using fs-laser irradiation under varying pulse energies and shot numbers. Cleaning performance was assessed via high-resolution colorimetric measurements (CIE L*a*b*), complemented by quantitative porosity analysis using optical microscopy and ImageJ segmentation, as well as grayscale-based surface plot visualizations. Pore size measurements further revealed a heterogeneous distribution of small and large pores, each affecting cleaning uniformity differently.

Fs-laser cleaning effectively reduced visible surface contamination, with ΔE^* values approaching the commonly accepted perceptibility threshold of 2.0. However, microstructural irregularities inherent to leather substrates introduced significant reflectance variability, highlighting the limitations of area-averaged colorimetry for fully capturing cleaning efficacy. Comparative measurements on graphite-contaminated parchment, characterized by negligible porosity but micro-topographical undulations, further demonstrated that both pore structure and micro-relief govern the consistency and reliability of optical outcomes.

Overall, these findings underscore the importance of interpreting colorimetric data in the context of surface morphology and porosity. They also emphasize the importance of integrating spatially localized optical diagnostics into fs-laser cleaning protocols to ensure both treatment effectiveness and the long-term visual integrity of heritage leather objects.

3.1 Introduction

Cleaning historical objects demands meticulous care and precision, as preserving the authenticity and integrity of their surfaces requires methods that minimize the risk of damage [47]. Consequently, the development of non-invasive and non-contact cleaning techniques has become essential for ensuring both effective treatment and the long-term preservation of cultural artifacts.

Among these techniques, fast (ns) and ultra-fast (fs) pulsed lasers have recently gained significant attention due to their precision and commercial viability [51]. Unlike conventional methods, laser cleaning is a non-contact process that utilizes parameters such as wavelength, pulse duration, and polarization to tailor the treatment to delicate materials. This high degree of control not only mitigates the risks associated with mechanical or chemical cleaning approaches but also enhances efficiency and safety during treatment [54, 55]. Moreover, laser cleaning is often faster and more cost-effective, making it a compelling option for cultural heritage conservation [56].

Initially developed for stone surfaces, ultra-fast laser cleaning techniques have since been extended to a variety of materials, including metals, textiles, paintings, and paper [52]. Among these, paper artifacts pose a particular challenge due to the need to preserve their fragile, fibrous organic matrix [53]. Recognizing these complexities, numerous research groups have conducted extensive studies to optimize laser parameters for paper cleaning, producing valuable insights into safe and effective conservation practices [57, 58, 68]. Similarly, parchment, a substrate derived from animal hide and closely related to leather, has received considerable attention in laser cleaning research due to its fragility and historical significance [69, 70, 71].

Building on this foundation, leather, a material widely used from antiquity through the medieval period, remains a critical focus in cultural heritage preservation [72, 73]. The accumulation of dust and soot on leather surfaces presents a significant conservation challenge. These particles are not only abrasive but also difficult to remove due to the inherently porous and uneven structure of leather. Additionally, they can accelerate leather degradation by promoting microbial activity and increasing the material's moisture uptake [74]. In humid conditions, pollutant gases can react with water to produce acids, which in turn oxidize the lubricants and dressing layers on the leather surface and penetrate the collagen matrix. This oxidation process leads to darkening of the leather's appearance and contributes to accelerated deterioration [75].

Having previously focused on the laser cleaning of paper-based substrates, this study turns toward a more complex and heterogeneous material: tanned leather. Compared to paper, leather presents greater challenges for laser-based interventions due to its fibrous three-dimensional structure, variable porosity, and strong surface absorption characteristics.

In this study, we investigated the application of femtosecond (fs) laser cleaning on artificially aged lgoat leather and parchment contaminated with graphite powder. While fs-laser cleaning proved effective in visibly reducing surface soiling, post-cleaning colorimetric measurements exhibited notable variability. We hypothesize that this inconsistency arises from the porous and topographically heterogeneous nature of tanned leather, which affects laser–material interactions and reflectance behavior. To further examine this hypothesis, we performed comparative tests on a graphite-contaminated goat parchment sample, which possesses a markedly smoother and less porous surface. Our findings suggest that surface morphology critically influences the consistency and reliability of optical outcomes in fs-laser cleaning of collagen-based substrates.

3.2 Materials and methods

3.2.1 Leather samples

Goat leather is the most widely utilized leather type in the production of historical bookbindings all over the world [47]. In classical Turkish bookbinding crafts, goat leather (*sahtiyan*) produced by vegetable tanning has also been extensively preferred for bookbinding and book covers over sheep, lamb, or cattle leather for centuries. This preference is due to its greater durability, uniform thickness, and superior workability compared to those other materials. Goat leather tanning is an important step to enhance the durability of the leather and enable its use in various applications. The first stage involves cleaning the leather and removal of hair. Vegetable tannins have historically been the primary materials used for tanning leather. For traditional vegetable tanning, the leather was soaked in baths including extracts from tannin-containing plants like oak, acacia, and chestnut. After tanning, the leather was dyed and then treated with various vegetable oils to make it ready for use.

At this stage, if the wet skin is stretched and dried without tanning, it becomes parchment; if it is tanned, it becomes leather. In this study, parchment samples were prepared following traditional methods: the goat skins were first washed to remove dirt acquired during slaughtering or transport, then underwent chemical hair removal in large drums, followed by repeated enzymatic washes to eliminate any residual depilatory agents. Flesh and fat residues were then mechanically removed from the inner surface, a process known as “etleme” in Turkish practice. Finally, the wet skins were stretched by experienced craftsmen called “çiviciler”, who secured them onto steel frames and dried them in heated cabinets without any tanning treatment. This process yielded parchment rather than leather, preserving its characteristic untanned collagen structure. Fully processed goat parchment was supplied by KAREDERİ Leather & Parchment (Turkiye), a workshop specializing in artisanal parchment production for this project.

For experimental use, both the vegetable-tanned goat leather and the parchment samples were treated similarly to replicate historical soiling and aging conditions. The leather samples were dyed with RODA Dye NF brand dyes in the colors Orange and Havana. These colors correspond to the color codes of RODA brand leather dyes, which were supplied by GMW, a German supplier that specializes in equipment and materials for paper conservators and bookbinders.

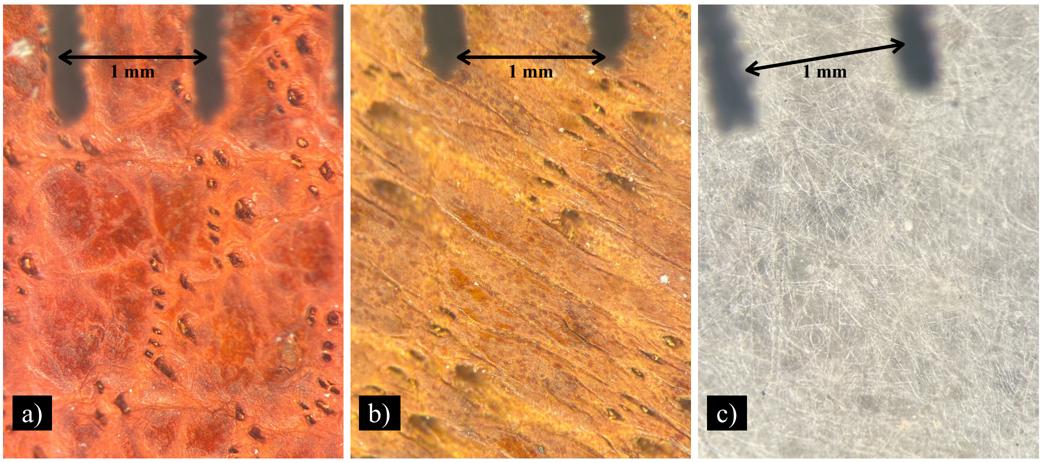


Figure 3.1. Surface morphology of representative samples captured at 2× magnification using an Olympus stereo microscope. a) Orange colored goat leather, b) Havana colored goat leather, and c) parchment. A 1mm scale bar is shown in each image. The perceived colors in the images may slightly differ from the actual samples due to lighting and imaging parameters.

To provide visual context for the material surfaces after aging but prior to any soiling or laser treatment, low-magnification optical images were acquired using an Olympus stereo microscope at 2× magnification. Figure 3.1 presents representative examples of a) Orange colored leather, b) Havana colored leather, and c) parchment. A 1 mm scale bar is included in each image. These photographs are intended to give the reader a qualitative impression of the substrate appearance following artificial aging, helping to contextualize later observations on cleaning performance and material response.

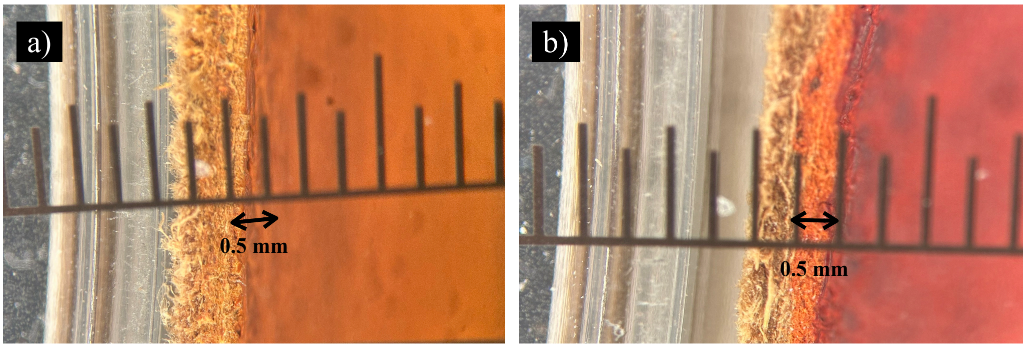


Figure 3.2. Cross-sectional optical images of leather samples dyed with a) Havana and b) Orange RODA NF dyes. Images were acquired using an Olympus stereo microscope at 2× magnification. Each scale division corresponds to 0.5 mm. The dyed layers are visible as distinct surface coatings. The perceived colors in the images may slightly differ from the actual samples due to lighting and imaging parameters.

Cross-sectional microscopy was used to estimate the thickness of the dye layer. Scale-referenced optical images of the leather cross-sections, shown in Figure 3.2a and 3.2b, were acquired using an Olympus stereo microscope at 2× magnification. The dyed coatings were approximately 500 μm thick for both the Havana and Orange colored samples. While the Havana dye appeared visually less distinct due to its similarity to the natural color of the leather substrate, its thickness was consistent with that of the orange dye, suggesting comparable dye penetration depth. These values confirm that the leather surface was uniformly dyed to a moderate depth, which may influence both laser–material interaction and post-cleaning optical behavior.

To simulate soiling, graphite powder (AF spezial, Kropfmühl, D90 $\leq$ 25 μm) was applied to represent unburned carbon particles (soot). The D90 specification indicates that 90% of the graphite particles are smaller than 25 μm, reflecting the fine particulate nature of the contaminant and its potential to penetrate porous substrates. The leather samples, including both soiled and unsoiled specimens, were prepared following the aging conditions recommended in the BS EN ISO 17228:2015 standard. The aging process was conducted at a controlled temperature of 50°C and a relative humidity of 90% for a duration of four days. This procedure ensured that the leather surfaces closely simulated aged and soiled conditions commonly observed in historical artifacts, providing a realistic context for laser cleaning experiments.

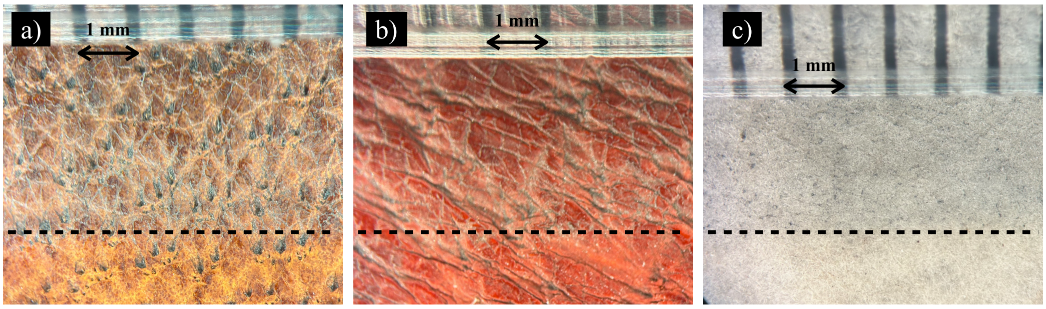


Figure 3.3. Surface images of the three sample types after contamination and aging process, acquired using an Olympus stereo microscope at 2× magnification. Panels a), b), and c) correspond to Havana colored leather, Orange colored leather, and parchment, respectively. In each image, the upper region above the dashed black line shows the graphite-contaminated area, while the lower region corresponds to the laser-cleaned zone. The scale bar represents 1 mm. The perceived colors in the images may slightly differ from the actual samples due to lighting and imaging parameters.

To provide visual context for the soiling and cleaning conditions examined in this study, Figure 3.3 presents surface images of three artificially soiled and aged sample types following partial fs-laser cleaning. Panel a) shows a Havana colored leather sample, panel b) displays an Orange colored leather sample, and panel c) corresponds to a parchment specimen. All samples were subjected to the same soiling and accelerated aging procedures prior to laser application. The images were acquired using an Olympus stereo microscope at 2× magnification. In each case, the region above the black dashed line represents the graphite-contaminated surface, while the area below the line corresponds to the portion where fs-laser cleaning was applied. A 1 mm scale bar is included to provide spatial reference. This figure is intended to support visual interpretation of the cleaning effect and offers an illustrative comparison between soiled and laser-treated regions under optical magnification; it does not imply complete contaminant removal within the treated areas.

3.2.2 Femtosecond Laser Cleaning Setup

The laser cleaning experiments were conducted using a Yb:KGW femtosecond laser source (Pharos, Light Conversion), delivering pulses at a central wavelength of 1030 nm with a pulse duration of 170 fs. The laser beam was propagated through free space and focused onto the sample surface using a plano-convex lens (Thorlabs LA1254-B) with a focal length of 1500 mm, positioned at a working distance of approximately 1430 mm from the target. The resulting beam diameter at focus was estimated to be approximately 500 μm. A schematic illustration of the experimental configuration is shown in Figure 3.4.

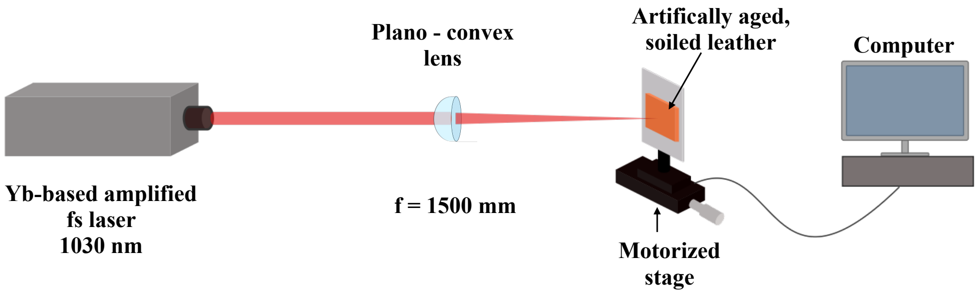


Figure 3.4. Schematic of the femtosecond laser cleaning setup.

Samples were mounted vertically on a motorized translation stage (Thorlabs MTS50A-Z8), enabling precise control of the sample position along the z-axis. All irradiation was performed under normal incidence and ambient air conditions. Fs-laser cleaning was performed by raster-scanning the beam across the surface under normal incidence and ambient air conditions. The scanning parameters were adjusted to ensure homogeneous pulse overlap and uniform fluence distribution across the treated area.

To investigate the influence of pulse energy and cumulative dose on cleaning performance, pulse energies were varied from 90 to 120 μJ for Havana colored leather, from 110 to 150 μJ for Orange colored leather, and from 120 to 180 μJ for parchment. For all materials, each energy condition was tested using both 50 and 100 laser pulses per site to evaluate the effect of cumulative exposure on cleaning efficacy.

These energy windows were determined based on preliminary screening tests. Microscopic inspection after laser exposure showed that Havana colored leather could be overcleaned even at relatively low energies, whereas Orange colored leather required higher energy levels to exhibit measurable cleaning. Accordingly, different energy ranges were assigned to each leather type to ensure material-specific optimization, rather than direct comparison between them.

