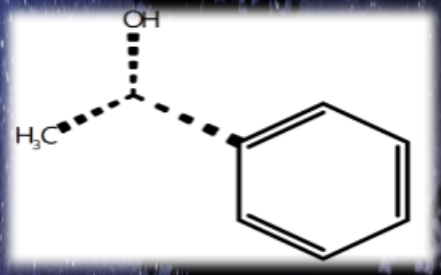
Science Ascend

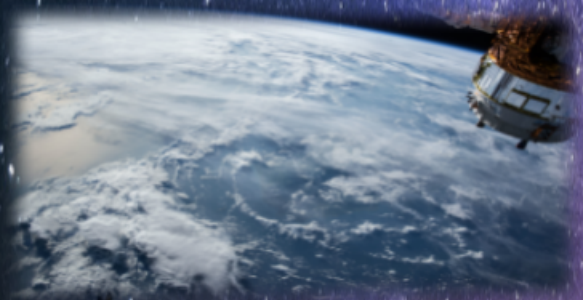
From September 16-23, 2024!

Rising to new heights of discovery with Science!
Every week!

Better theoretical understanding, TESS, HST, XMM, VLT, JWST-MIRI and more!

More robust MALDI-MSI, forensic with fluorescence, robust biosensors, ML in chemistry and more!





Manifolds for hyperspectral image sharpening, Sentinel 1 and 2 fusions, much better segmentation algorithms, and better UAV monitoring!

Aqueous ozone oxidation of thiosulfate, reaction kinetics and pathway determination, and sub 10 nm size particle mass measurement by MS!

cout<<solutions;

Decomposing model and data uncertainty, patch PCA, much better subset selection for algorithms and more!

ISSN: 3062-0090
FIRE Araştırma Eğitim Ltd. Şti., Vol:1, Issue:3

Science Ascend

Rising to New Heights of Discovery!

Science Ascend teleports you to the frontiers of science. It compiles and discuss the scientific research preprints from arXiv, bioRxiv, chemRxiv just from the previous week to be cognizant of the *state-of-the-art* of knowledge in astrophysics, chemistry, environmental chemistry, remote sensing, and applied statistics/data science. Light from the *Science Ascend* will keep brightening the dark horizon beyond the limits of our comprehension. FIRE Araştırma Eğitim Ltd. Şti. guarantees the weekly publication and dissemination of this journal, and make it available for everyone at most fifteen days after its publication freely.

Publisher: FIRE Araştırma Eğitim Ltd. Şti.
Media: Online Journal
Responsible person: Yasin Güray Hatipoğlu
Editor-in-chief: Yasin Güray Hatipoğlu
Editor: Yasin Güray Hatipoğlu
Frequency: Once a week
Address: Yıldızevler Mah. Kişinev Cad. No:10 Çankaya/Ankara/Türkiye
Website: https://fire-ae.github.io

This issue: September 23, 2024

Volume: 1

Issue Number: 3

Bilim Yükselişi

Keşfin Yeni Yükseklerine Ulaşmak!

Science Ascend sizi bilimin sınırlarına ışınlar. Astrofizik, kimya, çevre kimyası, uzaktan algılama ve uygulamalı istatistik/veri bilimi alanlarındaki bilgi birikiminin *en son durumu* hakkında bilgi sahibi olmak için arXiv, bioRxiv, chemRxiv'den sadece bir önceki haftaya ait bilimsel araştırma ön baskılarını derler ve tartışır. *Bilim Yükselişi*'nden gelen ışık, kavrayışımızın sınırlarının ötesindeki karanlık ufku aydınlatmaya devam edecektir. FIRE Araştırma Eğitim Ltd. Şti. bu derginin haftalık olarak yayınlanmasını, dağıtılmasını ve yayınlandıktan en geç on beş gün sonra ücretsiz olarak herkesin erişimine açılmasını garanti eder.

Yayıncı: FIRE Araştırma Eğitim Ltd. Şti.
Ortam: Online Journal
Sorumlu Kişi: Yasin Güray Hatipoğlu
Yazı İşleri Müdürü: Yasin Güray Hatipoğlu
Editör: Yasin Güray Hatipoğlu
Yayınlanma Sıklığı: Haftada bir kez
Adres: Yıldızevler Mah. Kişinev Cad. No:10 Çankaya/Ankara/Türkiye
Website: https://fire-ae.github.io

Bu sayı: September 23, 2024

Cilt: 1

Sayı Numarası: 3

Last week in Astrophysics

Author: Yasin Güray Hatipoğlu

The preprints summarized here were published between September 16 - September 23, 2024. These are from arXiv's astro.EP cross-fields without high-energy main cross-list papers.

Stellar Systems - Populations - Clusters

Cooper et al.[1] acquired low and middle resolution (R 300 and R 2500)[1] optical spectroscopy of 53 late M and L ultracool dwarfs in the Solar neighborhood with the Optical System for Imaging and low-intermediate Resolution Integrated Spectroscopy - OSIRIS of the Gran Telescopio Canarias at El Roque de los Muchachos, La Palma, Spain. They stated that this study is the fourth of the Gaia UltraCool Dwarf (UCD) sample series, which aims to characterize all Gaia visible UCDs. They determined their samples' spectral types, spectral indices, and radial velocities from OSIRIS spectra, and effective temperature, surface gravity, and metallicity [Fe/H][2] astrophysical parameters were inferred by comparing with atmospheric models. They reduced the data with the PypeIt[3] reduction procedure and compared it with the standard IRAF tasks on the J1745-1640 object, and found agreement in the part of the spectrum with longer wavelengths than 9800 angström. Analysis was done with the rvfitter package. Throughout the analysis procedure, specutils, kastredux packages were also utilized. They determined the radial velocity values from the middle resolution (R 2500) spectra. With this study, 29 of the 46 objects gained a radial velocity estimation for the first time. Two objects were outside of the thin disc[4] and four in the young stellar kinematics group.

[1]This R is the resolving power and is calculated as R = $\lambda/\delta\lambda$, where λ is the wavelength and the denominator is the minimum wavelength difference that can be resolved with the resolving power R. A higher R-value indicates a finer detail in terms of the wavelength in the spectrum.

[2]This iron/hydrogen abundance is in dex - decimal exponent, e.g., Fe/H 1 dex means 10^{+1} times more metallicity than Sun.

[3]PypeIt can be reached here.

[4]Thin disk is a part of galactic components. From inside out,the Galactic center is the center, the bulge is around this center, the thin disc is a mostly planar disk around this bulge, the thick disc is farther away over this thin disc and slightly more rarefied. Finally, the Galactic halo is the most rarefied component of a galaxy, a halo encapsulating all other components. Objects are placed in these groups here according to the total space velocities.

Simulation-based studies

Gautam, Farias, and Tan[2] made N-body simulations to measure the impact of the surrounding interstellar environment within star clusters on protoplanetary disk (PPD) formation. Their method was a gradual semi-analytical model of PPD formation in N-body way together. Their model basis is Turbulent Clump Model[5]. PPD was assumed to be 10% of its star mass and modelled for every star. Additionally, the perturbing star also impacts the disk radius and its truncation. Finally, external photoevaporative irradiation in the interstellar medium also results in mass loss at the outskirts of these disks. Nbody6++ was used for the gradual star cluster formation outlined in this study. They reported many insights regarding the impact of initial conditions in the stellar cluster on disk truncation, and in the comparison of their results with the discs in the ONC and G286 star-forming region, they could match some of their models with these discs. They also stated that the inner core of star clusters has more truncated discs than the outer ones with more active disc depletion mechanisms.