3.2.3 Optical Microscopy for Surface Morphology Analysis

Optical microscopy was performed using a Zeiss Axio Scope A1 upright microscope (Carl Zeiss Microscopy GmbH, Carl-Zeiss-Promenade 10, 07745 Jena, Germany) equipped with a 10× objective lens. For each sample, images were collected from three different regions on clean (uncontaminated) but artificially aged leather substrates to assess surface morphology, pore distribution, and porosity parameters. Images were captured using ZEN 2 core software (version 2.5; Carl Zeiss Microscopy GmbH, Germany) under consistent illumination and contrast settings to ensure comparability across samples.

Porosity analysis was conducted using ImageJ software (NIH, USA). Scale calibration was performed by drawing a line along the 100 μm scale bar embedded in each image and using the *Set Scale* function to define the known distance (100 μm) with micrometers as the unit of length. Following calibration, pore areas were measured manually using the *Polygon Selection* tool. This tool was chosen because pores exhibited irregular shapes that could not be accurately captured with predefined circular or rectangular selections. This allowed free-form outlining of each pore boundary to ensure precise area measurement. Each pore was individually outlined and measured via the *Analyze > Measure* function in ImageJ, with results recorded in the measurement table for subsequent analysis. Finally, the total image area was measured by selecting the entire image and using the *Analyze > Measure* function.

Porosity (%) was calculated using the following equation (3.1):

$$\text{Porosity (\%)} = \frac{\text{Total pore area}}{\text{Total image area}} \times 100 \tag{3.1}$$

where total pore area is the sum of all individually measured pore areas, and total image area is the area of the entire image field.

These procedures were repeated for Orange colored leather, Havana colored leather, and parchment samples. For each material, three separate images from different regions were analyzed to account for spatial variability in surface porosity. The porosity values obtained from these replicates were used to calculate the mean porosity and standard deviation for each material type, providing a quantitative estimate of both average porosity and its variability. It should be noted that this approach provides a two-dimensional (2D) estimation of porosity based on surface images and does not account for the whole three-dimensional (3D) pore network within the material.

After calibration, pore size measurements were performed using ImageJ software with the known scale bar in each optical microscopy image. Following scale calibration, the *Line* tool was used to draw a straight line across the maximum diameter of each visually identifiable pore, and the measured length was recorded in micrometres. This approach allowed for rapid estimation of pore diameters to complement the quantitative area-based porosity analysis.

Grayscale surface plots were generated to visualize pixel intensity variations across sample surfaces. Optical microscopy images were first converted to 8-bit grayscale format in ImageJ. The built-in *Surface Plot* function was then used to produce plots representing grayscale intensity distributions, which reflect local reflectance differences and microstructural heterogeneity. It should be noted that these plots do not provide true topographical height information but instead illustrate relative brightness variations within each image.

3.2.4 Color Measurement

To evaluate the effectiveness of the laser cleaning process, color measurements were performed on laser-cleaned, artificially aged, soiled leather samples and compared with non-soiled, aged reference samples. The VSC® 8000/HS (Foster + Freeman, UK) was used for this purpose, utilizing the CIE L*a*b* color space to analyze the lightness (L^*), green-to-red (a^*), and blue-to-yellow (b^*) components. The differences in these values, expressed as ΔL^*, Δa^*, and Δb^*, were calculated to assess the relative color changes between the cleaned areas and reference samples.

Achieving a ΔE^* value, which combines ΔL^*, Δa^*, and Δb^* into a single metric,as defined in equation 3.2:

$$\Delta E^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \tag{3.2}$$

is crucial for maintaining the aesthetic integrity of historical artifacts [76, 77]. A ΔE^* value below approximately 2 is particularly desirable, as it is imperceptible to the human eye and indicates minimal visual alteration [78]. Ensuring minimal color change not only preserves the aesthetic properties but also avoids chemical degradation, as discoloration often signifies underlying structural changes [79].

For Havana colored leather, Orange colored leather, and parchment samples, optical microscope images were acquired from three different regions on clean (uncontaminated) but artificially aged samples to assess surface morphology and quantify porosity prior to laser exposure. For each laser-treated area, L^*, a^*, and b^* values were collected from three different positions within the ablation zone. The same procedure was applied to the corresponding reference area. Mean values and standard deviations were computed for each color coordinate in the treated group, while only the mean values were considered for the unsoiled reference.

Since the surface of artificially aged leather is inherently rough and morphologically irregular, uniform laser–material interaction across the entire irradiated area cannot be assumed. The presence of microtopographic features such as pores, ridges, and depressions alters the local angle of incidence and effective fluence, which in turn affects the energy absorbed at each point. Consequently, certain zones, such as elevated protrusions, may undergo more efficient cleaning, while recessed or shadowed areas may remain partially soiled despite identical laser parameters. To better capture this inherent spatial variability, reflectance measurements were taken from multiple locations within the raster-scanned area, rather than a single fixed location. Our spectrophotometer enabled spot-wise measurements on areas smaller than the 500 µm laser beam diameter, allowing reflectance data to be recorded from peak, valley, and intermediate zones across the cleaned area. The leather surface presents a heterogeneous micro-relief dominated by pores and fine fiber bundles, each with randomly oriented facets. In principle, such structures can lead to anisotropic or direction-dependent reflectance, particularly when measured under highly oblique illumination geometries. In this work, color was measured under a near-normal illumination/observation geometry (illumination angle $\leq 10°$ from the surface normal) using broadband visible illumination. At this spatial scale, corresponding to only a few pores per acquisition area, the reflected signal represents a local average of diffuse reflectance and is effectively insensitive to the sample orientation. Consequently, rotating the specimen would not alter the recorded CIE L*a*b* values beyond instrumental repeatability.

The total color difference ΔE^* between treated and untreated regions was calculated using equation 3.2, as described before. To account for surface-related reflectance variability, the dispersion of L^*, a^*, and b^* values across different points within the treated region was used to estimate the measurement uncertainty.

The combined uncertainty in ΔE^* was estimated using first-order Gaussian error propagation. In this approach, σ_{L^*}, σ_{a^*}, and σ_{b^*} denote the standard deviations of L^*, a^*, and b^* across multiple measurement spots within the treated region, and covariance terms between L^*, a^*, and b^* were neglected because of the limited number of spatial replicates. Consequently, $\sigma_{\Delta E^*}$ should be interpreted as an approximate measure of spatial reflectance variability rather than as a rigorous statistical confidence interval. The following equation 3.3 was used:

$$\sigma_{\Delta E^*} = \sqrt{\left(\frac{\Delta L^*}{\Delta E^*}\right)^2 \sigma_{L^*}^2 + \left(\frac{\Delta a^*}{\Delta E^*}\right)^2 \sigma_{a^*}^2 + \left(\frac{\Delta b^*}{\Delta E^*}\right)^2 \sigma_{b^*}^2} \tag{3.3}$$

This yielded a single uncertainty value ($\pm\Delta E^*$) that reflects the spatial heterogeneity of the laser–material interaction and allows a more representative evaluation of cleaning performance

3.3 Results and Discussion

3.3.1 Optical Microscopy-based porosity comparison

The surface porosity of Orange colored leather samples was evaluated using optical microscopy and ImageJ-based analysis. Representative images from two different regions are shown in Figure 3.5, highlighting the spatial variability of pore distribution across the sample surface. Quantitative analysis of three replicate images yielded an average porosity of approximately 7% with a standard deviation of 4%, indicating a moderately low porosity with considerable heterogeneity between different areas.

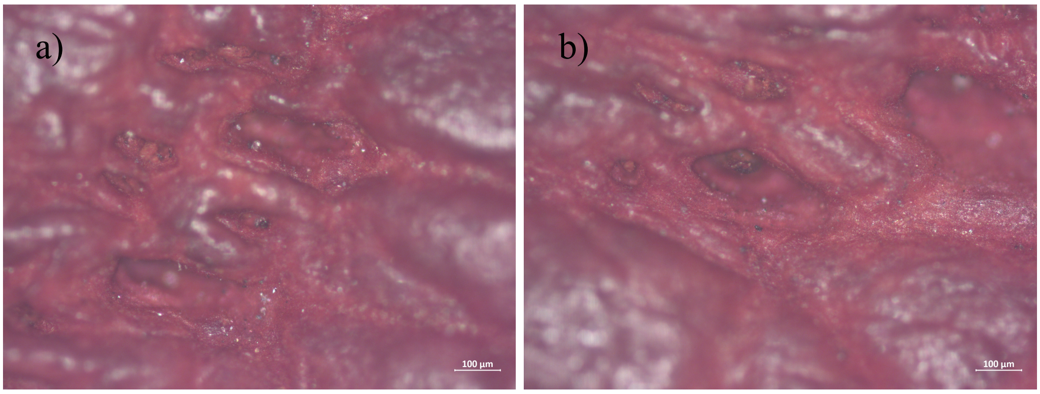


Figure 3.5. Optical microscopy images of Orange colored leather showing pore distribution at 10× magnification with a 100 μm scale bar. Images a) and b) were taken from different regions to assess spatial variability in surface porosity. The field of view for each image is approximately 1120 μm × 840 μm.

The surface porosity of Havana colored leather sample was assessed using optical microscopy and ImageJ-based analysis. Representative images are presented in Figure 3.6, where a) shows the general surface morphology and b) illustrates a pore-focused view. Quantitative analysis of three replicate regions yielded an average porosity of approximately 14% with a standard deviation of 3%, indicating a moderately porous and heterogeneous surface.

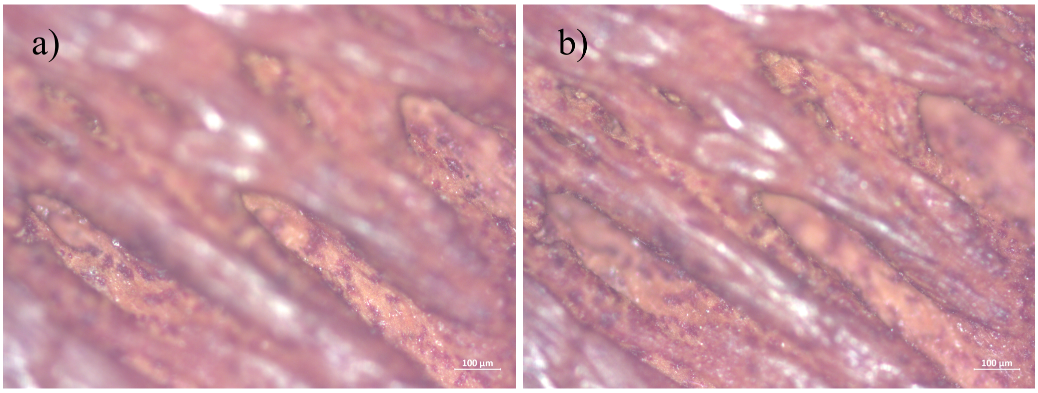


Figure 3.6. Optical microscopy images of Havana colored leather used for porosity analysis at 10× magnification with a 100 μm scale bar. Image a) shows the general surface morphology, while image b) illustrates a pore-focused view. The field of view for each image is approximately 1120 μm × 840 μm.

Quantitative analysis of three replicate regions yielded an average porosity of 0.195% with a standard deviation of 0.165%, a value insignificant compared to the Orange and Havana colored leather samples. Representative optical microscopy images from two different regions are shown in Figure 3.7. Overall, the parchment exhibited an almost pore-free structure, with only a few isolated pores detected across the analyzed areas. Such negligible porosity is consistent with the intended function of parchment, which was historically manufactured to provide a smooth, dense surface suitable for writing; the presence of pores would allow ink to penetrate and spread, compromising legibility. Nevertheless, despite its minimal pore content, the microstructure of parchment remains highly irregular, characterized by a complex landscape of peaks and valleys rather than a paper-like flatness, reflecting micro-topographical heterogeneity. This microstructural irregularity is expected to influence the colorimetric measurements discussed in the following sections.

Pore size measurements were conducted on Orange and Havana colored leather samples to complement the porosity analysis. As shown in Figure 3.8, both leathers exhibited a heterogeneous distribution of pore sizes, with diameters ranging from approximately 71 μm to 260 μm, although some larger pores are also observed in the microscopy images. This variability indicates that the surfaces contain both small and large pores, each influencing laser cleaning outcomes differently. If the same overall porosity were composed predominantly of numerous small pores, the surface would present less pronounced topographical variation, potentially enabling more uniform laser–material interaction and reduced reflectance variability. Conversely, if the porosity were concentrated in fewer but larger pores, shadowing effects and local

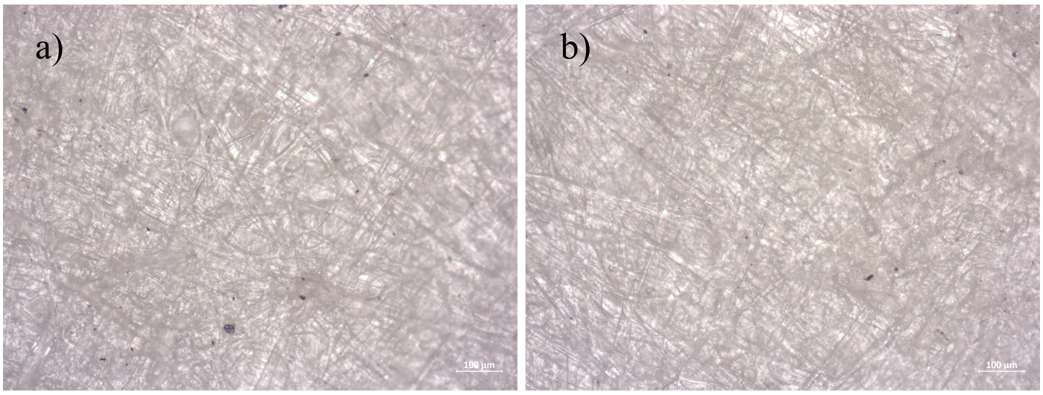


Figure 3.7. Optical microscopy images of parchment sample showing pore distribution at 10× magnification with a 100 μm scale bar. Images a) and b) were taken from different regions to assess spatial variability in surface porosity. The field of view for each image is approximately 1120 μm × 840 μm.

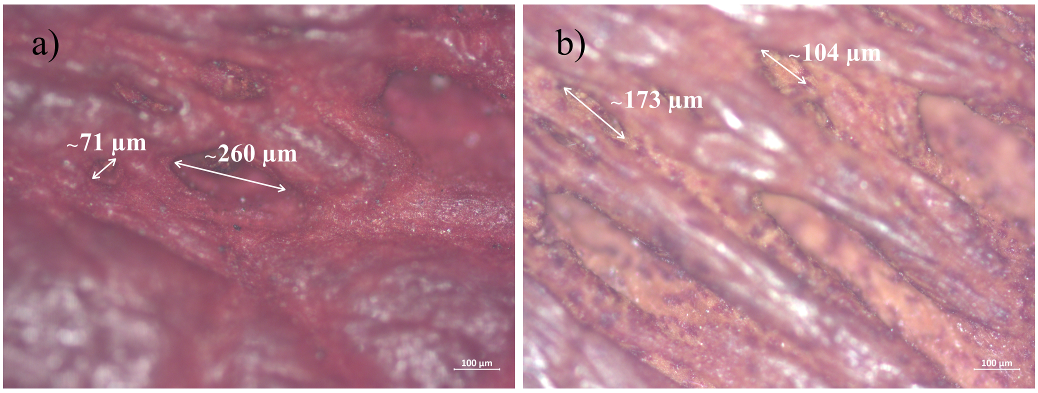


Figure 3.8. Representative pore size measurements from optical microscopy images: a) Orange colored leather, b) Havana leather. The field of view for each image is approximately 1120 μm × 840 μm. Approximate pore diameters are indicated in micrometers.

fluence disparities would become more significant, leading to uneven cleaning and patchy visual results. The mixed pore size distribution observed in these samples likely amplifies reflectance variability and imposes inherent limitations on achieving consistent fs-laser cleaning performance.

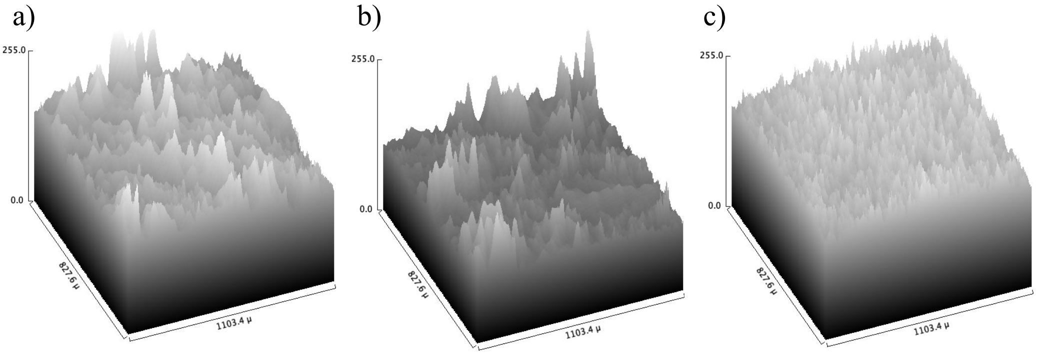


Figure 3.9. Grayscale-based surface plots generated from 8-bit converted optical microscopy images, showing pixel intensity distribution for a) Havana colored leather, b) Orange colored leather, and c) parchment samples. These plots visualize reflectance variation and surface texture but do not represent true topographical height measurements.

Grayscale-based surface plots for Havana colored leather, Orange colored leather, and parchment samples are shown in Figure 3.9, illustrating their distinct microstructural characteristics. Images a) and b) depict Havana and Orange colored leather samples, respectively, both characterized by textured surfaces with numerous intensity peaks and valleys reflecting their porous and microstructurally heterogeneous nature. Image c) presents parchment, with an almost flat grayscale surface plot showing minimal intensity fluctuations, suggesting a generally non-porous but subtly irregular surface morphology. Overall, these surface plot visualizations corroborate the optical microscopy observations and porosity measurements, highlighting the contrast between the more textured and porous leather samples and the nearly pore-free but micro-topographically uneven parchment.

3.3.2 Colorimetry and Reflectance Variability

As shown in Figure 3.10 and Table 3.1, ΔE^* values for the Orange colored leather sample exhibited a non-monotonic response to increasing pulse energy, particularly under 50-shot scanning conditions. The color difference ranged from 8.70 ± 4.38 to 2.56 ± 0.89, with the lowest value recorded around 137 μJ, indicating optimal cleaning performance at this energy level. At lower energies, ΔE^* values were elevated, likely due to incomplete removal of graphite residues. At higher energies such as 153 μJ, ΔE^* values remained elevated: 5.43 ± 3.17 for 50 shots and 7.73 ± 0.63 for 100 shots. These results indicate the onset of overcleaning, where not only residual graphite is removed but the dyed surface layer itself may be locally altered or partially ablated, resulting in changes in both lightness and chromatic coordinates. The substantial standard deviations observed across both lower and higher energy levels, including 8.70 ± 4.38 at 120 μJ and 8.15 ± 2.68 at 153 μJ, reflect notable reflectance variability across the cleaned areas. This variability is likely caused by uneven laser–surface interaction, where protruding regions received excessive exposure while recessed areas remained insufficiently treated.

Table 3.1. Total color difference (ΔE^*) for fs-laser cleaned Orange colored leather sample at varying pulse energies and shot numbers. Values are reported as $\Delta E^* \pm$ propagated uncertainty ($\sigma_{\Delta E^*}$).

Sample Name	Pulse Energy [μJ]	ΔE^* (50 shots)	ΔE^* (100 shots)
Orange colored	110	8.15 ± 2.68	8.02 ± 0.35
	120	8.70 ± 4.38	5.70 ± 2.57
	130	5.38 ± 2.23	2.63 ± 2.26
	137	2.56 ± 0.89	2.70 ± 1.59
	153	5.43 ± 3.17	7.73 ± 0.63

In contrast, the 100-shot condition yielded a more consistent trend, with generally lower ΔE^* values and a minimum of 2.63 ± 2.26 at 130 μJ. Although some variability persisted, the narrow error margins, such as 8.02 ± 0.35 at 110 μJ, indicate more uniform laser-material interaction. This supports the interpretation that increasing pulse count can mitigate the influence of microstructural irregularities, even though overcleaning remains a risk at higher energy levels.

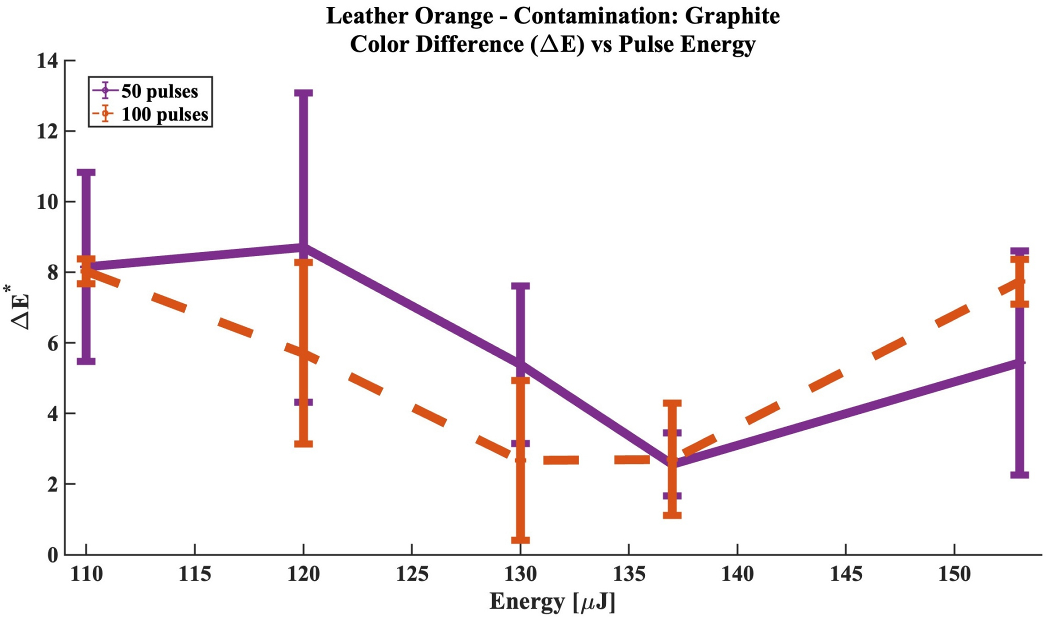


Figure 3.10. Color difference (ΔE^*) values measured on Orange colored leather samples contaminated with graphite after fs-laser cleaning, as a function of pulse energy for 50 and 100 laser shots. Error bars represent standard deviations from three replicate measurements.

As shown in Figure 3.11 and Table 3.2, ΔE^* values for the Havana colored leather sample remained relatively low and stable across the tested pulse energy range under 50-shot conditions. Between 90 and 112 μJ, ΔE^* ranged from 3.07 ± 1.50 to 2.12 ± 0.91, without any significant spikes in variability. The lowest value was recorded at 112 μJ, suggesting a favorable energy window where femtosecond laser interaction effectively removes graphite contamination while minimizing surface alteration. The narrow standard deviations observed throughout this range indicate a consistent and uniform cleaning response across the treated surface.

Table 3.2. Total color difference (ΔE^*) for fs-laser cleaned Havana colored leather sample at varying pulse energies and shot numbers. Values are reported as $\Delta E^* \pm$ propagated uncertainty ($\sigma_{\Delta E^*}$).

Sample Name	Pulse Energy [μmJ]	$\boldsymbol{\Delta E^*}$ (50 shots)	$\boldsymbol{\Delta E^*}$ (100 shots)
Havana colored	90	3.07 ± 1.50	2.91 ± 2.45
	100	3.57 ± 1.30	2.41 ± 1.71
	110	3.55 ± 1.15	6.63 ± 1.69
	112	2.12 ± 0.91	5.40 ± 0.78
	120	5.13 ± 2.45	7.70 ± 0.37

In contrast, the results obtained under 100-shot exposure revealed a distinct trend. At energies below 100 μJ, ΔE^* values were relatively high and variable, pointing to insufficient cleaning. A clear minimum was observed at 100 μJ, where ΔE^* dropped to 2.41 ± 1.71, indicating optimal performance for contaminant removal. However, beyond this point, the color difference increased rapidly, reaching 6.63 ± 1.69 at 110 μJ and 7.70 ± 0.37 at 120 μJ. These elevated values likely result from overcleaning effects, where excessive laser energy causes undesirable changes to surface reflectance or fiber morphology. This behavior emphasizes that while extended exposure can enhance cleaning efficacy, it also narrows the safe processing window and increases the risk of aesthetic degradation.

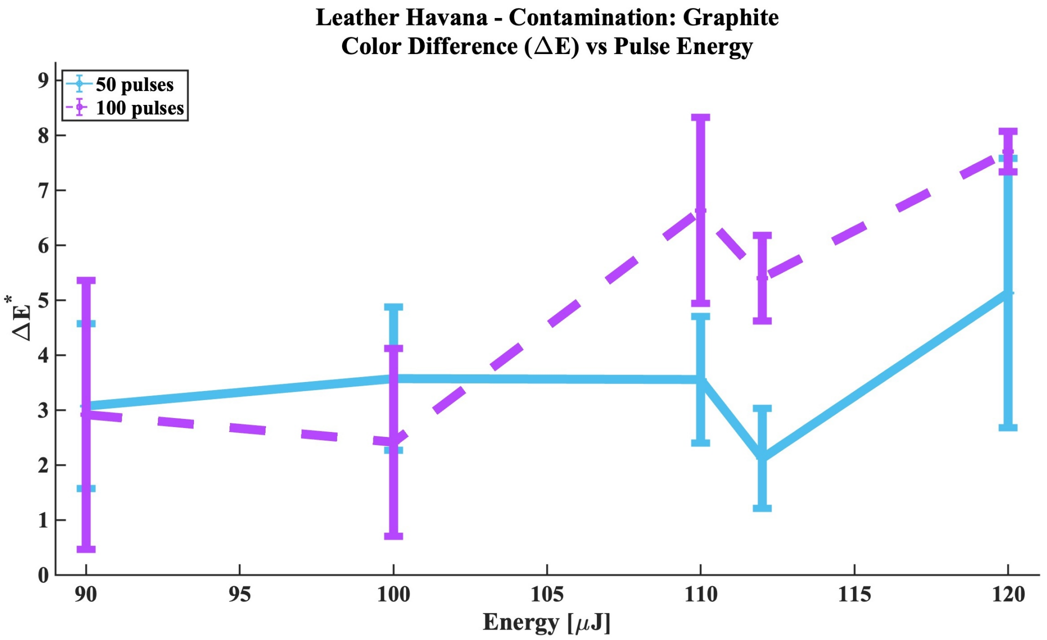


Figure 3.11. Color difference (ΔE^*) values measured on Havana colored leather samples contaminated with graphite after fs-laser cleaning, as a function of pulse energy for 50 and 100 laser shots. Error bars represent standard deviations based on three replicate measurements.

The parchment samples exhibited notably lower reflectance variability compared to the leather substrates, which aligns with their smoother and less porous surface morphology. As shown in Figure 3.12 and Table 3.3, ΔE^* values followed a more coherent and predictable trend across both 50-shot and 100-shot conditions, highlighting the reduced impact of microtopographic irregularities.

Table 3.3. Total color difference (ΔE^*) for fs-laser cleaned Parchment sample at varying pulse energies and shot numbers. Values are reported as $\Delta E^*\pm$ propagated uncertainty ($\sigma_{\Delta E^*}$).

Sample Name	Pulse Energy [μmJ]	ΔE^* (50 shots)	ΔE^* (100 shots)
Parchment	120	9.34 ± 1.05	5.74 ± 0.46
	130	7.81 ± 0.57	5.75 ± 0.55
	140	7.74 ± 1.04	4.10 ± 2.58
	150	3.50 ± 2.38	3.29 ± 1.79
	160	5.89 ± 3.85	6.14 ± 1.30
	170	6.33 ± 1.74	7.01 ± 2.34
	180	9.04 ± 1.04	10.85 ± 1.64

Under 50-shot exposures, ΔE^* steadily decreased from 9.34 ± 1.05 at 120 μJ to a minimum of 3.50 ± 2.38 at 150 μJ, before increasing again to 9.04 ± 1.04 at 180 μJ.

A similar pattern was observed with 100-shot cleaning, where ΔE^* dropped to 3.29 ± 1.79 at 150 μJ and then rose sharply to 10.85 ± 1.64 at 180 μJ. This consistent dip around 150 μJ suggests an optimal fluence window in which contamination is effectively removed without compromising the visual or structural integrity of the substrate.

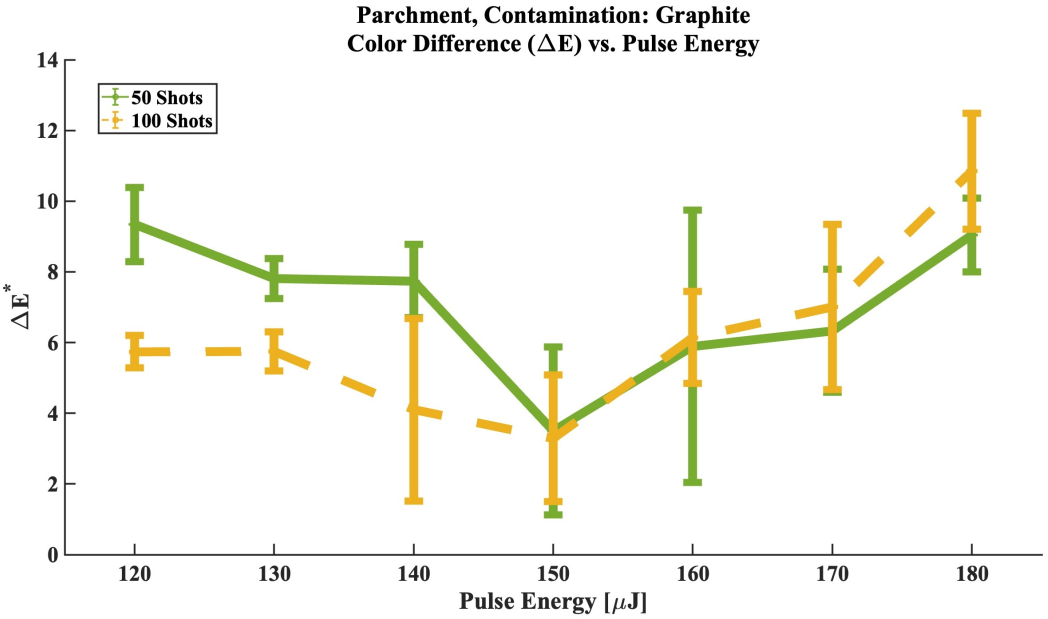


Figure 3.12. Color difference (ΔE^*) values measured on Parchment sample contaminated with graphite after fs-laser cleaning, as a function of pulse energy for 50 and 100 laser shots. Error bars represent standard deviations from three replicate measurements.

Despite the parchment's uniform morphology, the increase in ΔE^* values beyond 150 μJ indicates that excessive energy delivery can still lead to optical or structural degradation, likely due to overexposure effects such as collagen denaturation. This demonstrates that even substrates with minimal surface roughness require careful control of energy parameters to avoid unintended damage.

None of the tested conditions for Orange colored leather, Havana colored leather, or parchment achieved ΔE^* values below 2.0, the commonly accepted visibility threshold for the human eye. This suggests that residual contamination likely remained trapped within microscopic recesses or that reflectance was altered due to partial ablation of exposed fibers. These interpretations align with the porosity analysis results, which revealed that Orange and Havana colored leather surfaces exhibit notable porosity and microstructural heterogeneity. In contrast, parchment displayed negligible porosity, as confirmed by both optical microscopy and grayscale-based surface plots, yet its microstructure still exhibited subtle undulations that could influence reflectance measurements. Overall, such surface characteristics contribute to the observed variability in colorimetric outcomes, highlighting that, regardless of parameter optimization, the complex interaction between the laser and heterogeneous or micro-irregular surfaces imposes inherent limitations on achieving completely uniform and visually imperceptible cleaning results.