Single Star System (Star, Exoplanet)

Barnes et al.[3] stated that radial velocity (RV)[6] method of exoplanet detection requires the understanding and removal of stellar activity from the estimated exoplanet-originated signal. Their approach is analytically modeling cool and facular spots and calculating their impact on central line moments[7] (CLM), which is used in RV calculation. They also applied this method to the AU Mic[8] M dwarf star. The common methods mostly include cross-correlation function[9] to find the spectral shift magnitude, and bisector inverse span, to distinguish stellar noise-related spectrum

[5]The clump is a polytropic sphere of gas, polytropy comes from pressure-density relation in this context. Details can be found here.

[6]The second highest number of exoplanet-detecting method. Basically, it measures stellar movements indirectly from the blue-red shifting of its light owing to the system's center-of-mass distance from the stellar core because of exoplanets.

[7]It is a statistical way to describe the sample distribution. These are the mean, standard deviation, kurtosis/skewness and similar terms regarding the distribution.

[8]AU comes from the variable-star designation according to the detection order of the star in its respective constellation, Mic is from the Microscopii constellation.

[9]Basically, it is used to estimate a time-series with another time-series in different lags. One may correlate eating *Salmonella spp.* containing food and getting diarrhea in let's say, 16 hours of time lag, or generation of sunlight with reaching it to our Earth by 8 minutes, etc. In radial velocity calculation, the index of the data is wavelength instead of time, and the aim is matching the spectral line of a reference or synthetic spectrum with that of a real observation.

line distortion from the *Keplerian* signal. The details of their methodology are as follows: They used the Doppler imaging code DoTS to simulate line profiles for selected rotation angles with a specific signal-to-noise ratio then directly calculates CLM for each of them. They generated absorption lines from three different temperature models and specific spot-filling factors. Line profiles were calculated by considering cool spots, M dwarf and G dwarf facular contrasts, and convective blueshift[10]. They constructed the stellar surface pixel-by-pixel with different intensities of cool spots, facular regions, or the photosphere. Then, they experimented and reported central line moments from the single spot models and scaled multi-spot models. They reported that methods such as the generalised Lomg-Scargle periodogram interpretation mostly result in the recovery of a harmonic of the true stellar rotation period, as they assume sinusoidal variability. They instead recommend *string length minimisation*[11] for this.

Maggio et al.[4] concentrated on the youngest known star HIP[12] 67522 (around 1.2 solar mass) with a transiting multi-planetary system. They wanted to have extreme UV (EUV), but it was unavailable owing to the strong interstellar absorption, so they estimated EUV from the Far UV (FUV) from the Hubble Space Telescope (HST) and X-ray from the X-ray Multi-Mirror Misson (XMM) Newton. Time-resolved spectra acquired signals from the star's quiescent and flare periods separately, as well as corona and chromospheric temperature using Reflection Grating Spectrograph (RGS) for XMM-Newton and Cosmic Origin Spectrograph for HST. They reduced XMM-RGS data with XMM-SAS version 20.0.0 whereas for HST, analysis-ready data was downloaded from the HST archive. Emission Measure Distribution was obtained with PoA Software Monte Carlo Markov Chain procedure with the atomic line emissivity database of Chianti v. 7.13. They used ATES Hydrodynamic code[13] to model atmospheric mass loss rate which results from the stellar irradiation. As expected from a young star, even its quiescent period is very active with x-ray luminosity around $3x10^{30} ergss^{-1}$. For the range of core and atmosphere fraction, the planet may lose its atmosphere in several million years versus it might take more than 10 billion years.

Biagini et al.[5] made a photometric analysis in multiple wavebands for V1298 Tau to model spots' contribution to the stellar photometry so that they can remove this factor and study the young active planetary systems later. They chose to approach this problem with multiple waveband spectra since they expected that at longer wavelengths stellar and spot fluxes differ from each other less. Their telescope was Ritchey-Chreiten from the Osservatorio Polifunzionale del Chianti (OPC), an 80 cm diameter main mirror, 20 x 20' fielf of view, and Johnson filters in BVRI (blue, green (v is similar to vert, meaning green in Roman languages), red, and infrared). The stellar rotation period was known to be approximately 2.8 days and their observation campaigns took place in 1 and 2 days generally. They estimated the spots with the PyMultiNest v2.12 Python package which used a fitted sinusiodial with an offset to the lightcurve. ExoTETHyS package was used for limb-darkening effect[14]. They found their approach beneficial and also validated it on solar data.

Thanathibode et al.[6] constructed a new model for accretion shock-originated[15] carbon IV 1548-1550 Angström doublet line of T Tauri Stars[16]. This carbon emission was generally found related with the accretion of the inner protoplanetary disk materials. They consider both the geometry of this accretion via star and disk geometry, as well as 1D pre- and post-shock calculations. Cloudy[7] emissivities were used for post-shock density and temperature calculations. The temperature and flow velocity of the materials were decisive in the broadening of the line peaks and which line was to be the dominant emission line. After these, they used Hubble Ultraviolet Legacy Library of Young Stars as Essential Standards Director's Discretionary Program (ULLYSES)'s data of con-temporaneous UV-Infrared spectroscopy and ODYSSEUS, and PENELLOPE programs, and also Very Large Telescope's high/moderate resolution spectra. Their chosen sample set was from Orion OB1 and σ Ori regions. They obtained all required parameters for their modelling except truncation radii from the literature fo such

[10] Blueshift of the spectrum originated from convection in the stellar atmospheres.

[11] Repeated phase-folding of the data and iteratively finding the best match for estimating the sinusoidal pattern. In the perfect case of folding a single sinusoidal wave according to its period, there is 0 string length between the data points in consecutive phases. The real data has more than a perfect single sinusoidal wave, but getting this distance between the consecutive phases in the same location in the x-axis shorter and shorter, the period for phase-folding gets near to the stellar rotation period.

[12] From the European Space Agency's *Hipparcos* catalog.

[13] ATES GitHub repository

[14] The limb is darker than the center/core of the star. Basically, the light from the core directly comes towards us, while the ones from the limbs pass through a more stellar/solar atmosphere and can reduce, in the Sun's case, by as much as 20 %. More information can be found here.

[15] Matter falls on the star and this generates a shock.

[16] The name originates from the ternary-triple variable star T Tauri, which means from *Taurus* constellation.

studies, calculated truncation radii[17] from the PENELLOPE program's spectra with the magnetospheric flow model from Muzerolle et al.[8]. In the end, they separated C IV emission lines into broad (preshock) and narrow (postshock) components and they found differing origins and factors behind these narrow and broad components. Four out of seven of their samples from the ULYSESS Orion OB1 were fitted well, and they had subsolar carbon abundances in their inner disks. Their approach was beneficial and had room for development considering different geometries.