When examining the colorimetric data in more detail, it becomes clear that for each sample, both 50-shot and 100-shot conditions yielded ΔE^* values that approached the commonly accepted threshold, despite not falling below it:

- **Orange colored leather** reached 2.56 ± 0.89 at 137 μJ (50 shots) and 2.63 ± 2.26 at 130 μJ (100 shots),
- **Havana colored leather** showed 2.12 ± 0.91 at 112 μJ (50 shots) and 2.41 ± 1.71 at 100 μJ (100 shots),
- **Parchment** achieved 3.50 ± 2.38 at 150 μJ (50 shots) and 3.29 ± 1.79 at 150 μJ (100 shots).

These values represent averages across the entire measurement spot, including both elevated zones and recessed micro-areas. It is likely that if measurements had been confined to well-cleaned, elevated surface regions, where laser interaction is most direct and effective, ΔE^* values would have fallen below 2.0 for all substrates. At the macroscopic scale, such regions may appear visually clean and homogeneous. However, contaminants trapped within fibrous pores or topographic recesses would remain undetected, leading to an overestimation of cleaning efficacy.

In colorimetric analysis, it is also important to evaluate the individual CIE L*a*b* coordinates (L^*, a^*, b^*), as they provide complementary information on whether the changes arise primarily from brightness (L^*), from red–green balance (a^*), or from yellow–blue balance (b^*). While ΔE^* summarizes the overall color difference in a single value, it does not indicate which component dominates the variation. In heterogeneous substrates such as leather, this distinction is particularly relevant because surface roughness and porosity can affect brightness and chromaticity differently across the irradiated area.

To illustrate this point, Table 3.4 represents L^*, a^*, and b^* values at the irradiation conditions that produced the lowest ΔE^* (closest to the perceptibility threshold). These examples highlight how the three coordinates contribute differently to the overall ΔE^* and reveal the role of substrate morphology in shaping the observed variability.

Table 3.4. CIE L*a*b* coordinates (L^*, a^*, b^*) are reported as mean ± standard deviation, while the total color difference (ΔE^*) is reported as value ± propagated uncertainty. All values correspond to 100-shot exposures.

Sample Name **Energy**	ΔE^*	L^*	a^*	b^*
Orange colored 130 μJ	2.63 ± 2.26	1.40 ± 1.10	-2.07 ± 1.67	1.03 ± 1.97
Havana colored 100 μJ	2.41 ± 1.71	-2.17 ± 1.51	0.30 ± 1.50	1.00 ± 2.46
Parchment 150 μJ	3.29 ± 1.79	-3.13 ± 1.83	-0.37 ± 0.26	-0.93 ± 1.35

As shown in Table 3.4, the analysis of individual CIE L*a*b* coordinates provides valuable insight into the origins of the observed color changes. For orange leather at 130 μJ/100 shots, ΔE^* was relatively low (2.63 ± 2.26), yet the underlying coordinates revealed a brightening trend (L^* positive) combined with a shift toward green (a^* negative), likely reflecting residual graphite contamination. At the same time, the yellow–blue component (b^*) fluctuated considerably, with some regions appearing more yellowish while others shifted toward blue. These variations are consistent with the spatial heterogeneity of the surface: measurements collected from protruding peaks, intermediate zones, and recessed pores inherently capture different local cleaning efficiencies. Havana leather at 100 μJ/100 shots also approached the perceptibility threshold ($\Delta E^* = 2.41 \pm 1.71$), but the decrease in lightness (L^* negative) together with wide fluctuations in a^* and b^* again reflects heterogeneous outcomes across zones of different topography. In contrast, parchment at 150 μJ/100 shots exhibited a somewhat higher ΔE^* (3.29 ± 1.79), but its chromatic coordinates remained comparatively stable, with the change dominated by a consistent darkening (L^* strongly negative). These examples demonstrate that leather substrates exhibit pronounced coordinate-level variability driven by porosity and micro-topography, whereas parchment shows more uniform responses. At higher pulse energies, where ΔE^* values increase, such fluctuations in L^*, a^*, and b^* become even more pronounced, reinforcing the importance of evaluating coordinate-level changes alongside total color difference.

This finding highlights the significance of high-resolution, spatially localized reflectance measurements in assessing the performance of fs-laser cleaning on porous and morphologically heterogeneous organic heritage materials. Area-averaged colorimetry, although informative, may obscure critical microstructural variations and lead to misleading conclusions about cleaning success. For conservation tasks where visual consistency and material integrity are paramount, localized optical evaluation is crucial to accurately assess cleaning outcomes and to prevent long-term preservation risks that may be hidden beneath apparent surface uniformity.

Although post-cleaning reflectance measurements revealed variability across the leather surface, one might question whether targeting smaller, well-cleaned regions using a reduced laser spot size could improve measurement consistency or even cleaning precision. In theory, focusing the laser beam to smaller spot sizes, approaching the diffraction limit, could enable more localized cleaning and better alignment with the heterogeneous pore distribution observed on the leather surface. The minimum achievable spot size is ultimately governed by the diffraction limit, which is dependent on the laser wavelength and the numerical aperture (NA) of the focusing optics. For a laser operating at 1030 nm, the Rayleigh criterion suggests that spot diameters in the range of 2–3 μm are physically attainable under ideal conditions using high-NA objectives. However, such a reduction in beam size poses several challenges in the context of conservation science.

First, as the spot size decreases, the area that can be treated per unit time shrinks significantly, leading to a dramatic increase in cleaning duration. For example, reducing the spot from 500 μm to 5 μm corresponds to a 10,000-fold decrease in beam area. While the fluence (J/cm^2) can theoretically be kept constant by proportionally reducing the pulse energy, most fs-laser systems have limitations in how low the energy can be tuned without compromising pulse stability or triggering nonlinear effects. Lower fluences might also be achieved by dumping part of the laser fluence; however, this approach would not be economically viable.

Second, the spatial resolution of the detection system must match the spatial resolution of the laser focus, which almost certainly implies a confocal setup, thereby removing ATR-FTIR from the compatible options. Raman spectroscopy would operate with equal or even larger selectivity, but the substantially lower cross-section of the Raman effect as compared to the dipole-excitation of FTIR would come at the expense of much longer detection times and slower processing.

While the system is currently operating at a convenient pulse repetition rate of 1 kHz, increasing the spatial resolution from 1 mm to 1 micron, for example, would require 1 million localized pulses, whereas a single, wide pulse was sufficient before. An increased spatial resolution would be detrimental to the processing time, and a meaningful balance between resolution and speed would have to be established under the pressure of economic viability.

Overall, although reducing the spot size offers theoretical advantages for selective cleaning and measurement consistency, practical constraints related to optical hardware, pulse energy stability, and processing time render this approach less suitable for bulk laser cleaning of historical collagen-based materials. Future work involving beam shaping or adaptive focusing methods may offer viable pathways to overcome these limitations.

3.4 Conclusion

Colorimetric analysis remains a widely used diagnostic tool in laser cleaning studies; however, its interpretive reliability becomes questionable when applied to porous and morphologically heterogeneous materials such as tanned leather. A representative example is the study by Elnaggar et al.[80], where color recovery after picosecond-laser cleaning was assessed using a colorimeter with a relatively large sampling aperture. Reported ΔE^* values consistently fell below the perceptibility threshold of 2.0, leading to the conclusion that cleaning was both effective and visually unnoticeable. However, such devices inherently average reflectance from both protruding and recessed zones within their field of view. Given the complex micro-topography of leather surfaces, this averaging effect may conceal localized overcleaning on peaks or insufficient cleaning in valleys.

In contrast, our measurements were performed with high spatial resolution ($\sim$500 μm), revealing substantial variability across replicate readings. Even under identical cleaning conditions, ΔE^* values spanned wide ranges, with standard deviations often exceeding 1.0 or more. This observation reflects the uneven interaction between the laser and the surface, where elevated regions receive higher fluence and may be overcleaned, while deeper pores remain partially shielded and retain soiling residues.

This spatial mismatch in laser–material interaction introduces a significant challenge: ΔE^* values obtained from point-based reflectance measurements may not faithfully represent the overall effectiveness or safety of the cleaning process. For instance, a low ΔE^* could simply result from a balance between heavily cleaned peaks and still-contaminated valleys, rather than indicating a uniformly successful treatment.

Among the tested substrates, parchment showed the most consistent colorimetric response with narrower error bars, attributed to its smoother and more homogeneous surface structure. In contrast, both Havana and Orange colored leather types exhibited pronounced reflectance variability that cannot be solely attributed to contamination distribution or laser parameters.

Additionally, this study highlights that not only overall porosity but also pore size distribution critically influences laser cleaning outcomes. The presence of both small and large pores creates complex surface topographies that exacerbate reflectance variability and limit cleaning uniformity, underscoring the need to consider microstructural characteristics when evaluating treatment efficacy.

These findings underscore the need to interpret colorimetric data in conjunction with morphological context. Reflectance-based evaluation is not intrinsically misleading, but when used in isolation, especially on textured surfaces, it risks oversimplifying or misrepresenting the outcome. Without surface-level diagnostics such as SEM, Raman imaging, or profilometric mapping, color recovery alone is insufficient to draw definitive conclusions about laser cleaning performance or material preservation.

While this study identifies spatial reflectance variability as a major interpretive challenge in fs-laser cleaning of porous substrates, it also points to the theoretical feasibility of addressing such heterogeneity through ultra-localized laser targeting. However, pursuing this path would significantly compromise operational efficiency. Due to the quadratic relationship between lateral resolution and processing time, achieving the necessary spatial selectivity to clean microstructural valleys without overexposing surface peaks would require prolonged treatment durations. As a result, while technically possible, such an approach may not be practically viable for large-scale or time-sensitive conservation tasks, highlighting the need for optimized trade-offs between precision and throughput in future developments of fs-laser cleaning protocols.

Chapter 4

Integrating Uncertainty Distribution into Pulsed Laser Cleaning Models: A Case Study on Paper Artifacts

Foreword

This chapter is based on the content of the following article; as a result, some passages may appear verbatim from this publication:

Integrating Uncertainty Distribution into Pulsed Laser Cleaning Models: A Case Study on Paper Artifacts, submitted, 2025.

C. Y. Boynukara, K. Kohlmann, P. Antici, and A. Ruediger

This work was led by C. Y. Boynukara, who developed the simulation framework, conducted the data analysis, and coordinated the overall research. K. Kohlmann contributed to the debugging and validation of the simulation code and supported the drafting of the manuscript. P. Antici reviewed the final manuscript and provided editorial feedback. A. Ruediger guided the overall research direction, advised on model development and interpretation, and contributed to the scientific framing of the study.

In the previous chapters, we established how the saturation and exponential decay model allows for the extraction of three particularly meaningful values: first, the saturation cleaning that might be achieved for individual, high fluence laser pulses. We also indicated that for these very high fluence laser pulses, the risk of irreversible damage through single- or multi-photon absorption was elevated, so that cleaning with a sequence of lower fluence pulses was a safer way. That bears the question about the lower fluence threshold, i.e., the minimum required fluence to accomplish a detectable change of either the artwork or the contamination signal. We demonstrated that this value is readily available through extrapolation of the experimental data towards lower fluence values. The third relevant parameter that can be extracted from the curves relates to the exponential decay curve where the residual signal of the pollutant, the signal that might not even be removed for very high laser fluences, represents a quantification of a particular phenomenon related to laser cleaning: laser ablation can be modelled as a local explosion of material upon exposure to high fluence pulses. As a result, most of the material will be ablated into the open space above the sample, but some of it might be propelled into voids of porous materials such as paper and parchment. While we showed how these values can be extracted through a numerical fitting procedure, we have so far ignored the experimental uncertainties that come with every single experiment. They relate to uncertainties of the fluence (that we shall ignore here again, as it only depends on the skills of the experimentator to establish these values with the desired precision) and to uncertainties of the detected signal. These uncertainties have different origins: each signal has statistical noise of, e.g., the detector; then there is statistical noise from sample heterogeneities that were addressed in this thesis through averaging over signals from three different, neighbouring locations of the sample. To this, depending on the experimental setup, it might be required to add additional noise sources through, e.g., laser instabilities and parasitic light sources. All these sources of noise have one thing in common: they add uncertainty to the extrapolated value from the numerical fitting procedure. This chapter, therefore, addresses in a statistically rigorous way how the experimental uncertainties propagate and how reliable the extrapolated values really are. The authors of the respective manuscript are convinced that knowledge of these uncertainties is a cornerstone of a responsible cleaning approach in order to avoid jeopardizing the physical integrity of cultural heritage.

Abstract

Laser cleaning is widely used in cultural heritage conservation, yet its evaluation often prioritizes visual metrics such as colorimetry, which offer limited chemical specificity and typically lack a systematic approach to uncertainty quantification. This study proposes a model-based framework to assess pulsed laser cleaning of paper artifacts, integrating physical modeling and uncertainty quantification. An exponential saturation and exponential decay model were fitted to ATR-FTIR data to describe cleaning dynamics. Monte Carlo simulations were used to propagate parameter variability and experimental noise, generating confidence intervals around predicted cleaning trajectories. The results show that uncertainty expands symmetrically at low fluence and at high fluence. This modeling approach identifies critical fluence zones where cleaning is both effective and predictable. By visualizing uncertainty and modeling both contaminant and substrate signals, this approach supports more informed, risk-aware decision-making and overcomes limitations of traditional evaluation methods.

4.1 Introduction

Laser cleaning has become a critical technique in cultural heritage conservation, offering precise, selective, and non-contact removal of contaminants from delicate historical artifacts. It provides significant advantages over conventional mechanical and chemical methods [50, 51]. Over the past five decades, various research groups have explored the effects of different wavelengths, pulse durations, and pulse sequences to optimize the removal of unwanted layers and to expand the applicability of laser techniques in preservation [49, 52, 57]. As automation advances, the need for quantitative assessment of the cleaning progress grows, but this must be accompanied by even greater care to protect the artwork, such as preventing laser-induced damage.

While femtosecond lasers offer the clear advantage of surface ablation with often negligible heat transfer into the sample compared to nanosecond lasers, the optimal wavelength selection remains less straightforward. As they are used in dental and eye surgery, UV lasers can be considered as optical scalpels due to their very small penetration depth and low selectivity to the absorbing surface. They enable progressive ablation of surface contaminants, typically removing only a few tens of nanometers per pulse. This fine control is not time-limiting, given the high repetition rates of UV sources. However, exposing cellulose-based materials to UV irradiation often causes yellowing, which restricts the suitability of UV lasers for certain applications.

Femtosecond infrared lasers offer a promising alternative to overcome this limitation, though they introduce other considerations. Due to the significantly greater penetration depth of IR radiation compared to UV, it can be expected that: (a) energy is deposited over a thicker layer, thereby increasing the ablation rate per pulse; and (b) even the first laser pulse may interact with underlying substrate layers beneath the contamination. This interaction tends to intensify as the absorbing overlayer is progressively thinned during cleaning. Therefore, it should be ensured that none of the substrate materials exhibit strong intrinsic absorption in the IR range.

The verification of cleaning requires appropriate signals for both the contamination to be removed as well as the sample to be restored. Ideally, during laser cleaning, the signal associated with the contaminant should gradually decrease and eventually vanish, while the signal originating from the underlying sample surface should increase and reach a stable plateau. If this dual behavior is not observed, that is, if the contaminant signal does not significantly decline or the sample signal fails to stabilize, the cleaning can be considered incomplete. Conversely, if the sample signal begins to decrease over time, this may indicate that the laser is no longer selectively removing the contaminant but instead affecting the substrate itself, marking a transition toward a destructive process.

In order to establish a quantitative signal from contaminants and sample surfaces, a wide range of non-destructive analytical techniques have been employed in the evaluation of laser cleaning processes, including Optical Coherence Tomography (OCT), Fourier-transform infrared spectroscopy (FTIR, including Attenuated Total Reflectance mode), Raman spectroscopy, UV/VIS/NIR spectroscopy, colorimetry, and, in certain cases, low-vacuum mode SEM/EDS, though the latter's non-invasiveness is occasionally debated. These techniques offer complementary insights into cleaning effects: OCT reveals surface morphology and subsurface features; FTIR and Raman spectroscopy provide chemical and molecular information; UV/VIS/NIR spectroscopy detects changes in optical properties; and colorimetry quantifies visual differences through colorimetric metrics [68, 71, 81, 82, 83, 84]. Collectively, these methods help assess both cleaning efficacy and the preservation of original material integrity.

For the sake of quantification, a signal should be proportional to the quantity of the material under investigation. This is, for example, the case for IR absorption bands or Raman emission lines, both of which are proportional to the quantity of molecular oscillators in a specific, user-defined volume. This is in strong contrast to the use of colorimetry, which has become one of the most widely adopted non-invasive tools for evaluating laser cleaning efficacy due to its accessibility and ease of use.

In absorption studies, the intensities of individual absorption bands linearly scale with the quantity of absorbers under otherwise identical conditions (interaction volume of light and sample, temperature, humidity, etc.), so that a direct correlation between this signal and the progress of cleaning may be established. The same is true for the intensity of Raman peaks, even though the experimental complexity to ensure the "otherwise identical conditions" might be challenging in configurations such as attenuated total reflection (ATR) or a confocal set-up. Colorimetry is based on a metric that averages over a specific spectral range that might include signatures from both the contaminants as well as the sample surface (at this point, we should also point out the possible presence of layers of varnish that may or may not need to be preserved, but that play a role for future considerations). The metric of colorimetry, based on the specific color scheme that is used, is that of a Pythagorean distance in a color map, where the respective change of color is calculated with equation 4.1:

$$\Delta E = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \tag{4.1}$$

Here, L refers to brightness, while a and b refer to spectral features, usually averaged over a given range. Not only do L, a, and b potentially contain contributions from both the contaminants and the underlying sample, but changes in material concentration are not guaranteed to produce linear shifts in these parameters. Consequently, a progressive cleaning protocol, such as one employing consecutive laser pulses, may yield colorimetric values that do not reflect a linear trend, and may not even behave monotonically.

For the remainder of this manuscript, we assume access to a signal that is either directly proportional or can be reliably correlated to the quantity of the respective signal source. The critique of colorimetry outlined above does not exclude its use, but leveraging it effectively would require more sophisticated analytical strategies, such as machine learning, to link observed colorimetric responses to the actual state of the sample surface.

Furthermore, the consistency of colorimetric measurements is strongly influenced by the geometry of the measurement device and the specific acquisition protocols, which makes cross-instrument comparability inherently difficult [85]. Point-based colorimetric readings often fail to reflect the full spatial variability present on heterogeneous surfaces, leading to partial or misleading evaluations [86]. Automated color profiling methods, such as those based on the Munsell system, are also sensitive to ambient lighting conditions and the observer's viewing angle, emphasizing the need for standardized measurement protocols [87]. In addition, the absence of well-established reference standards specifically designed for cultural heritage materials further complicates the accurate interpretation of color changes and degradation phenomena [88].

Accurate and reliable assessment of laser cleaning outcomes is thus essential, as inappropriate cleaning or misinterpretation of cleaning effectiveness can lead to irreversible damage or inadequate preservation of culturally valuable items. This underscores the necessity for robust analytical methods capable of quantifying not only the effectiveness of the cleaning process but also the uncertainty associated with the evaluation itself.

The least invasive laser cleaning protocol is based on subsequent laser pulses that all have sufficient fluence (energy per area) to accomplish a certain degree of material ablation, i.e., the fluence must exceed a certain minimum threshold. This is advantageous over the use of the integrated fluence in a single pulse, which not only provides less control but also increases the risk of inflicting damage to the sample surface. For that matter, two thresholds of laser fluence should ideally be established: a lower threshold to determine the lowest fluence that creates a tangible cleaning effect at all, and a higher threshold that establishes either full cleaning with a single shot or the damage threshold of the sample, whichever value is smaller. While the determination of such thresholds is undoubtedly important, we do not address it here, since it lies beyond the primary focus of this study.

The signals of contamination and of the sample to be restored will follow opposite trends. With each laser pulse, the contamination signal will further decrease, while the signal of the sample surface to be restored should increase. Both trends do not have to be necessarily proportionate, as the contamination layer may have varying thickness, the exposure of the underlying surface may not immediately result in strong signal changes, and saturation or overlapping spectral features can distort the linearity between the two.

We base our analysis on an exponential saturation model for the restoration of the sample signal and an exponential decay model for the signal of the contaminant. These models capture the asymptotic behavior commonly observed in laser cleaning: cleaning efficiency increases rapidly at lower fluences, where loosely bound contaminants are easily removed, and then gradually levels off as remaining residues become more chemically or physically resistant. The model assumes that cleaning progress slows with each pulse as absorption by the contaminant decreases due to prior exposure. Unlike linear models, which assume constant material response, the exponential form reflects diminishing returns per pulse and provides a more realistic, conservative estimate of cleaning efficiency.

Fitting this model to spectral data, specifically obtained through ATR-FTIR, allowed us to describe the material response as a function of fluence in a structured and interpretable way. However, since most laser setups provide only a limited range of fluence values, threshold behavior is not always directly observable and often requires extrapolation. Because all experimental data points carry some degree of uncertainty, the extrapolated predictions must also incorporate this variability. To address this, we extended the model with a Monte Carlo-based uncertainty quantification approach. First, the model was fitted to data from one representative surface type. Then, Gaussian-distributed noise was added to the original spectral data based on measured standard deviations, and the model was re-fitted to each perturbed dataset. This process was repeated over multiple iterations to generate a distribution of plausible parameter sets and associated predictions. The resulting ensemble was used to compute confidence intervals for cleaning efficiency, providing a more robust and reproducible method to evaluate laser cleaning performance.

4.2 Modeling and Uncertainty Quantification

We will begin our discussion with the signal of the surface to be restored, acknowledging that it may not always be possible to track both the contamination (or any of the contaminants) and the signal of the surface to be restored simultaneously.

The exponential saturation model and its application to ATR-FTIR data were previously introduced and validated in Chapter 2. For completeness, we briefly summarize the model structure here before presenting the uncertainty quantification framework developed in this study.

To describe the nonlinear relationship between laser fluence and cleaning efficiency, we employed a four-parameter exponential saturation model defined by equation 4.2:

$$y(x) = a \cdot \left(1 - e^{-\frac{(x-b)}{c}}\right) + d \tag{4.2}$$

Mathematical interpretation

From a mathematical perspective, each parameter shapes the curve in a specific way:

- a defines the vertical stretching factor of the function, determining the total amplitude of the signal.
- b is a horizontal offset that shifts the curve along the fluence axis, effectively defining the fluence threshold.
- c is a horizontal scaling factor that controls how quickly the curve approaches saturation, i.e., the steepness of the exponential rise.
- d represents the initial signal level before laser exposure, reflecting how much of the substrate's spectral contribution is already detectable despite the presence of the contamination layer. Lower values indicate more opaque or strongly absorbing contamination, which prevents the underlying material from contributing to the measured signal.

These parameters allow the model to flexibly fit experimental data, capturing the typical asymptotic behavior observed in laser cleaning, where contaminant removal increases rapidly at first and then levels off.

Physical interpretation

Physically, each parameter carries meaning related to the laser–material interaction and the detection setup:

- a represents the maximum signal achievable under ideal cleaning conditions. It may also reflect the sensitivity of the spectroscopic system—larger values indicate stronger signals and potentially lower noise.
- b corresponds to the minimum fluence required to initiate a detectable cleaning effect. It is associated with the ablation threshold.
- c describes the fluence-dependent efficiency of contaminant removal. It reflects how energy absorption translates into ablation; smaller values indicate faster saturation.
- d represents the signal before laser exposure, indicating the extent to which the substrate is initially visible through the contaminant layer. Lower values suggest heavier soiling or more opaque contamination.

During numerical fitting, the algorithm seeks the parameter set that minimizes the difference between the fitted curve and the original data. However, mathematically optimal parameters may not always be physically meaningful (e.g., negative signal values or unrealistically low thresholds). Therefore, parameter boundaries were constrained based on physical plausibility to ensure interpretable and reliable modeling. Specifically, the following constraints were applied:

- $a \in [0, \infty]$: This parameter reflects the signal amplitude in arbitrary units. Its lower bound is zero, and the upper bound is theoretically unconstrained, consistent with potential amplification in the detection system.
- $b \in [0, 0.39]$: Since laser fluence is strictly positive, the lower limit of b was set to zero. Its upper bound was set to 0.39 J/cm^2, which corresponds to the lowest fluence at which a measurable signal was experimentally observed. This ensures that the model does not predict cleaning onset at fluence values beyond the range of collected data.
- $c \in [0, 0.2 \cdot \Delta x]$: The steepness parameter c was bounded above by 20% of the total experimental fluence range Δx. Higher values would inhibit the model from reaching saturation within the accessible fluence window.
- $d \in [0, \infty]$: The offset parameter d accounts for the initial spectroscopic signal prior to laser exposure. As it is derived from integrated absorbance, it is physically required to remain positive.

Two physically meaningful values can be directly extracted from the exponential saturation model fitted to experimental data.

First, for very large laser fluence values, the exponential term in equation 4.2 approaches zero, and the signal $y(x)$ asymptotically converges to $a + d$:

$$\lim_{x \to \infty} y(x) = a + d$$

This value represents the *maximum recoverable signal* and corresponds to the best possible cleaning scenario achievable by intense single-pulse irradiation. However, it is important to note that such high fluence levels may not be ideal in practice, as they may cause irreversible damage to the artwork through linear and nonlinear absorption processes. The goal is not only efficient cleaning but also preservation of the material integrity.

Second, the model allows for an estimation of the *minimum laser fluence required to initiate a detectable cleaning effect.* This can be obtained by solving the model equation for the fluence x when the output signal $y(x) = 0$, i.e., when the signal begins to deviate from the baseline:

$$0 = a \cdot \left(1 - e^{-\frac{(x-b)}{c}}\right) + d \Rightarrow x = b - c \cdot \ln\left(\frac{a+d}{a}\right)$$

This expression shows that the minimum effective fluence is primarily determined by parameter b, which defines the fluence threshold where cleaning begins. The logarithmic correction term becomes small when $a > d$, a condition that was consistently observed in our experiments. In such cases, the detected signal before cleaning d is negligible compared to the signal after cleaning $(a + d)$, indicating that the surface is initially heavily contaminated and the effect of cleaning is significant. On the contrary, if d were close to a, this would suggest that the signal is already strong even before cleaning, and further laser exposure would have little effect, making the process unnecessary or inefficient.

Altogether, this analysis offers two critical insights: (1) the best achievable signal recovery through intense cleaning, and (2) the minimum required fluence for initiating cleaning. The latter is particularly relevant when using high-repetition-rate pulsed lasers, where lower fluence can be compensated by pulse accumulation without significantly increasing cleaning time.

This formulation captures the asymptotic behavior commonly observed in laser cleaning, where cleaning efficiency increases rapidly at first and then levels off as remaining contaminants become less accessible or more strongly bound. Unlike linear models, this approach accounts for the physical limits of contaminant removal and avoids overestimating cleaning effects at high fluence levels.

Model fitting was carried out using non-linear least squares optimization in MATLAB. Initial parameter values were estimated through visual inspection of the data, and the quality of the numerical fit was evaluated using metrics such as R^2, RMSE, and MAE. The fitted parameters allowed us to characterize the material response across the applied fluence range. To ensure physical interpretability, parameter boundaries were constrained as described above.

To extend the deterministic model into a probabilistic framework, we implemented an uncertainty quantification strategy based on Monte Carlo simulation. The goal was to assess how uncertainties in model parameters propagate through the exponential saturation model and influence the predicted cleaning outcomes, particularly the estimation of lower and upper thresholds for complete cleaning, across a range of fluence values.

To characterize the variability of the fitted model parameters, a, b, c, and d, we performed Monte Carlo simulations individually for each dataset. Gaussian noise was added to the spectral data based on experimentally measured standard deviations, and the model was re-fitted to each perturbed dataset. This procedure was repeated 100 times per dataset, resulting in an ensemble of plausible parameter sets derived from that dataset alone. These empirical parameter distributions were then used to generate corresponding predictions and confidence intervals for laser cleaning performance across the fluence range.

A total of 100 Monte Carlo samples were drawn by independently sampling each model parameter from a normal distribution based on its fitted mean and standard deviation. For each sampled parameter set, the model was used to compute a predicted cleaning trajectory over a defined fluence range. Additionally, to reflect experimental noise, synthetic datasets were created by perturbing the original spectral data, where the error bars were directly calculated as the standard deviation of the integrated band area from repeated ATR-FTIR measurements at each fluence level. Fitting was repeated for each perturbed instance.

This simulation yielded an ensemble of predicted outcomes that allowed us to calculate confidence intervals and visualize the uncertainty associated with the cleaning process. Rather than relying on a single best-fit curve, this method captures the variability and robustness of predictions, providing a more comprehensive and realistic assessment of laser cleaning efficiency.

4.2.1 Noise Modeling and Data Perturbation

In addition to sampling parameter sets from independent normal distributions, we also simulated experimental variability by introducing Gaussian noise into the original dataset. This was done to replicate the natural fluctuations observed in spectroscopic measurements due to instrumental uncertainty and sample inhomogeneity.

Using the standard deviation values associated with each fluence point (as estimated from repeated measurements), synthetic datasets were created by perturbing the original ATR-FTIR signal using zero-mean normally distributed noise scaled by the experimental error bars. Specifically, for each Monte Carlo iteration, a new dataset was generated using equation (4.3):

$$A_i^{\text{noisy}} = A_i^{\text{measured}} + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_i) \tag{4.3}$$

where A_i^{measured} is the original band area at fluence i, and σ_i is the corresponding measurement error (standard deviation).

Each noisy dataset was then re-fit using the exponential saturation model, starting from the baseline parameter set (i.e., corn starch), with the same fitting constraints and algorithm used in the original fit. This process was repeated across 100 Monte Carlo iterations, resulting in a collection of model outputs that reflected not only parameter variability but also the influence of experimental noise on the fit outcome.

4.2.2 Ensemble Prediction and Confidence Interval Construction

Ensemble Prediction and Confidence Interval Construction

For each fluence value within the defined range, predicted signals from all 100 Monte Carlo iterations were collected to form an ensemble of outcomes. The pointwise mean of these predictions was used to define the average cleaning trajectory. To quantify the uncertainty associated with these predictions, the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles of the ensemble at each fluence point were computed to establish a 95% confidence interval.

This percentile-based confidence band captures the combined effect of parameter variability and stochastic measurement noise introduced during the Monte Carlo sampling process. Its interpretation and relevance to model robustness are discussed in the following section.

4.3 Results and Discussion

The modeling framework described above was applied to the ATR-FTIR dataset to evaluate the predictive capability of the exponential saturation model and to quantify the uncertainty associated with the cleaning process. This section presents the results of both the initial model fitting and the Monte Carlo simulation, highlighting the performance of the fit, the variability of the parameters, and the resulting confidence intervals.

4.3.1 Model Fit and Performance Metrics

The exponential saturation model was successfully fitted to the ATR-FTIR dataset corresponding to the corn starch-sized sample, as illustrated in Figure 4.1. The model captured the nonlinear trend of the cleaning response with respect to laser fluence, showing a rapid increase in the signal area followed by a gradual saturation, consistent with the expected physical behavior of contaminant removal.

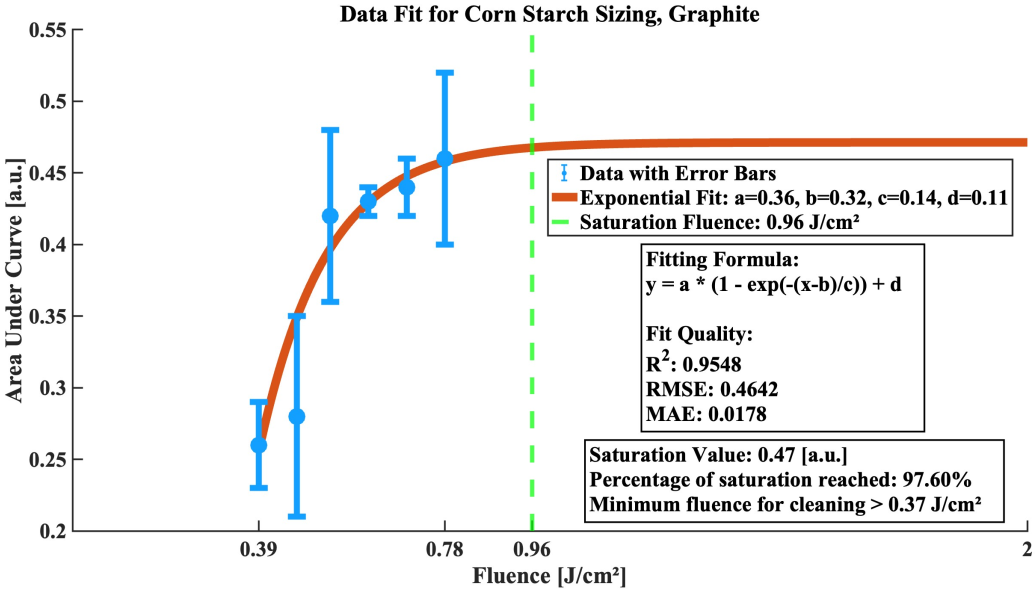


Figure 4.1. Exponential saturation model fit for corn starch sizing with graphite contamination for fs laser cleaning.