Malla et al.[9] reported on the inconclusive status of low-luminosity (retired A-stars) mass offset in measurements with the spectroscopic, interferometric, and asteroseismologic methods. They used the data from the Transiting Exoplanets Survey Satellite (TESS) with zero pixel quality flags and PDCMAP[18] (Pre-search Data Conditioning Maximum A Posteriory), high-pass filtered (3 microHertz, around 4 days) the data to remove slow trends (and for 24 Sex and HD 185351 50 microHertz). The outliers were assumed to be higher than four times the standard deviation in the sample, hence they applied $4-\sigma$ clipping. After also several manual removal of the sudden changes in time-series above 4-σ, they normalized the lightcurves from every sector by sector-median division and combined them. They applied discrete Fourier Transform to calculate power density spectra, and then, for the maximum acoustic power and large frequency of separation, SYD pipeline[19] was applied. After these, 92 low-luminosity and 157 high-luminosity targets were selected to further work out of all 451 star samples. The stellar mass determination step had two different ways for asteroseismic-based ones, one from individual estimations using the maximum acoustic frequency or large frequency separation, and one combining both. These are empirical equations with ratios of maximum acoustic frequencies, luminescence, and temperature, and combining two of them also discards the luminosity end reduces the temperature measurement impact on the mass determination. They also tried to fit for the mass-dependent mass offset in comparing spectroscopic and seismic mass estimations with polynomials and benefitting from the Bayesian Information Criterion. In the end, they were able to expand the mass offset with 249 new samples.

[17] Truncation radius is related to the protoplanetary disk truncation, and was generally assumed to be 5 times the stellar radius.

[18] Coming from co-trending basis vectors, as explained in previous issues, it measures cotrend in irrelevant locations in the images and assumes that this comes from non-celestial redundant instrumental error nad similar origins.

[19] A Python adaptation can be found here, SYD comes from Sydney, Australia.

Simulation-based studies

Li[10] created the Python package Toy Coronagraph[20] to estimate, understand, and develop ways to overcome exozodiacal dust[21] problem in coronagraphic image acquisition studies better and faster. The author emphasizes the light-scattering impact of such dust and its detrimental effect on the process of direct imaging of exoplanets. Such light scattering would make the entire background brighter and requires even higher contrasting power to detect exoplanet-specific signals. Toy Coronagraph package's contribution allows more details of this light-scattering impact to be modeled and considered compared to the available simulators. Both the input dust disk image and coronagraph point spread function (PSF) are circularly symmetric for better computational efficiency. The package multiplies the dust disk and PSF values for each radius with the $2\pi r$ factor, and this is being done for the entire circle, and finally, the sum result was divided by the number of rotations it took to calculate this. Furthermore, the author defined a "Core Throughput Region" where the '*core* from a point-source like a star starts from the brightest pixels and expands until coming to a neighboring pixel where it is less than half of its brightest neighboring pixel. This approach permits the calculation of both the effectiveness of stellar light suppression by the coronagraph and the possibility of detecting the exoplanet in the middle of dust light-scattering. Many parameters can be tuned. The package comes with the MCFOST[11]-generated Epsilon Eridani dust disk, allows different PSF generations, with a user-friendly interface.

Marigo et al.[12] extended the AESOPUS Rosseland mean gas opacities[22] to high-pressure environments and broader density and temperature conditions to also accommodate sub-stellar objects and very low mass stars. The Rosseland mean opacity is calculated by: $R = \rho T_6^{-3}$ where ρ is density and T_6 is million Kelvins. They also explained ionization potential depression (IPD) impacts at higher densities, where this becomes relevant to calculate in their extended higher-density regions. An increased number density of electrons and ions reduces the generalized Debye length[23], which increases the IPD.

[20] Toy Coronagraph can be retrieved from here.

[21] This is related to the dust in other stellar systems than Sun, zodiacal is related to the solar system objects' appearance in the sky, and in this context zodiacal dust would mean solar system dust.

[22] Inversely related to the transmissivity. It is a measure of the extent of how much light is absorbed or scattered while passing through a specific medium.

[23] Measures a single charge's electrostatic effect in a so-

AESOPUS also considers this IPD. They include thermal motion-induced Doppler broadening and pressure-induced pressure (Lorentzian) broadening, as well. Additional broadening parameters were retrieved from the ExoMol database when available, and a total of 65 astrophysics-relevant molecular species were included. Later, they reported Rosseland mean opacities with solid grains and also considered the novel approach's impact on stellar models of especially low mass ones, between 0.1-0.85 solar masses. They provided the results and the new AESOPUS web interface here.

Sajadian and Afshordi[13] simulated the TESS light curves of self-lensing and eclipsing cases by compact companions (white dwarf (WD), neutron star (NS), or stellar-mass black hole (SBH)). They considered edge-on cases, hence transit/eclipse events were included in the simulations. Elliptical orbit by Kepler's equation was solved numerically, the Einstein radius regarding the self-lensing events was explained and relevant constants and equations were given. While lensing takes place, occultation (the compact object masks a part of the source star directly) and eclipse (the source star is in front of the compact object) occur, too, and they should be considered. They proceeded to make different light curve simulations and the peak shapes and relative intensity changes in the normalized flux with changing compact object mass, source star radius, orbital eccentricity, and limb-darkening coefficient. They chose events shorter than 365 days, used 2-min cadence TESS data, and in general, observation time spans changed between 27.4 days to 356 days according to the chosen candidate target lists for this study. Especially WD and NS binaries had detectable impacts in both high-confidence (higher signal-to-noise ratio) and low-confidence cases.

Sun

Liu et al.[14] reported the last two decades largest magnitude geomagnetic storm from May 2024, using GOES EXIS (while trying to associate the coronal mass ejections with the flares, SOHO, WIND, and STEREO-A. They focused halo eruptions[24] as candidates most effective on the Earth, further examined the *geo-effectiveness*, and also checked solar wind at a distance of 1 astronomical unit (AU). They compared two separate complex ejecta from Sun and in halo eruptions also plotted GOES on top of SOHO and STEREO A data acquired between May 8-14, 2024.

Penza et al.[15] estimated and reconstructed total solar irradiance (TSI) starting from the open solar magnetic flux dataset between 971 common era (CE) and 2020 CE. The idea is to use surrogate variables to estimate TSI as its direct measurements began in the 1970s. Very roughly, they estimate the historical solar magnetic field from the cosmic ray flux to the ^{14}C measurements in tree rings. They estimated spots and faculae as well. Empirical Mode Decomposition[25] was utilized to decompose data. Their results were generally in line with the literature.

Instrumentation - Software

Gordon et al.[16] reported James Webb Space Telescope (JWST) Mid-Infrared Instrument (MIRI) Imaging and Coronography absolute flux calibration. The data they used were from the same data of Cycle 1 and Cycle 2[26] flux calibrations, and commissioning period. The reduction pipeline was jwst version 1.51.1[27]. They chose to use the MIRI imaging reduction pipeline for the coronagraphy as well, and they included different subarray[28] formations. They utilized band-specific point spread functions (PSF) for the aperture and background annulus and with the photutils Python package. The temporal change in the instrument's flux measurement was estimated with a known very low intrinsically variable BD+60 1753 dwarf star. They reported an approximately 22 % degradation for the filter F2550W[29], and the degradation got smaller with the smaller wavelength. They also reported subarray, source, and detector dependencies, and no considerable dependency was found for source and detectors after the subarray correction.

Blunt et al.[17] reported the version 3 of *orbitize!*[30].

lution. The higher it is, the more isolated/screened the ions are and the lower influence.

[24] They appear as halo/ring-like structures on white-light coronagraph that monitors the Sun.

[25] Cubic spline fitting of data with several criteria to separate it into intrinsic mode functions which capture oscillations, and the residual contains trend.

[26] Cycles are yearly observation programs that start on July 1st of the year.

[27] jwst pipeline is in three different steps, detector1, image2, and image3, in which the first one aims to correct instrumental errors, the second one subtracts background and generates rectified 2D products after the flux calibration, and the third one register it in world coordinate system (WCS) better, match the background, segment the image, and create a source catalog.

[28] A part of the imaging array, which can be utilized for other purposes.

[29] F means filter, there can also be grating. 2550 is 25.50 micrometers wavelength, and W is wide, basically, a wide filtering window instead of a very narrow peak.

[30] The GitHub repository for *orbitize!* is here. Bayesian modeller software in the Journal of Open Source Software. This software aims to aid in retrieving the orbital parameters of binaries from time-series data, especially for high-contrast imaging. The paper in this case informs the readers of new features since its version 1, such as joint-fitting of radial velocities of both the primary star and its companion, fit absolute astrometry including Hipparcos-Gaia ac-

Cox et al.[18] introduced the optical-photothermal infrared (O-PTIR) method to the planetary science community. They stated that this technique is better in performance considering the spatial and spectral resolution **together**. Most peaks in this and FTIR also coincided with each other. Moreover, this method may still work without a sample preparation step.

celerations and intermediate astrometric data, nested sampling backend via the dynesty[31] Python package, 2 ways to handle multiplanet effects, and the ability to fit in different orbital bases and using priors from observations. The static nested sampling improves error in evidence, but for better posterior estimation efficiency, a dynamical way to trade-off evidence accuracy in favor of posterior estimation is choosing live points in variable numbers through the runs. Shortly, this can be seen as an alternative to the Monte Carlo Markov Chain methods in handling the data and estimating the posterior.

Last week in Chemistry

Author: Yasin Güray Hatipoğlu

The preprints summarized here were published between September 16 - September 23, 2024. They are more in nature of spectroscopy alone, and hence several studies regarding biochemistry, chromatography, and several other disciplines might be missed here.

Mass Spectroscopy

Zemaitis et al.[19] worked with fresh frozen pancreas tissues and tried to disentangle matrix impact on the matrix-assisted laser desorption/ionization (MALDI)-mass spectrometry imaging (MSI) analysis of these living tissue samples. Firstly, they report that washing the sample with ammonium formate reduced the polymer contamination of the spectrum, especially in negative ion mode. Positive ion mode had higher persistent polyethylene glycol polymer contamination. They reported the data in the following link.

Rosales et al.[20] worked on better detections of opioids, amphetamines, and psychedelics with trimethylation enhancement using diazomethane (TrEnDi) to increase MS sensitivity on 13 different drugs. Aprotic solvent system for electrospray ionization made it even much more sensitive.

Fluorescence Spectroscopy

Elliott et al.[21] used fluorescence spectroscopy with parallel factor analysis (PARAFAC) to determine the time-since-deposition of bloodstains, an important metric in forensic investigations. They tested this in bovine blood collected with and without anticoagulants. They conducted excitation-emission matrix (EEM) measurements with 250-500 nm excitation in 5 nm steps and measuring every nm between 280-750 nm, and in the second EEM, excitation was done between 350-600 in 5 nm steps and 400-700 nm were measured. After several post-processing steps, they also removed Rayleigh and Raman scattering peaks from the spectra. They analyzed the data with staRdom's library in R Statistical Software. They fused all of the EEM data and constructed four different regression models to predict the time-since-deposition (TSD) values: 1) Boruta feature selection and random forest regression, 2) Random Forest regression with no feature selection, 3) partial least-square regression (orthogonal scores algorithm, pls library), and 4) Elastic Net regression (glmnet library). Hyperparameter tuning and model performance evaluation were done with 5-fold cross-validation. A total of 42 eluates were prepared for mass spectrometry analysis with a High-Performance Liquid Chromatography-Mass Spectrometry/Mass Spectrometry (HPLC-MS/MS). Even though the results were generally above 0.95 adjusted R^2 values, among them Boruta feature selection followed by random forest regression was the most accurate one.

Microscopy

Mukherjee et al.[22] applied data analysis techniques to disentangle heterogeneity on materials via the k-means clustering technique. They worked on caesium lead bromide nanocrystals, and data analysis included estimating the optimum number of clusters, categorizing the ata, and cluster-wise power spectral density analysis. silhouette score, Calinski-Harabasz index, and the elbow method were utilized to find the optimum number of clusters. They ran the code in epochs of 100 to reduce the discrepancy potentially originating from the randomly chosen initial centroids in k-means clustering and took the mean of indices mentioned above. Their initial clustering results were unsatisfactory, and to mitigate this, they looked for discrete wavelet transformation assisted noise-flattening and data binning techniques to reduce the dimension of the data. While the wavelet analysis did help, subsequent data binning resulted in even better misclassification scores with values reaching 0.044 out of 100 iterations. The optimum cluster number was found 3 in all approaches and power spectral density profiles were different for each classes. They also applied this method to real time spectroscopy data

Biosensors

Zhang et al.[23] worked on a universal oxygen scavenger for extending oxidase-based biosensor lifetime and facilitating their storage, as well as reducing the oxygen interference. They reported that a combination of alcohol oxidase, paraformaldehyde[32], and catalase enzyme scavenges oxygen and prevents its interference. They specifically focused on glucose oxidase, commonly used for blood sugar sensing.

Infrared Spectroscopy

Olivares et al.[24] measured and corrected the impact of temperature fluctuations on near-infrared spectra of raw milk samples. They worked with

[32]Formaldehyde polymer

270 samples and took the spectra at 20, 25, 30, 25, and 40 degrees of celsius. The methods to correct these variations in the study were piecewise direct standardization (PDS), continuous PDS, external parameter orthogonalization (EPO), and dynamic orthogonal projection (DOP). The actual method to monitor the milk composition is partial least square regression (PLSR), and the preprocessing methods above were applied before that to see if they better the results. These preprocessing methods significantly improved the results, and the EPO and DOP were superior to the other cases and the case where there was no preprocessing method.

Nuclear Magnetic Resonance

Silva[25] utilized machine learning classification methods for molecular structure estimation via ^1H NMR spectra interpretation. The suite of methods included k-nearest neighbors, decision tree, support vector classifier, extra trees, random forerst, and gradient boosting in sklearn package in Python environment. The ^1H NMR spectra database used in the study was from NMR Challenge[33]. The class imbalances in the functional groups were mitigated with oversampling via Synthetic Minority Oversampling Technique (SMOTE) variations, Borderline SMOTe and SVMSMOTE, and adaptive synthetic sampling (the reason for choosing them was practical, as they were easy-to-use from imblearn Python package). The best result was the application of SMOTE, then, random forest with an 0.88 accuracy score.