The fitting process yielded a coefficient of determination (R^2) of 0.9548, indicating a strong agreement between the model and experimental data. Additional goodness-of-fit metrics included a root mean square error (RMSE) of 0.4642 and a mean absolute error (MAE) of 0.0178. These results suggest that the exponential saturation model provides an acceptable description of the laser-induced cleaning dynamics in this case.

The estimated model parameters were as follows:

- $a = 0.36$,
- $b = 0.32$,
- $c = 0.14$,
- $d = 0.11$.

Using these parameters, the saturation fluence, which we here defined as the fluence level at which 99% of the maximum response is reached, was calculated according to the relation:

$$F_{\text{sat}} = b - c \cdot \ln(0.01) \tag{4.4}$$

The computed saturation fluence was 0.96 J/cm^2, providing a quantitative estimate of the fluence threshold required for near-complete cleaning with a single laser pulse. For all the aforementioned arguments, this may not necessarily be the best option in terms of process control and sample safety.

In addition to the saturation model, an exponential decay model was applied to a dataset corresponding to corn starch sizing contaminated with graphite and kaolinite mixture (Figure 4.2). This model describes cases where the contaminant signal decreases with fluence but asymptotically approaches a nonzero baseline due to intrinsic material or process limitations.

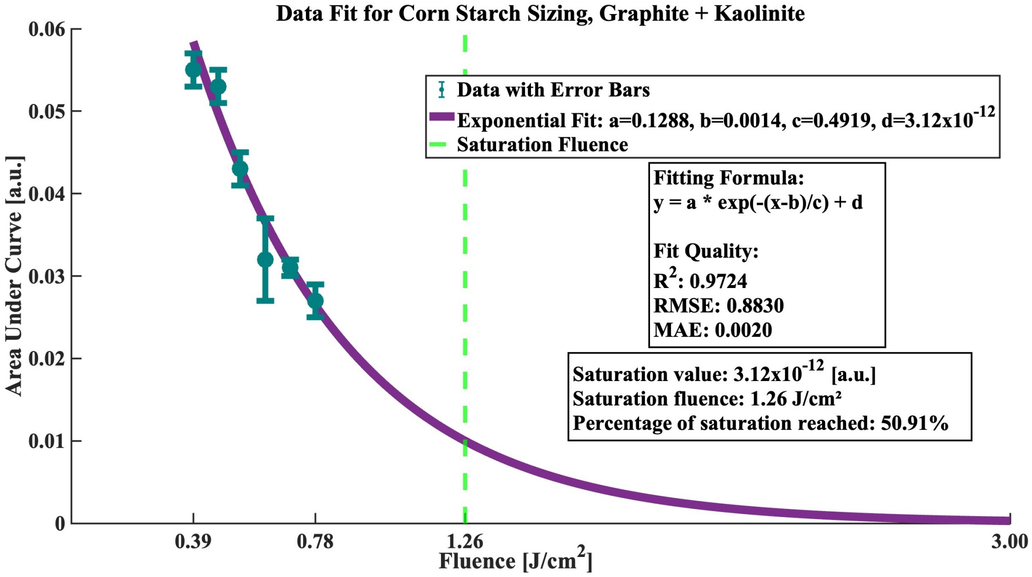


Figure 4.2. Exponential decay fit for corn starch sizing contaminated with graphite and kaolinite.

The high coefficient of determination ($R^2 = 0.9724$), low RMSE (0.8830), and MAE (0.0020) confirm the model's ability to represent the experimental trend. In this case, the baseline offset d was clearly nonzero, indicating the persistence of residual contamination even at higher fluence levels. The fluence value corresponding to 99% decay was computed using the same relation as before, yielding a saturation fluence of 1.26 J/cm^2.

This behavior contrasts with the typical saturation curve, as the contaminant signal decreases rather than increases with fluence. The presence of a nonzero asymptote in the decay model may reflect physical limitations in contaminant removal, such as dust particle trapping or forward scattering effects that prevent complete elimination of the material.

4.3.2 Parameter Variability and Monte Carlo Results

Figure 4.3 presents the results of Monte Carlo simulations performed on the exponential saturation model, illustrating how parameter variability and experimental noise propagate into model predictions. In this figure, the red curve represents the best-fit exponential saturation model based on the original dataset, serving as the deterministic reference. The blue dashed curve indicates the ensemble mean prediction derived from 100 Monte Carlo iterations, each involving randomized perturbations of the original spectral slopes. The shaded yellow region corresponds to the 95% confidence interval constructed using the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles of the ensemble predictions at each fluence value. Experimental data points are plotted with vertical error bars to indicate the uncertainty in measured signal values. Additionally, a vertical green dashed line denotes the estimated saturation fluence, calculated as $0.96\,\mathrm{J/cm^2}$, and a horizontal red line marks the asymptotic saturation level of $0.47\,\mathrm{a.u.}$ reached by the exponential fit.

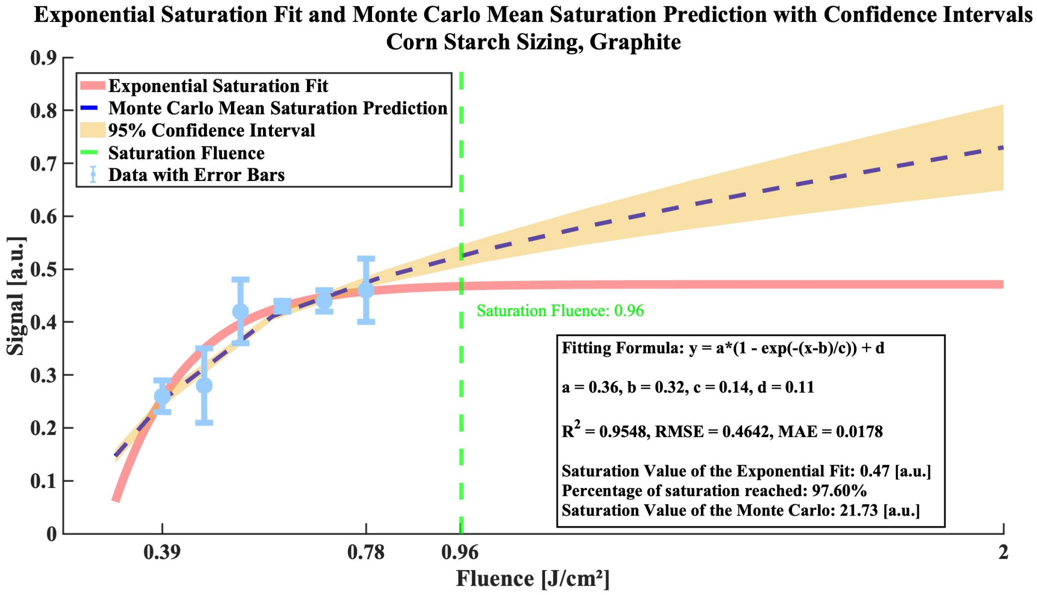


Figure 4.3. Monte Carlo simulation results for the exponential saturation model applied to corn starch sized paper contaminated with graphite. The red curve represents the best-fit deterministic model, while the dashed blue curve corresponds to the ensemble mean prediction across 100 Monte Carlo realizations. The shaded yellow region denotes the 95% confidence interval (percentile-based), capturing both experimental noise and parameter uncertainty. Experimental data are shown as circular markers with error bars. The vertical green dashed line indicates the saturation fluence ($F_{\mathrm{sat}} = 0.96\,\mathrm{J/cm^2}$), and the horizontal red line marks the saturation value of the deterministic fit ($0.47\,\mathrm{a.u.}$). The widening of the confidence band at higher fluences reflects increased model uncertainty and reduced predictive stability beyond the measured data range.

Several key observations emerge from the figure:

- At low fluence values (0.39–$0.54\,\mathrm{J/cm^2}$), where experimental data are dense and the signal increases rapidly, the confidence band remains narrow. This indicates strong agreement across model realizations and low uncertainty in predicted signal behavior.
- In the intermediate fluence range (0.54–$0.78\,\mathrm{J/cm^2}$), the confidence band starts to widen, reflecting increased sensitivity to parameter fluctuations—particularly the stretching parameter c.
- Beyond $0.78\,\mathrm{J/cm^2}$, the ensemble trajectories diverge significantly, with the upper bound of the confidence interval increasing steeply. The widening of the band suggests reduced model certainty and greater risk of overestimation in extrapolated regions.

The shape of the confidence band is not uniform across the fluence range and carries important physical and statistical implications. At low fluence levels, the narrow confidence band indicates strong data support and high sensitivity of the signal to small fluence changes. As fluence increases, the saturation curve flattens, and the confidence band widens significantly. This widening reflects the reduced sensitivity of the signal to fluence variations in the saturated regime, as well as the growing influence of parameter variability. The increasing uncertainty at high fluences highlights the risks associated with extrapolating beyond the measured data range and emphasizes the need for caution when interpreting predicted cleaning performance under these conditions.

While Figure 4.3 illustrates only the asymptotic saturation level of the exponential fit (0.47 a.u.), it is important to note that the Monte Carlo ensemble yielded a significantly higher saturation value of 21.73 a.u. This discrepancy stems from the fact that the ensemble prediction reflects the average outcome across 100 model realizations, each influenced by randomized perturbations of the original spectral data. In some of these iterations, particularly those in which the parameter c assumed high values, the saturation curve was excessively stretched along the fluence axis, resulting in slow signal growth and extremely delayed saturation behavior.

Although mathematically valid, this extrapolated saturation behavior is not physically meaningful. Achieving signal levels near 21.73 a.u. would require operating at fluence levels well beyond what is considered safe for fragile historical materials. Such high exposures could compromise the physical integrity of the substrate and increase the likelihood of nonlinear interactions such as multiphoton absorption.

The parameter c, which governs the slope of the saturation curve, plays a critical role in this behavior. Large values of c imply a low rate of signal change with increasing fluence, corresponding to inefficient contaminant removal. This situation may arise when the contaminant has poor absorption at the laser wavelength (1030 nm), rendering even high-energy laser pulses ineffective.

To address this issue, we imposed an upper bound constraint on c during the Monte Carlo simulations. This constraint ensures that saturation behavior remains observable within the experimentally tested fluence range and that the resulting predictions are both numerically stable and physically plausible. Without such a limitation, the model risks producing extrapolations that are mathematically consistent but incompatible with the actual behavior of the system.

To ensure that the uncertainty propagation remains physically meaningful, an upper bound was imposed on the steepness parameter c during the Monte Carlo simulations. This constraint was derived from the general condition $c \in [0, 0.2 \cdot \Delta x]$, where Δx corresponds to the experimental fluence range. In this case, since the simulations extended up to 2 J/cm^2, the upper limit was set at 0.4. The rationale behind this choice was to prevent the generation of saturation curves that are excessively stretched along the fluence axis, which would otherwise fail to reach a plateau within the experimentally accessible range. Such behavior, although mathematically permissible, leads to physically implausible predictions and diminishes the model's interpretability.

Figure 4.4 illustrates the outcomes of the bounded simulation, showing several

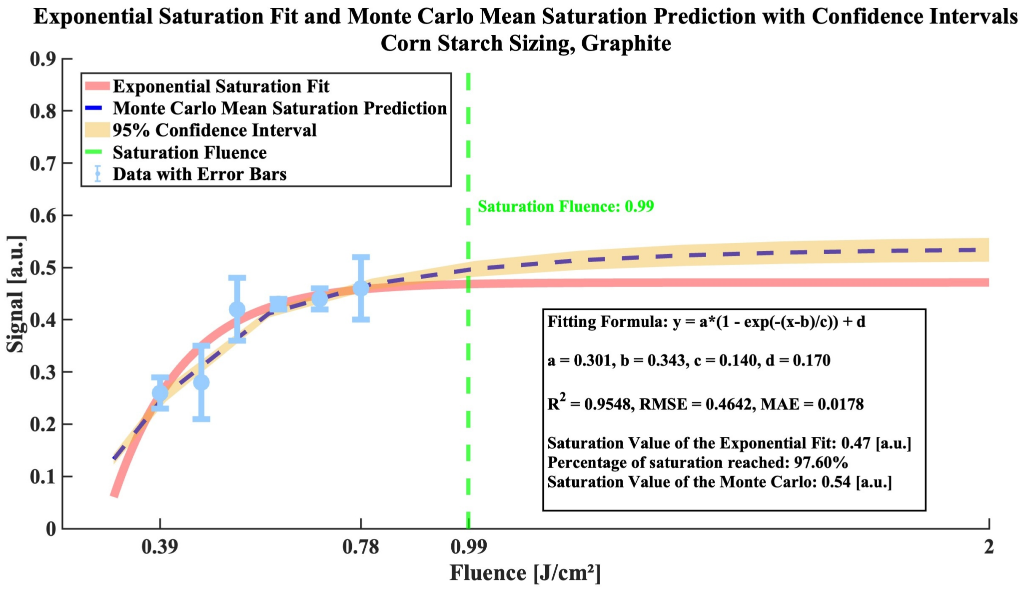


Figure 4.4. Exponential saturation fit and Monte Carlo mean prediction with bounded c parameter (Corn Starch Sizing, Graphite). The Monte Carlo ensemble used an upper bound of $c = 0.4$, corresponding to 20% of the total fluence range (2 J/cm^2).

notable improvements over the previous unconstrained case. The Monte Carlo ensemble mean trajectory (blue dashed line) now aligns more closely with the best-fit exponential saturation curve (red line), and the extrapolated saturation value of the ensemble (0.54 a.u.) is in much better agreement with the asymptotic saturation level of the deterministic model (0.47 a.u.). This indicates that the simulated predictions remain within a realistic operating range and no longer suggest unfeasibly high signal values.

The 95% confidence interval also exhibits a narrower and more symmetric band across the fluence domain, particularly in the high-fluence region. This suggests reduced extrapolation uncertainty and higher predictive stability. Unlike the previous scenario, where the upper edge of the confidence interval rose sharply beyond the last data point, the current simulation displays a confidence band that gently levels off toward the saturation plateau.

Importantly, this improvement does not come at the cost of model flexibility. The parameter ensemble still captures natural variability in the signal response, particularly in the mid-fluence range (around 0.54–0.78 J/cm^2), where sensitivity to parameter fluctuations is highest. However, by limiting c, we ensured that all ensemble trajectories eventually reach saturation within the fluence window, preserving physical consistency and enabling more reliable threshold estimation.

In summary, the application of an upper bound to the parameter c enhances both the numerical robustness and physical plausibility of the Monte Carlo simulations. The model is now better suited to generate realistic predictions that are applicable to fragile heritage substrates, for which high-fluence exposure is neither feasible nor safe.