Electrochemistry

Seymour et al.[26] devised a method and instrument that can detect dissolved oxygen down to 0.36 pm detection limit even in the presence of chlorine and iron by converting it to hydrogen peroxide and measuring that species. Another superior characteristic of this device is that it is membrane free.

[33]The challenge website can be found here.

Last week in Remote Sensing

Author: Yasin Güray Hatipoğlu

The preprints summarized here were published between September 16 - September 23, 2024. These are generally based on the preprints retrieved when "remote sensing" words are given between quotation marks within arXiv's cs.CV and similar cross-fields.

Sharpening-like Methods

Wang et al.[27] worked on fusing hyperspectral images with multispectral ones having higher spatial and temporal resolution. They reported a blind fusion method based on deep Tucker decomposition[34] and spatial-spectral manifold learning[35] to overcome previous algorithms' shortcomings on data representation, compression, and feature extraction. They reported that their novel approach performed better than available cases, and their code can be found here.

Diakogiannis et al.[29] worked with Sentinel-2 Multispectral Imager and Sentinel-1 (radar) time series data together to eliminate cloud cover-related problems in field delineation with satellite remote sensing data[36]. They worked with Australia data for the training and inference area. Their base was on Patch Tomato Attention Vision Transformer 3D (PTAViT3D). They estimated the segmentation efficiencies and Sentinel-2 Sentinel-1 fusion 3D dataset was generally the best in all performances, followed by Sentinel-2 3D. Finally, there was the test set ePaddocks for the evaluation and the novel approach was successful and robust.

Segmentation

Wang and Wang[30] proposed a low computing power-requiring and low-sample size-sufficient Bilateral Attention Fusion Network(BAFNet) that segments urban hyperspectral images. The visual attention network extracts the dependency path with depthwise convolution (separate channel-by-channel convolution then stacking on top of each other), depthwise dilation convolution (dilates the kernel with inserted holes between them[37]), and 1x1 convolution (directly taking a single pixel information). The remote-local path, on the other hand, has the ability to locally represent the data and exchange information between them later. Their datasets were Vaihingen and Potsdam images. In Vaihingen, BAFNet performed the best in many metrics and classes except floating-point operations and tree land cover. In Potsdam, it was best in all cases except floating-point operations.

Blushtein-Livnon et al[31] made a cross-check study on evaluating the accuracies of the segmentation and object detection labeling tasks by the human annotators and cautioned that they are simply not 100 % true, and several biases are present, such as being more prone to false negatives than false positives. Moreover, a low number of targets in an image makes it more difficult for annotators to locate/segment/detect the targets.

Broni-Bediako et al.[32] proposed the generalized few-shot semantic segmentation method. They stressed the importance of also predicting base classes well from the background while trying to model the novel class, as sometimes learning a new class better might be detrimental to the base class prediction accuracies. The paper is related to a data science challenge and among the winners, one proposed the SegLand where the base-class predictors were frozen, and the novel class ones were orthogonal. There were four other winners with different approaches and algorithms which were still competitive or best in one or more land cover classes both base and novel.

Diaconu et al.[33] reviewed the methods that can utilize multi-sensor earth observation remote sensing data and deep learning to delineate glaciers and monitor them temporally. They summarized the studies in the standard glacier-extent mapping, glacier area change analysis, rock glacier mapping, and calving-front (ocean neighboring glaciers) detection. Another comprehensive review[34] was on knowledge distillation, transferring large-complex model knowledge to computationally-efficient ones on tasks such as object detection and segmentation.

Ren et al.[35] worked on a tuning method in multimodal model training. There were two targets, one was reducing the required time and resources without compromising accuracy, and the other was higher generalizing power using a lower amount of data. For this, the input data was clustered with minibatch kmeans, and the training data was selected so as to have a general, not overfitted model at the end. They used the

[34] A tensor is decomposed into a core tensor and matrices. It can be thought of as follows: a 2D matrix is decomposed into a specific matrix and unit vectors that in case they are multiplied by the matrix, they will restore the original. "Unitness" is on the vectors in this case.

[35] A non-linear decomposition method, in contrast to linear ones like singular value decomposition. For further information[28].

[36] GitHub repository for their code can be found here.

[37] Let's say we have a 4x4 matrix and the kernel is 2x2, we decide to put one row and column between the kernels. In the end, we will have 5x5 result

AID, UCMerged, and LBREN dataset with multiple baselines, such as, MiniGPTv2, GeoChat, SkyEyeGPT, and InstructBLIP. In the end, their model trained on one-third of the data it was initially trained on only showed a 1% reduction in all performance metrics, and it was the best among the models used in this study.

Liu et al.[36] proposed a new segment-anything model (SAM), pointly-supervised[38] SAM which significantly outperformed the (vanilla) SAM model and several other baselines. The algorithm has an offline prototype generation where it encodes the image in a feature map, generates feature points, and cluster them. Then, an input image is strongly and weakly augmented with image encoders, a point prompt is provided and it is encoded followed by a mask decoder, the mask is refined and the offline prototype cluster is matched with the cluster of this SAM with the self-training step. NWPU VHR-10, HRSID, and WHU datasets were used in this study. Their algorithm was mostly the best in all datasets.

Ma et al.[37] stated the need for a realistic adversarial attack in algorithm development and presented their cloud adversarial example. Their cloud mask was related to the Perlin noise, and they provided clouded image examples.

UAV

Farid et al.[38] proposed the Proximal Policy Optimization to train UAVs in a way that prevent them from colliding with each other, with obstacles, and "discovering" an already covered area. This is a type of reinforcement learning activity where unwanted activities outlined in the previous sentence were penalized. The actor-critic networks (meaning the critics to the actors - UAVs according to their actions) involved long-short term memory (LSTM) and convolutional neural network (CNN). This novel approach can be robust against new, untrained environment-related challenges, and using LSTM rather than CNN improved its overall performance.

Diao et al.[39] focused on aerial remote sensing and developed the Frequency-Enhanced Multi-Head Self-Attention (FE-MSA) algorithm and an affine transformation-based contrastive learning in the pre-learning step. The latter aims to enlarge the distant-small targets in oblique views to augment its detection frequency, and the former improves occluded targets' detection, among other things. Additionally, the paper introduces the first aerial remote sensing (ARS)-specific foundational model RingMo-Aerial, and ARS-Adapter with improved fine-tuning capabilities. Their proposed algorithm has beaten other baselines in several different datasets for many metrics most of the time. The categories in performance comparisons were the pretraining phase, scene classification, change detection, 3D reconstruction, semantic segmentation, object detection, object tracking, ARS-adapter, and ablation experiments. They recommended future research targets for better generalization capability, computational efficiency, and experiments in real-world applications.

[38] Rather than a polygon or shape, a point indicates the location of a specific object, which is more challenging to train and work.

Last week in Environmental Chemistry

Author: Yasin Güray Hatipoğlu

The preprints summarized here were published between September 16 - September 23, 2024. chemRxiv's Earth, Space, and Environmental chemistry preprints are being surveyed, and unfortunately, not many preprints are published under environmental topics in this field.