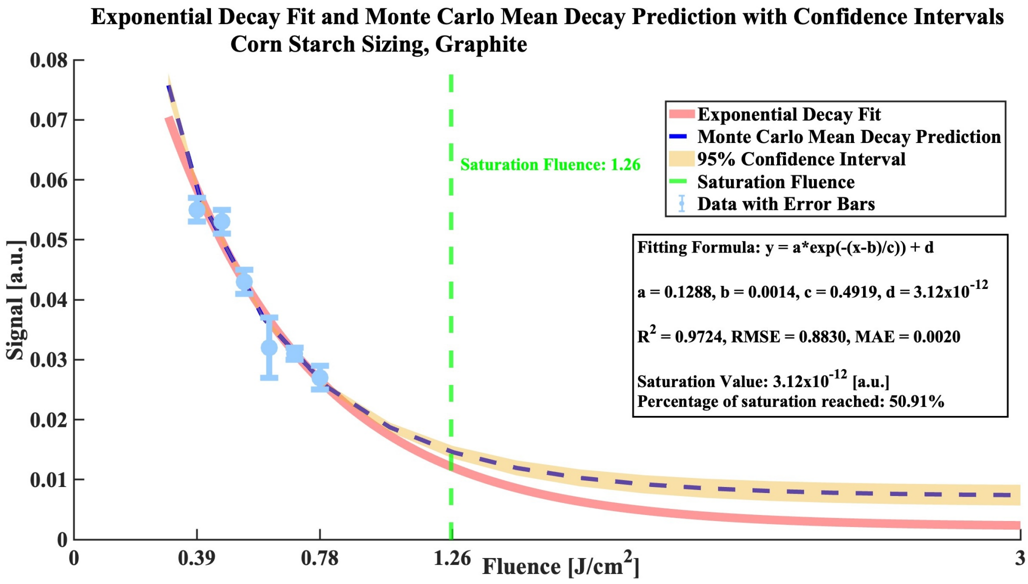


Figure 4.5. Monte Carlo simulation results for the exponential decay model applied to corn starch sizing paper contaminated with graphite. The red curve represents the best-fit deterministic model, while the dashed blue line indicates the ensemble mean from 100 simulations. The yellow band denotes the 95% confidence interval.

As illustrated in Figure 4.5, the deterministic exponential decay fit and the Monte Carlo ensemble mean prediction show excellent agreement within the measured fluence range. This alignment supports the physical plausibility of the model and reinforces its predictive reliability.

The 95% confidence interval remains relatively narrow across the experimental domain, indicating stable model behavior and low extrapolation uncertainty. The band broadens slightly beyond the last data point, which is expected as prediction reliability decreases in regions lacking experimental constraints.

The calculated saturation fluence of 1.26 J/cm^2 corresponds to the point at which the contaminant signal has decreased by 99% relative to its initial value. However, a small fraction ($\sim$1%) of residual contamination remains, likely due to mechanical embedding into fibrous microstructures. These deeply trapped residues cannot be removed efficiently, as increasing the fluence beyond this point risks burning or irreversibly damaging the paper substrate.

This analysis confirms that the exponential decay model, enhanced with uncertainty quantification, offers a robust and physically meaningful approach for evaluating contaminant removal and safe operating limits in laser cleaning.

In contrast to traditional methods such as colorimetry, which are often influenced by lighting conditions and lack molecular sensitivity, the approach presented here quantifies not just the extent of cleaning but also the confidence in its reproducibility. By generating a confidence band around the predicted cleaning trajectory, our model provides a more nuanced view of treatment outcomes, particularly in cases where visual appearance alone may not capture underlying chemical or morphological changes, such as sub-surface residues or altered material structures.

Monitoring both the signal of the contaminant and the restored surface is also crucial for reliable evaluation. Tracking only the contaminant may overlook laser-induced damage to the substrate, while focusing solely on the substrate signal could miss residual contamination. Such residuals may result from unwanted migration into the substrate, which can accelerate degradation or alter the perceived appearance of the artifact.

We would like to point out that the interpretation of the signal originating from the artwork (saturation curve) will almost always be a more useful indicator of the cleaning process as compared to the signal of the pollutant. Our saturation model is justified through the law of Lambert-Beer under the assumption that each laser pulse with a given fluence will remove the same amount of thickness and that the amount of removed thickness scales linearly with the laser fluence. While there will almost always be a detectable signal of the artwork through diffuse scattering in the ATR-FTIR volume, the initial signal of the pollutant may not actually change as long as the total thickness of the layer does not drop below the penetration depth of the ATR-FTIR setup.

Nevertheless, the proposed method presents certain limitations. The current Monte Carlo framework assumes independent parameter sampling, which may not fully reflect interdependencies observed in real experimental settings. While three representative datasets were used to define plausible parameter ranges, these may not encompass the full spectrum of materials and contamination conditions encountered in conservation practice. Expanding the parameter base and integrating complementary diagnostic techniques, such as optical coherence tomography (OCT) or low-vacuum SEM/EDS, could further enhance the interpretive power of the model and improve its generalizability across heritage substrates.

For the case of the saturation model, from which we extract the asymptotically possible recovered signal of the artwork and the lowest possible laser fluence to accomplish any cleaning at all, we should now respond to the question about the sensitivity of our extrapolated values to these four fitting parameters. Which is the parameter (they all have a physical meaning, see above) that our predictions are most sensitive to? This leads us to a comforting finding: The saturation value of the recovered signal only depends on the sum of parameters a and d. So the signal strength that the experiment will reproduce upon complete cleaning will only depend on the amplification of the detector and its possible offset. After complete cleaning (i.e. at saturation), neither the absorption of the pollutant nor the lowest required energy should play a role. As for the determination of the lowest threshold of laser fluence (equation 2.2), the situation only appears to be more complicated. In essence, the value is determined by the single fitting parameter b and all the remaining term will be a minor correction at most.

Overall, by combining physical modeling with statistical simulation, this work contributes to a more rigorous and risk-aware approach to laser cleaning in heritage science. The ability to quantify and visualize uncertainty not only enhances predictive accuracy but also enables more informed and conservative decision-making in the context of cultural heritage conservation.

4.4 Conclusion

This study introduced a model-based framework for evaluating laser cleaning performance on paper artifacts, emphasizing the integration of uncertainty quantification into the analytical process. By applying a four-parameter exponential saturation model to ATR-FTIR data and extending it with Monte Carlo simulations, we captured both the nonlinear dynamics of contaminant removal and the variability inherent in experimental and material conditions.

The proposed method not only produced accurate fits to experimental data but also visualized predictive uncertainty through confidence bands. This approach addresses key limitations of conventional metrics such as colorimetry, offering a more informative and reproducible assessment of laser treatment outcomes. It also underscores the importance of monitoring both the signal of the contaminant and the underlying surface, as this dual perspective enhances diagnostic sensitivity and minimizes the risk of misinterpretation.

It allows the establishment of meaningful threshold values of laser fluence for experiments under similar conditions, bearing in mind that consecutive cleaning with sufficiently strong laser pulses should, in general, be given priority over the attempt to clean with a single, strong laser pulse.

By embedding uncertainty directly into the model, this framework supports more transparent and risk-aware decision-making in conservation practice. Future developments may include expanding the dataset, incorporating additional surface types, and integrating complementary non-invasive diagnostic techniques to further improve robustness and applicability.

Conclusion and perspective

This thesis aimed to advance the field of fs-laser cleaning for cultural heritage by integrating quantitative spectroscopy, morphological characterization, and predictive modeling to assess femtosecond laser cleaning performance across paper and leather substrates. The main motivation to move forward from the existing practice of laser cleaning with femtosecond UV lasers towards femtosecond NIR lasers is the possibility of avoiding yellowing. At the same time, the lower selectivity of NIR radiation as compared to UV radiation does require a more detailed investigation into the exact nature of the light-matter interaction in this particular wavelength regime.

Chapter 1 introduced the fundamental physical principles of laser–material interactions, diagnostic techniques such as ATR-FTIR spectroscopy, SEM, profilometry, and colorimetry, and statistical methods including ANOVA and Monte Carlo simulations. These sections provided a conceptual foundation for understanding cleaning mechanisms and evaluating outcomes in cultural heritage contexts.

Chapter 2 investigated NIR femtosecond laser cleaning of artificially soiled mock-up paper samples, prepared to replicate historical papers, sized with alum-egg, wheat starch, and corn starch, as well as non-sized controls. Cleaning outcomes were evaluated using ATR-FTIR spectroscopy to quantify underlying substrate signals through exponential saturation modeling and residual contaminant signals through exponential decay modeling. The results demonstrated that effective contaminant removal with minimal substrate alteration could be achieved for non-destructive laser fluences below the threshold of significant multiphoton absorption, although substrate type, contamination characteristics, and burnishing substantially influenced cleaning efficiency. Future work should expand to include additional contaminants, such as metal-based soiling or aged natural varnishes, and assess their interaction with femtosecond laser irradiation to broaden the generalizability of the presented quantitative modeling framework.

Chapter 3 assessed NIR femtosecond laser cleaning of laboratory-prepared mock-up samples replicating historical goat leather, emphasizing the influence of surface porosity and microstructural heterogeneity on cleaning outcomes. Optical microscopy-based porosity quantification revealed heterogeneous distributions of small and large pores, which critically affect laser–material interactions. Grayscale reflectance surface plots further highlighted spatial variability in cleaning responses. Importantly, this chapter demonstrated that for leathers with complex surface morphologies, even widely adopted evaluation methods such as colorimetry may fail to capture true cleaning effectiveness, highlighting the critical role of surface morphology-induced measurement uncertainty. While it has become clear that an unfocussed laser is not able to provide appropriate cleaning conditions to account for surface roughness on the scale of tens to hundres of microns, future research could extend these analyses to three-dimensional porosity mapping and examine the effects of diverse tanning processes on laser–material interaction, while at the same time considering the use of focussed lasers (at the already established fluence levels) adding equally spatially resolved characterization techniques.

Chapter 4 focused on enhancing quantitative cleaning assessments by integrating uncertainty distributions into nonlinear cleaning models. Noise modeling and data perturbation approaches were used to simulate experimental variability, enabling the construction of confidence intervals around cleaning predictions. These results underscore the value of uncertainty quantification in defining safe operational windows and improving the reliability of model-based cleaning protocols. We have shown that, while our model extrapolation depends on the quality (low noise level) of the experiments, it is not sensitive against influence from the sample (such as e.g. the absorption), it remains robust for regardless of the chosen detection system (fitting parameters a and d and even the absorption of the sample). Future studies should incorporate parameter correlation matrices whenever appropriate to more accurately represent interdependencies within model parameters.

While this research has established that the method deserves serious attention as a technologically reliable and commercially attractive alternative to conventional or UV-laser cleaning, its selectivity to specific absorption bands of pollutants and inks or dyes implies that parametric studies to establish the specifics for different material classes are required for a broader distribution of this technique. This is merely a matter of time; the proof of concept has been delivered.

Bibliography

[1] M.D. Shirk and P.A. Molian. A review of ultrashort pulsed laser ablation of materials. *Journal of Laser Applications*, 10:1173, 1997.

[2] W. Kautek and J. Krüger. Ultrashort pulse laser interaction with dielectrics and polymers. *Advances in Polymer Science*, pages 247–290, 2004.

[3] B. Chichkov, C. Momma, S. Nolte, F. von Alvensleben, and A. Tünnermann. Femtosecond, picosecond and nanosecond laser ablation of solids. *Applied Physics A: Materials Science and Processing*, pages 109–115, 1996.

[4] A. Semerok, C. Chale Iard, V. Detalle, J.L. Lacour, P. Mauchien, P. Meynadier, C. Nouvellon, B. Salle, P. Paliano, M. Perdrix, and G. Petite. Experimental investigations of laser ablation efficiency of pure metals with femto, pico and nanosecond pulses. *Applied Surface Science*, 1999.

[5] N. Anderson, L. Wang, and T. Erdogan. Ion-beam-sputtered (ibs) coatings for ultrafast lasers and applications. Semrock White Paper Series.

[6] R. E. Russo, X. Mao, J. J. Gonzalez, and S. S. Mao. Femtosecond laser ablation icp-ms. *Journal of Analytical Atomic Spectrometry*, pages 1072–1075, 2002.

[7] Z. Lin and M. Hong. Femtosecond laser precision engineering: From micron, submicron, to nanoscale. *Ultrafast Science*, 2021:Article ID: 9783514, 2021.

[8] Geoff Shannon and Stephen Hypsh. Femtosecond laser processing of metal and plastics in the medical device industry. *Industrial Laser Solutions*, September 2014. Published online September 12, 2014.

[9] Katherine C. Phillips, Hemi H. Gandhi, Eric Mazur, and S. K. Sundaram. Ultrafast laser processing of materials: A review. *Advances in Optics and Photonics*, 7(4):684–712, 2015.

[10] Q.Z. Zhao and Z. Wang. Manipulation of tribological properties of metals by ultrashort pulsed laser micro-/nanostructuring. In *Advances in Tribology*, pages 247–290. InTech, 2016.

[11] N. Chen, C.-K. Liu, E. M. Brown, and N. Latona. Environment-friendly treatment to reduce photoyellowing and improve uv-blocking of wool. *Polymer Degradation and Stability*, 181:109319, 2020.

[12] D. Rosu, L. Rosu, and C. N. Cascaval. Ir-change and yellowing of polyurethane as a result of uv irradiation. *Polymer Degradation and Stability*, 94(4):591–596, 2009.

[13] B. C. Smith. *Fundamentals of Fourier Transform Infrared Spectroscopy.* CRC Press, 2011.

[14] D. L. Pavia, G. M. Lampman, G. S. Kriz, and J. R. Vyvyan. *Introduction to Spectroscopy.* Cengage Learning, 2014.

[15] S. Wartewig. Basic principles of vibrational spectroscopy. In *IR and Raman Spectroscopy: Fundamental Processing*, chapter 4, pages 27–33. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2003.

[16] D.A. Skoog, F.J. Holler, and S.R. Crouch. *Principles of Instrumental Analysis.* Cengage Learning, Boston, USA, 7 edition, 2016.

[17] P. Larkin. *Infrared and Raman Spectroscopy: Principles and Spectral Interpretation.* Elsevier, Oxford, 2011.

[18] G.-L. Liu and S.G. Kazarian. Recent advances and applications to cultural heritage using atr-ftir spectroscopy and atr-ftir spectroscopic imaging. *Analyst*, 147:1110–1128, 2022.

[19] A. Mohammed and A. Abdullah. Scanning electron microscopy (sem): A review. In *Proceedings of the International Conference on Hydraulics, Pneumatics, Sealing Elements, Tools, Precision Mechanics, Specific Electronic Equipment and Mechatronics*, Romania, 2019.

[20] E. R. Fischer, B. T. Hansen, V. Nair, F. H. Hoyt, and D. W. Dorward. Scanning electron microscopy. *Current Protocols in Microbiology*, Chapter 2:Unit 2B.2, 2012.

[21] M. Schreiner, M. Melcher, and K. Uhlir. Scanning electron microscopy and energy dispersive analysis: Applications in the field of cultural heritage. *Analytical and Bioanalytical Chemistry*, 387(3):737–747, 2007.

[22] A. Kontogeorgopoulou, M. Koui, A. Alexopoulos, M. Spiliopoulou, M. Karoglou, and A. Moropoulou. Scanning microscopy techniques as an assessment tool of materials and interventions for the protection of built cultural heritage. *Scanning*, 2019:1–20, 2019.

[23] H. Lukesova and B. Holst. Identifying plant fibres in cultural heritage with optical and electron microscopy: how to present results and avoid pitfalls. *Heritage Science*, 12:Article number: 12, 2024.

[24] S. Burattini and E. Falcieri. Scanning and transmission electron microscopy in cultural heritage: State of the art. *Microscopie*, 31(1), 2020.

[25] S. Botti, F. Bonfigli, V. Nigro, A. Rufoloni, and A. Vannozzi. Evaluating the conservation state of naturally aged paper with raman and luminescence spectral mapping: Toward a non-destructive diagnostic protocol. *Molecules*, 27(5):1712, 2022.

[26] S. Fu, J. Liu, and H. Yu. A non-contact measuring system for in-situ surface characterization based on laser confocal microscopy. *Sensors*, 18(8):2657, 2018.

[27] T. Buchenau, P. de Groot, and E. Brinksmeier. Comparison of optical and stylus methods for surface texture characterisation in industrial quality assurance of post-processed laser metal additive ti-6al-4v. *Materials*, 16(13):4815, 2023.

[28] D. Ambrosini, D. Paoletti, S. Sfarra, and G. Pasqualone. The potential of optical profilometry in the study of cultural stone weathering. *Journal of Imaging*, 5(6):60, 2019.