Deal et al.[40] reported the detailed reaction mechanism behind thiosulfate oxidation by ozone in aerosols. They simulated such conditions with levitated microdroplets to study pH-dependent ozone oxidation of sulfate under atmospheric and other aqueous and wastewater treatment conditions. A quadrupole electrodynamic trap (QET) trapped and reacted the microdroplets and these were transferred to an open-port sampling interface mass spectrometry (OPSI-MS). Roughly, the idea is to take time series measurements of several relevant sulfur and oxygen-bearing molecules with a continuous gas-phase ozone flow, and this ozone was also analyzed by a UV ozone analyzer. Kinetiscope commercial software was used to simulate stochastic kinetics of the reaction and experimental data benchmarked these. While analyzing the droplets, they also considered the necessary corrections on bulk-to-surface transfers. The air-water interface sulphur species were monitored via Deep-UV Second Harmonic Generation spectroscopy. Their results were consistent with the literature while identifying sulfate, trithionate, and tetrathionate as reaction products and sulfite as an intermediate. Conversely, they identified dithionite as a reaction intermediate as well, and again contrary to the literature, they did not detect dithionate. In the end, they were able to propose a reaction mechanism and specific kinetics.

Hariharan and Johnson[41] stated that the mass of atmospheric particles below 10 nm in diameter in size can be measured via MS techniques while they remain intact. The curious review not only provided an in-depth analysis of what has been done but also pointed out clear future recommendations for developing this field.

Last week in Data Science-Applied Statistics

Author: Yasin Güray Hatipoğlu

The preprints summarized here were published between September 16 - September 23, 2024. This is generally from arXiv's stat.ML (and correspondingly cs.LG) cross-list. Large-language model-related, or text-mining and similar studies are omitted, also several less-application oriented studies (they, too, are important, but currently *Science Ascend* can't accomodate to review them within the time constraints.

Data Decomposition

Dharamshi et al.[42] worked on the decomposition of the data with Gaussian distributed components having *a priori* unknown mean and/or standard deviation. They consider more than one realization of a dataset, then also consider a single realization, like a single time-series photometry of a star, or water temperature change against time, etc., where in the latter they stated that separate independent components are not possible, but dependent components are. Indeed, they provided a graphical representation of an algorithm, where with a known covariance matrix, Gaussian data thinning[39] is recommended for getting independent folds. After theoretical underpinnings, applied examples are being provided[40] from a simulation case and University of California Irvine (UCI) encephalography data and reported the added benefit of their approach.

Ameri et al.[43] worked on quantitative ultrasound data with Homodyned-K distribution[41] model and estimated their parameters Bayesian Neural Networks. For decomposing the uncertainty, when they first calculate the standard deviation of multiple model inferences and averaging these over 10 different input datasets, they pick epistemic (model) uncertainty, while doing this in the opposite order gave the aleatoric (data/measurement) uncertainty. In the end, the error in the Bayesian Neural Network estimations was found mostly from the data uncertainty.

Feofanov et al.[44] utilized principal component analysis and neural network-related methods to overcome the resource-intensive problems arising from foundation models for time-series analysis, especially with multi-channel, very high dimensional data. Basically, even pre-trained models are sufficiently resource intensive when requiring to utilize the same high-dimensional data on runtime. Hence, the authors considered a dimensional reduction approach to later work through the data in an efficient and faster (one order of magnitude similar speed enhancement with the same resources) way. They tested this concept with 12 different datasets having more than 9 channels. They evaluated MOMENT[45] and ViT. They implemented a PCA-like method as Patch PCA, where they distribute different timesteps of different temporal datasets by multiplying and then patching into different patches, essentially having two dimensions at the end, one with the multiplication of dataset number and patch number, and the other with the multiplication of number of dimensions and patch window size.

Time Series

Chen et al[46] combined recurrent neural network and stochastic interpolants[42] to develop a time series forecasting algorithm without too much computational cost. In this stochastic mapping, velocity and score functions can be modelled with different functions, such as recurrent neural networks, and conditional generation of such estimations are also possible. In only one in four datasets, their novel algorithm was surpassed by another one, flow-matching. Furthermore, the denoising diffusion probabilistic model, vec-LSTM and score-based generative model were competed against the novel algorithm.

Bhattacharya et al.[47] worked on anomaly detections on time series data. Their Balanced Point Adjustment removes the internal bias of previous algorithms on true positive measuring rate and resulting inflated high score. Different than the point adjustment, their balanced adjustment introduced an island around false positives and additional adjusted scores according to them. Then, they tested the robustness, discrimination power order, and exclusivity and proved that their algorithm is superior on these. They also tested them in simulated data and again, their novel approach was better.

[39] As opposed to splitting data into several parts, data thinning separates value in the data into different columns where if summed, reach the actual value of that row.

[40] The relevant GitHub repository is here..

[41] For ultrasound measurements, this distribution's parameters 1)scatterer clustering parameter, and 2) coherent-to-diffuse scattering ratio are correlated with physical variables

[42] A dependent coupling between two variables/features are reflected in the subsequent time steps according to their probability densities.

Machine Learning / Deep Learning

Jain et al.[48] reported a generalized training method SubSelNet. The idea is choosing such a subset that it will be beneficial for different model architectures, rather than the *status quo* where training - test separation steps are model-specific, and at the same time, all these should be achieved without heavy computing power and resources. Furthermore, SubSelNet is not only able to adapt itself for its trained model architectures and develop a subset appropriate for all, but also it can do this for *unseen* architectures, hence the generalized adjective. In three steps, it encodes the architecture in a way that can be done for other architectures as well, into a vector space. Then, the second step makes an approximate prediction of this trained model. Finally, the subset selection is according to the predictions from the previous step. This last part has transductive and inductive part, in which the former takes the predicted training step and optimizes according to the candidate subset of that model, in contrast, in the latter no optimization takes place in much faster way. They applied these and available pruning, selection-via-proxy, facility location, grad-match, EL2N, and GraNd with the following datasets: FMNIST, CIFAR10, CIFAR100, Tiny-Imagenet-200, Caltech-256. In the end, they report a best-case scenario of efficiency with this approach as a matter of trade-off between runtime, accuracy, and memory usage.

Chatterjee and Sudijono[49] explored the theoretical background of how neural networks (also deep neural networks) can be so complex and learn well and at the same time generalize on test sets. They attributed this to the fact that usually studied images and similar datasets have themselves quite a complex structure so that they can be learnt with complex architectures while having not so much noise so that they are *general* still. They studied this with if- and for-loop containing simple neural programs on prime number estimation in a given range.

Karumuri et al.[50] worked on efficient neural operators and randomized sampling. They stated that the superior performance of DeepONet, a deep neural operator approach, over resource and time-intensive finite element method-like numerical solvers, or physics-informed neural network (PINN) where they are mostly reliable over their set boundaries. DeepONet has two sub-networks, branch and trunk taking different kinds of inputs (branch learns the parameters affecting the value at that specific location/time, and trunk takes the spatial and temporal information separately and directly. The authors in this paper reported a new randomized sampling method that reduces generalization error. They found the approach better and efficient and pointed the requirement of future research on finding the optimal number of evaluation points for better performance in terms of generalization.

Optimization

Biswas et al.[51] implemented a novel way to resolve the Bayesian Optimization issue of being stuck at a local optimum by creating a human-intervened way to generally estimate more than one optima and have a better understanding of the interactions among the parameter space. They applied the method to Piezoresponse spectroscopy data with noise and piezoresponse force microscopy[43] hyperspectral data. They reported that this method has a higher-value scientific output than the normal Bayesian Optimization approach.

Hong and Kratsios[52] worked on a multilayer perceptron (MLP) approach using ReLU (rectified linear unit) that can universally learn the data with large complexity without being too specific, in other words, that can still generalize well. Their report scrutinized the theoretical background of being both accurate and general using the template of the ReLU activated MLP.

[43] An atomic force microscopy variation (which applies force to a surface or chemicals with several different ways and graph its topography) utilizing piezoelectric properties of the substrate material under the investigation.

References

[1] W J Cooper, H R A Jones, R L Smart, S L Folkes, J A Caballero, F Marocco, M C Gálvez Ortiz, A J Burgasser, J D Kirkpatrick, L M Sarro, B Burningham, A Cabrera-Lavers, P E Tremblay, C Reylé, N Lodieu, Z H Zhang, N J Cook, J F Faherty, D García-Álvarez, D Montes, D J Pinfield, A S Rajpurohit, and J Shi. The gaia ultracool dwarf sample–iv. gtc/osiris optical spectra of gaia late-m and l dwarfs. *Monthly Notices of the Royal Astronomical Society*, 534(1):695–725, September 2024.

[2] Aayush Gautam, Juan P. Farias, and Jonathan C. Tan. Star cluster formation from turbulent clumps. iv. protoplanetary disc evolution, 2024, 2409.12378.

[3] JR Barnes, SV Jeffers, CA Haswell, M Damasso, F Del Sordo, F Liebing, M Perger, and G Anglada-Escudé. Identifying activity induced rv periodicities and correlations using central line moments. *Monthly Notices of the Royal Astronomical Society*, page stae2125, 2024.

[4] A Maggio, I Pillitteri, C Argiroffi, D Locci, S Benatti, and G Micela. Xuv irradiation of young planetary atmospheres. results from a joint xmm-newton and hst observation of hip67522. *arXiv preprint arXiv:2409.07229*, 2024.

[5] Alfredo Biagini, Antonino Petralia, Claudia Di Maio, Lorenzo Betti, Emanuele Pace, and Giuseppina Micela. Spot modeling through multiband photometry analysis of v1298 tau, 2024, 2409.11034.

[6] Thanawuth Thanathibodee, Connor Robinson, Nuria Calvet, Catherine Espaillat, Caeley Pittman, Nicole Arulanantham, Kevin France, Hans Moritz Günther, Seok-Jun Chang, and P. Christian Schneider. A model of the c iv $\lambda\lambda$ 1548, 1550 doublet line in t tauri stars, 2024, 2409.10361.

[7] G. J. Ferland, M. Chatzikos, F. Guzmán, M. L. Lykins, P. A. M. van Hoof, R. J. R. Williams, N. P. Abel, N. R. Badnell, F. P. Keenan, R. L. Porter, and P. C. Stancil. The 2017 Release Cloudy.
rmxaa, 53:385–438, October 2017, 1705.10877.

[8] James Muzerolle, Nuria Calvet, and Lee Hartmann. Emission-Line Diagnostics of T Tauri Magnetospheric Accretion. II. Improved Model Tests and Insights into Accretion Physics.
apj, 550(2):944–961, April 2001.

[9] Sai Prathyusha Malla, Dennis Stello, Benjamin T. Monet, Daniel Huber, Marc Hon, Timothy R. Bedding, Claudia Reyes, and Daniel R. Hey. Benchmarking the spectroscopic masses of 249 evolved stars using asteroseismology with tess, 2024, 2409.11736.

[10] Yu-Chia Lin. Unveiling habitable planets: Toy coronagraph tackles the exozodiacal dust challenge. *arXiv preprint arXiv:2409.05797*, 2024.

[11] Pinte, C., Ménard, F., Duchêne, G., and Bastien, P. Monte carlo radiative transfer in protoplanetary disks. *Astronomy and Astrophysics*, 459(3):797–804, 2006.

[12] Paola Marigo, Francesco Addari, Diego Bossini, Alessandro Bressan, Guglielmo Costa, Leo Girardi, Michele Trabucchi, and Guglielmo Volpato. Aesopus 2.1: Low-temperature opacities extended to high pressure, 2024, 2409.10905.

[13] Sedighe Sajadian and Niayesh Afshordi. Simulating self-lensing and eclipsing signals due to detached compact objects in the tess light curves, 2024, 2409.12441.

[14] Ying D. Liu, Huidong Hu, Xiaowei Zhao, Chong Chen, and Rui Wang. A pileup of coronal mass ejections produced the largest geomagnetic storm in two decades, 2024, 2409.11492.

[15] Valentina Penza, Luca Bertello, Matteo Cantoresi, Serena Criscuoli, Lorenza Lucaferri, Raffaele Reda, Simone Ulzega, and Francesco Berrilli. Reconstruction of the total solar irradiance during the last millenium, 2024, 2409.12648.

[16] Karl D. Gordon, G. C. Sloan, Macarena Garcia Marin, Mattia Libralato, George Rieke, Jonathan A. Aguilar, Ralph Bohlin, Misty Cracraft, Marjorie Decleir, Andras Gaspar, David R. Law, Alberto Noriega-Crespo, and Michael Regan. The james webb space telescope absolute flux calibration. ii. mid-infrared instrument imaging and coronagraphy, 2024, 2409.10443.

[17] Sarah Blunt, Jason Jinfei Wang, Lea Hirsch, Roberto Tejada, Vighnesh Nagpal, Tirth Dharmesh Surti, Sofia Covarrubias, Thea McKenna, Rodrigo Ferrer Chávez,

Jorge Llop-Sayson, Mireya Arora, Amanda Chavez, Devin Cody, Saanika Choudhary, Adam J. R. W. Smith, William Balmer, Tomas Stolker, Hannah Gallamore, Clarissa R. Do Ó, Eric L. Nielsen, and Robert J. De Rosa. orbitize! v3: Orbit fitting for the high-contrast imaging community, 2024, 2409.11573.

[18] Christopher Cox, Jakob Haynes, Christopher Duffey, Christopher Bennett, and Julie Brisset. Photothermal spectroscopy for planetary sciences: Mid-ir absorption made easy, 2024, 2409.11626.

[19] Kevin Zemaitis, Rashmi Kumar, Marija Veličković, Dušan Veličković, Wei-Jun Qian, and Ljiljana Paša-Tolić. Identification of remnant oct media artifacts after tissue washing by maldi mass spectrometry imaging. 2024.

[20] Christian Rosales, Noah Lepinsky, Wondewossen Gebeyehu, Karl Wasslen, Benjamin Warnes, Jasmine Chihabi, Jeffrey Manthorpe, and Jeffrey Smith. Improved lc-ms detection of opioids, amphetamines, and psychedelics using trendi. 2024.

[21] Colin Elliott, Denina Simmons, and Theresa Stotesbury. Integrating time since deposition estimation of bloodstains into a dna profiling workflow: A novel approach using fluorescence spectroscopy. 2024.