[29] I. Lunghi, E. Vannini, A. Dal Fovo, V. Di Sarno, A. Rocco, and R. Fontana. A performance comparison of 3d survey instruments for their application in the cultural heritage field. *Sensors*, 24(13):5283, 2024.

[30] A. Mironova, F. Robache, R. Deltombe, R. Guibert, L. Nys, and M. Bigerelle. Digital cultural heritage preservation in art painting: A surface roughness approach to the brush strokes. *Sensors*, 20(21):6269, 2020.

[31] S. Mazzocato and C. Daffara. A method for spatially registered microprofilometry combining intensity-height datasets. *Sensors*, 23(8):4144, 2023.

[32] B. C. K. Ly, E. Dyer, J. L. Feig, A. L. Chien, and S. Del Bino. Research techniques made simple: Cutaneous colorimetry: A reliable technique for objective skin color measurement. *Journal of Investigative Dermatology*, 140:3–12.e1, 2020.

[33] A. M. Gueli, S. Pasquale, G. Politi, and G. Stella. The role of scale adjustment in color change evaluation. *Instruments*, 3:42, 2019.

[34] C. Cabello Briones, S. Mayorga Pinilla, D. Vázquez Moliní, and A. Álvarez Fernández-Balbuena. Colorimetry to assess the visual impact of dust deposition on mosaics at sheltered archaeological sites. *Heritage Science*, 9:40, 2021.

[35] M. Á. Martínez-Domingo, M. Melgosa, K. Okajima, V. J. Medina, and F. J. Collado-Montero. Spectral image processing for museum lighting using cie led illuminants. *Sensors*, 19:5400, 2019.

[36] M. Melgosa, J. Hernández-Andrés, M. Sánchez-Marañón, J. Cuadros, and Á. Vicente-Retortillo. Some approaches for light and color on the surface of mars. *Applied Sciences*, 14:10812, 2024.

[37] M. A. Barbero-Álvarez, S. Brenner, R. Sablatnig, and J. M. Menéndez. Preserving colour fidelity in photogrammetry—an empirically grounded study and workflow for cultural heritage preservation. *Heritage*, 6:293–312, 2023.

[38] Martin G. Larson. Analysis of variance. *Circulation*, 117:115–121, 2008.

[39] M. Petraretti, A. De Natale, A. Del Mondo, R. Troisi, O. De Castro, N. Mormile, M. Avino, G. Tortino, G. Graziano, and A. Pollio. Deterioration-associated microbiome of a modern photographic artwork: the case of skull and crossbones by robert mapplethorpe. *Heritage Science*, 12, 2024.

[40] J. Qu, M. Sun, F. Wang, K. Liu, and W. Miao. Analysis of microenvironment characteristics and the impact on the preservation of heritage sites: a case study of the jinsha earthen site. *npj Heritage Science*, 13:109, 2025.

[41] C. E. Bottaini, A. Brunetti, I. Montero-Ruiz, A. Valera, A. Candeias, and J. Mirão. Use of monte carlo simulation as a tool for the nondestructive energy dispersive x-ray fluorescence (ed-xrf) spectroscopy analysis of archaeological copper-based artifacts from the chalcolithic site of perdigões, southern portugal. *Applied Spectroscopy*, 72:17–27, 2017.

[42] J. Qu, M. Sun, F. Wang, K. Liu, and W. Miao. Interactions between the ming yansui great wall heritage and geographical environment via monte carlo simulation. *npj Heritage Science*, 13:56, 2025.

[43] V. Morgunov, P. Hybler, and M. Zachar. Geant4 validation for x-ray treatment of wooden cultural heritage artefacts. *Applied Radiation and Isotopes*, 169:109565, 2021.

[44] D. Moro, G. Ulian, and G. Valdrè. Mineral diagnostics: Sem-eds monte carlo strategy for optimised measurements of ultrathin fragments in cultural heritage studies. *ACTA IMEKO*, 10:193, 2021.

[45] A. H. Barbat, F. Y. Moya, and J. A. Canas. Damage scenarios simulation for seismic risk assessment in urban zones. *Earthquake Spectra*, 12:371–394, 1996.

[46] T. Ersoy, T. Tunay, M. Uguryol, G. Mavili, and S. Akturk. Femtosecond laser cleaning of historical paper with sizing. *Journal of Cultural Heritage*, 15:259–265, 2014.

[47] Y.P. Kathpalia. *Conservation and Restoration of Archive Materials*, volume 1. Unesco, Geneva, Switzerland, 1973.

[48] NEDCC. Conservation treatment for works of art and unbound artifacts on paper. Technical report, Northeast Document Conservation Center, 1999.

[49] J. Larson, M. Cooper, and S. Sportun. Developments in the application of laser technology for conservation. *Studies in Conservation*, 45(sup1):107–110, 2000.

[50] M. Bertasa and C. Korenberg. Successes and challenges in laser cleaning metal artefacts: A review. *Journal of Cultural Heritage*, 53:100–117, 2022.

[51] L. Lazzarini, L. Marchesini, and J.F. Asmus. Lasers for the cleaning of statuary: Initial results and potentialities. *Journal of Vacuum Science and Technology*, 10:1039–1043, 1973.

[52] L. Lazzarini and J.F. Asmus. The application of laser radiation to the cleaning of statuary. *Bulletin of the American Institute for Conservation of Historic and Artistic Works*, 13:39–49, 1973.

[53] P. Rudolph, F.J. Ligterink, J.L. Jr. Pedersoli, H. Scholten, D. Schipper, J.B.G.A. Havermans, H.A. Aziz, V. Quillet, M. Kraan, B. van Beek, S. Corr, H.-Y. Hua-Ströfer, J. Stokmans, P. van Dalen, and W. Kautek. Laser-induced alteration of contaminated papers. *Applied Physics A*, 79:941–944, 2004.

[54] J. Krüger, Simone Pentzien, and A. Conradi. Cleaning of artificially soiled paper with 532-nm nanosecond laser radiation. *Applied Physics A*, 92:179–183, 2008.

[55] K. Ochocińska, A. Kamińska, and G. Śliwiński. Experimental investigations of stained paper documents cleaned by the nd:yag laser pulses. *Journal of Cultural Heritage*, 4(Supplement 1):188s–193, 2003.

[56] M. Farsari, P. Pouli, G. Totou, V. Zafiropulos, C. Kalpouzos, and C. Fotakis. The use of high-power lasers in diverse cleaning applications: an overview. In *Proceedings of SPIE*, volume 5777, pages 936–939, 2005. XV International Symposium on Gas Flow, Chemical Lasers, and High-Power Lasers, Prague, Czech Republic.

[57] S. Pentzien, A. Conradi, R. Koter, and J. Krüger. Cleaning of artificially soiled paper using nanosecond, picosecond and femtosecond laser pulses. *Applied Physics A*, 101:441–446, 2010.

[58] J. Kolar, M. Strlič, S. Pentzien, and W. Kautek. Near-uv, visible and ir pulsed laser light interaction with cellulose. *Applied Physics A*, 71:87–90, 2000.

[59] P. Rudolph, F.J. Ligterink, J.L. Pedersoli, M. van Bommel, J. Bos, H.A. Aziz, J.B.G.A. Havermans, H. Scholten, D. Schipper, and W. Kautek. Characterization of laser-treated paper. *Applied Physics A*, 79:181–186, 2004.

[60] T. Smausz, B. Kondász, T. Gera, T. Ajtai, N. Utry, M. Pintér, G. Kiss-Albert, J. Budai, Z. Bozóki, G. Szabó, and B. Hopp. Determination of uv–visible–nir absorption coefficient of graphite bulk using direct and indirect methods. *Applied Physics A*, 123:633, 2017.

[61] J. D. Lindberg and M. S. Smith. Visible and near infrared absorption coefficients of kaolinite and related clays. *American Mineralogist*, 59:274–279, 1974.

[62] D. Thickett and B. Pretzel. Ftir surface analysis for conservation. *Heritage Science*, 8:5, 2020.

[63] G.-L. Liu and S.G. Kazarian. Recent advances and applications to cultural heritage using atr-ftir spectroscopy and atr-ftir spectroscopic imaging. *Analyst*, 147:1777–1797, 2022.

[64] Wayne M. Bundy and Joseph N. Ishley. Kaolin in paper filling and coating. *Applied Clay Science*, 5:397–420, 1991.

[65] Martin A. Hubbe, Joel J. Pawlak, and Alexander A. Koukoulas. Paper's appearance: A review. *BioResources*, 3(2):627–665, 2008.

[66] J. Madejová. Ftir techniques in clay mineral studies. *Vibrational Spectroscopy*, 31(1):1–10, 2003.

[67] Cliff T. Johnston, Jessica Elzea Kogel, David L. Bish, Toshihiro Kogure, and Haydn H. Murray. Low-temperature ftir study of kaolin-group minerals. *Clays and Clay Minerals*, 56(4):470–485, 2008.

[68] C.Y. Boynukara, M. Uguryol, G. Mavili, and S. Akturk. An investigation of the cleaning performances of femtosecond and nanosecond laser pulses for artificially soiled papers with sizing. *Applied Physics A*, 127:275, 2021.

[69] S. Sportuna, M. Cooper, A. Stewart, M. Vest, R. Larsen, and D.V. Poulsen. An investigation into the effect of wavelength in the laser cleaning of parchment. *Journal of Cultural Heritage*, 1:S225–S232, 2000.

[70] W. Kautek, S. Pentzien, A. Conradi, D. Leichtfried, and L. Puchinger. Diagnostics of parchment laser cleaning in the near-ultraviolet and near-infrared wavelength range: a systematic scanning electron microscopy study. *Journal of Cultural Heritage*, 4(Supplement 1):179s–184s, 2003.

[71] W. Kautek, S. Pentzien, M. Röllig, P. Rudolph, J. Krüger, C. Maywald-Pitellos, H. Bansa, H. Großwang, and E. König. Near-uv laser interaction with contaminants and pigments on parchment: laser cleaning diagnostics by se-microscopy, vis-, and ir-spectroscopy. *Journal of Cultural Heritage*, 1(Supplement 1):S233–S240, 2000.

[72] A. W. Landmann. Lubricants. In C. Calnan and B. Haines, editors, *Leather, its composition and changes with time*, page 30. The Leather Conservation Centre, Northampton, 1991.

[73] J. W. Waterer. *Leather.* Clarendon Press, Oxford, 1962. Reprinted from Singer and Holmyard: *History of Technology*, Volume II, for the Museum of Leathercrafts.

[74] W. Kautek, S. Pentzien, J. Krüger, and E. König. Laser cleaning of ancient parchments. In *Laser in the Conservation of Artworks, LACONA I*, pages 69–78, Heraklion, Crete, 1995.

[75] H. A. B. van Soest, T. Stambolov, and P. B.. Hallebeek. Conservation of leather. *Studies in Conservation*, 29:21–31, 1984.

[76] A. Lo Monaco, M. Marabelli, C. Pelosi, and R. Picchio. Colour measurements of surfaces to evaluate the restoration materials. In *Proc. of SPIE*, volume 8084, pages 80840P–14, 2011.

[77] DIN 6174:1979–01, Colorimetric Evaluation of Colour Differences of Surface Colours According to the CIELAB Formula. Deutsches Institut für Normung (DIN), 1979. Standard.

[78] M. Strlič, J. Kolar, V. S. Selih, and M. Marincek. Surface modification during nd:yag (1064 nm) pulsed laser cleaning of organic fibrous materials. *Applied Surface Science*, 207:236–245, 2003.

[79] W. S. Mokrzycki and M. Tatol. Colour difference delta e – a survey. *Machine Graphics and Vision*, 20:383–411, 2011.

[80] A. Elnaggar, P. Fitzsimons, A. Lama, Y. Fletcher, P. Antunes, and K. G. Watkins. Feasibility of ultrafast picosecond laser cleaning of soiling on historical leather buckles. *Heritage Science*, 4(1):30, 2016.

[81] P. Moretti and et al. Laser cleaning of paintings: in situ optimization of operative parameters through non-invasive assessment by oct, ft-ir and lif. *Heritage Science*, 7:44, 2019.

[82] M. Walczak and et al. Evaluation of femtosecond laser pulse irradiation of ancient parchment. *Applied Surface Science*, 255:3179–3183, 2008.

[83] S. Cagno and et al. Comparison of four mobile, non-invasive diagnostic techniques for differentiating glass types in historical leaded windows: Ma-xrf, uv–vis–nir, raman spectroscopy and irt. *X-Ray Spectrometry*, 50:293–309, 2021.

[84] T. Dey and D. Naughton. Cheap non-toxic non-corrosive method of glass cleaning evaluated by contact angle, afm, and sem-edx measurement. *Environmental Science and Pollution Research*, 24:13373–13383, 2017.

[85] K. Sanderson. A comparative study of handheld reflectance spectrophotometers. *PMG Annual Meeting Postprints*, 19:100–110, 2015.

[86] S. Lorusso, A. Natali, and C Matteucci. Colorimetry applied to cultural heritage: examples of case studies. *Conservation Science in Cultural Heritage*, 6:1–18, 2006.

[87] F. L. M. Milotta, G. Furnari, C. Quattrocchi, S. Pasquale, D. Allegra, A. M. Gueli, F. Stanco, and D. Tanasi. Challenges in automatic munsell color profiling for cultural heritage. *Pattern Recognition Letters*, 125:574–580, 2019.

[88] B. Ramírez Barat, E. Cano, M. T. Molina, J. Santana, B. Díaz, and J. M. Bastidas. Design and validation of tailored colour reference charts for monitoring cultural heritage degradation. *Heritage Science*, 9:41, 2021.

Acknowledgments

My advisors, Patrizio and Mauro — This challenging journey would never have begun without you. Thank you for believing in me. Patrizio, I only wish we had had the chance to do more experiments together in the cold Canadian winter. I'll never forget how much fun they were.

Elias — First and foremost, thank you for tirelessly translating French for me with endless patience. Without you, every beam time would have turned into a nightmare. Thank you for always standing by me when I needed help. Your support has made all the difference.

Agnese — I am so happy our paths crossed during this journey. Knowing there's a friend who will greet me with a smile every time I visit Italy is a beautiful feeling. And never forget there will always be a place for you in our home, too.

Nils — Although we met later in this journey, I am truly glad to have gotten to know you. A long and fulfilling PhD path surely awaits you, and whenever you need anything, you know how to reach me.

Katharina, Amira, and Lilian — Three brilliant women. Even if our time together was short, I am genuinely grateful for your support throughout this journey. Without your scientific insight, my work would have been incomplete.

Diana Cagnina and Enrico Papiri — Thank you for all your support. Traveling to Rome has always been precious and joyful because of the time spent with you. Thank you for sharing the stress of this PhD with me and supporting me through this journey. I am so grateful our paths crossed. Here's to many more moments of happiness to celebrate together!

Mehmet — I am deeply grateful for your unwavering support since my master's studies. You were always there when I was struggling, patiently answering my questions. Thanks to you, I never gave up. It has been an incredible privilege to know you and work with you.

My dear friends, Ceyda and Midya — I was exhausted during this journey, but you never stopped supporting me. No matter where I was in the world, you were always by my side. Your presence has never been missing from my life; please never let it be. Let's keep meeting in different cities around the world. My holidays in Istanbul

are beautiful now, just because you're there. I love you both so much.

Müjde and Çelen — You are the most precious gifts this PhD journey has brought into my life. I truly appreciate how you've carried every challenge with me, throughout this journey and beyond. Chasing adventures together across cities worldwide has been unforgettable. We are still at the beginning, the road is long, and we are young.

My co-supervisor Andreas — I'm incredibly happy to have met you. You have influenced not only my PhD journey but also how I look at life and many things. Your discipline, boundless energy, and unwavering support for your students in all circumstances are extraordinary. What a blessing for any student! Every time my research progressed, I often found myself thinking, "If only we had had more time." Thanks to you, I now drink my coffee without milk, and here's to many more cups we'll share.

Fulya and Miktat Boynukara — To my beloved family, nothing I have achieved would mean anything without you. I'm forever grateful for your unconditional support. I'm so lucky to have you. Here's to many more shared successes!

Birtanem, Roberto — It's a miracle that this journey I began on my own brought me to you. Thank you for holding my hand every time I fell and never letting go. This journey wouldn't have ended so beautifully without you. There is no doubt that one of these diplomas belongs to you.