[22] Amitrajit Mukherjee, Robby Reynaerts, Bapi Pradhan, Sudipta Seth, Andreas T Rösch, Tamali Banerjee, Lata Chouhan, Handong Jin, Christian Sternemann, Michael Paulus, et al. Machine learning for microscopy data analysis: Toward real-time optical and electrical characterization of submicron materials. 2024.

[23] Huijie Zhang, Mohamed Saadeldin, Darren Buesen, Hamzah Elfaitory, Jakob Burger, Vincent Friebe, Jonas Honacker, Tobias Voepel, Alaa Oughli, and Nicolas Plumeré. A universal oxygen scavenger for oxidase-based biosensors. 2024.

[24] José Antonio Díaz Olivares, Stef Grauwels, Xinyue Fu, Ines Adriaens, Wouter Saeys, Ryad Bendoula, Jean-Michel Roger, and Ben Aernouts. Temperature correction of near-infrared spectra of raw milk. 2024.

[25] Igor d'Anciães Almeida Silva. Predicting structural groups of small molecules from 1h nmr spectral features using common machine learning classifiers. 2024.

[26] Ian Seymour, Fiona Barry, James F Rohan, and Alan O'Riordan. On-chip electrochemical detection of dissolved oxygen: Eliminating the requirement for permeable selective membrane. 2024.

[27] He Wang, Yang Xu, Zebin Wu, and Zhihui Wei. Unsupervised hyperspectral and multispectral image blind fusion based on deep tucker decomposition network with spatial-spectral manifold learning, 2024, 2409.09670.

[28] Marina Meilă and Hanyu Zhang. Manifold learning: what, how, and why, 2023, 2311.03757.

[29] Foivos I. Diakogiannis, Zheng-Shu Zhou, Jeff Wang, Gonzalo Mata, Dave Henry, Roger Lawes, Amy Parker, Peter Caccetta, Rodrigo Ibata, Ondrej Hlinka, Jonathan Richetti, Kathryn Batchelor, Chris Herrmann, Andrew Toovey, and John Taylor. Tackling fluffy clouds: field boundaries detection using time series of s2 and/or s1 imagery, 2024, 2409.13568.

[30] Wentao Wang and Xili Wang. Bafnet: Bilateral attention fusion network for lightweight semantic segmentation of urban remote sensing images, 2024, 2409.10269.

[31] Roni Blushtein-Livnon, Tal Svoray, and Michael Dorman. Performance of human annotators in object detection and segmentation of remotely sensed data, 2024, 2409.10272.

[32] Clifford Broni-Bediako, Junshi Xia, Jian Song, Hongruixuan Chen, Mennatullah Siam, and Naoto Yokoya. Generalized few-shot semantic segmentation in remote sensing: Challenge and benchmark, 2024, 2409.11227.

[33] Codruţ-Andrei Diaconu, Konrad Heidler, Jonathan L. Bamber, and Harry Zekollari. Multi-sensor deep learning for glacier mapping, 2024, 2409.12034.

[34] Yassine Himeur, Nour Aburaed, Omar Elharrouss, Iraklis Varlamis, Shadi Atalla, Wathiq Mansoor, and Hussain Al Ahmad. Applications of knowledge distillation in remote sensing: A survey, 2024, 2409.12111.

[35] Yi Ren, Tianyi Zhang, Zhixiong Han, Weibin Li, Zhiyang Wang, Wenbo Ji, Chenhao Qin, Chenbin Liang, and Licheng Jiao. A novel adaptive fine-tuning algorithm for multimodal models: Self-optimizing classification and selection of high-quality datasets in remote sensing, 2024, 2409.13345.

[36] Nanqing Liu, Xun Xu, Yongyi Su, Haojie Zhang, and Heng-Chao Li. Pointsam: Pointly-supervised segment anything model for remote sensing images, 2024, 2409.13401.

[37] Fei Ma, Yuqiang Feng, Fan Zhang, and Yongsheng Zhou. Cloud adversarial example generation for remote sensing image classification, 2024, 2409.14240.

[38] Ali Moltajaei Farid, Jafar Roshanian, and Malek Mouhoub. On-policy actor-critic reinforcement learning for multi-uav exploration, 2024, 2409.11058.

[39] Wenhui Diao, Haichen Yu, Kaiyue Kang, Tong Ling, Di Liu, Yingchao Feng, Hanbo Bi, Libo Ren, Xuexue Li, Yongqiang Mao, and Xian Sun. Ringmo-aerial: An aerial remote sensing foundation model with a affine transformation contrastive learning, 2024, 2409.13366.

[40] Alexandra Deal, Alexander Prophet, Franky Bernal, Richard Saykally, and Kevin Wilson. A detailed reaction mechanism for thiosulfate oxidation by ozone in aqueous environments. 2024.

[41] Annapoorani Hariharan and Christopher Johnson. Laboratory mass spectrometry of intact atmospherically-relevant particles. 2024.

[42] Ameer Dharamshi, Anna Neufeld, Lucy L. Gao, Jacob Bien, and Daniela Witten. Decomposing gaussians with unknown covariance, 2024, 2409.11497.

[43] Dorsa Ameri, Ali K. Z. Tehrani, Ivan M. Rosado-Mendez, and Hassan Rivaz. Uncertainty decomposition and error margin detection of homodyned-k distribution in quantitative ultrasound, 2024, 2409.11583.

[44] Vasilii Feofanov, Romain Ilbert, Malik Tiomoko, Themis Palpanas, and Ievgen Redko. User-friendly foundation model adapters for multivariate time series classification, 2024, 2409.12264.

[45] Mononito Goswami, Konrad Szafer, Arjun Choudhry, Yifu Cai, Shuo Li, and Artur Dubrawski. Moment: A family of open time-series foundation models, 2024, 2402.03885.

[46] Yu Chen, Marin Biloš, Sarthak Mittal, Wei Deng, Kashif Rasul, and Anderson Schneider. Recurrent interpolants for probabilistic time series prediction, 2024, 2409.11684.

[47] Debarpan Bhattacharya, Sumanta Mukherjee, Chandramouli Kamanchi, Vijay Ekambaram, Arindam Jati, and Pankaj Dayama. Towards unbiased evaluation of time-series anomaly detector, 2024, 2409.13053.

[48] Eeshaan Jain, Tushar Nandy, Gaurav Aggarwal, Ashish Tendulkar, Rishabh Iyer, and Abir De. Efficient data subset selection to generalize training across models: Transductive and inductive networks, 2024, 2409.12255.

[49] Sourav Chatterjee and Timothy Sudijono. Neural networks generalize on low complexity data, 2024, 2409.12446.

[50] Sharmila Karumuri, Lori Graham-Brady, and Somdatta Goswami. Efficient training of deep neural operator networks via randomized sampling, 2024, 2409.13280.

[51] Arpan Biswas, Rama Vasudevan, Rohit Pant, Ichiro Takeuchi, Hiroshi Funakubo, and Yongtao Liu. Sane: Strategic autonomous non-smooth exploration for multiple optima discovery in multi-modal and non-differentiable black-box functions, 2024, 2409.12295.

[52] Ruiyang Hong and Anastasis Kratsios. Bridging the gap between approximation and learning via optimal approximation by relu mlps of maximal regularity, 2024, 2409.12335.