Thermodynamic Profiling Through ASSIST Observations and TROPoe Retrievals

Stefano Letizia,[1] Vincent Michaud-Belleau,[2] David D. Turner,[3] and Aliza Abraham[1]

1 National Renewable Energy Laboratory
2 LR Tech Inc.
3 National Oceanic and Atmospheric Administration

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated under Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5000-92509
November 2025

Thermodynamic Profiling Through ASSIST Observations and TROPoe Retrievals

Stefano Letizia,[1] Vincent Michaud-Belleau,[2] David D. Turner,[3] and Aliza Abraham[1]

1 National Renewable Energy Laboratory
2 LR Tech Inc.
3 National Oceanic and Atmospheric Administration

Suggested Citation

Letizia, Stefano, Vincent Michaud-Belleau, David D. Turner, and Aliza Abraham. 2025. *Thermodynamic Profiling Through ASSIST Observations and TROPoe Retrievals*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5000-92509. https://www.nrel.gov/docs/fy26osti/92509.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated under Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5000-92509
November 2025

15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by NREL for the U.S. Department of Energy (DOE), operated under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Cover photos (clockwise from left): Josh Bauer, NREL 61725; Visualization from the NREL Insight Center; Getty-181828180; Agata Bogucka, NREL 91683; Dennis Schroeder, NREL 51331; Werner Slocum, NREL 67842.

Acknowledgments

The authors acknowledge the U.S. Department of Energy Wind Energy Technologies Office and the NOAA Atmospheric Science for Renewable Energy Program for funding this work, as well as the prominent scientists who championed the importance of thermodynamic profiles for wind energy and atmospheric research. We are immensely grateful to the crew of field researchers and engineers, particularly Andy Scholbrock, Dave Jager, Simon Thao, Dave Schwartz, Eve Cinquino, and Brian Cox, who worked hard to make the deployments of all the instruments possible. Dan Zalkind, Nicola Bodini, and the Wind Data Hub team are also acknowledged for their essential contribution to data management, as well as Bianca Adler and Laura Bianco in NOAA's Physical Sciences Laboratory for insightful conversations on the topic. Finally, we thank LR Tech for providing impeccable technical support and the members of the communities around the American WAKE Experiment and Wind Forecast Improvement Project-3 sites for their help and hospitality. The views expressed herein do not necessarily represent the views of the DOE, NOAA, or the U.S. Government.

Disclaimer of Endorsement

List of Acronyms

ABB	Ambient Blackbody
AERI	Atmospheric Emitted Radiance Interferometer
ASSIST	Atmospheric Sounder Spectrometer by Infrared Spectral Technology
AWAKEN	American WAKE experimeNt
DFS	Degrees of Freedom of the Signal
DFT	Discrete Fourier transform
FFOV	Finite Field of View
IRS	Infrared Spectrometer
HBB	Hot Blackbody
LWP	Liquid Water Path
NOAA	National Oceanic and Atmospheric Administration
NREL	National Renewable Energy Laboratory
OE	Optimal Estimation
PBLH	Planetary Boundary Layer Height
QC	Quality Control
RH	Relative Humidity
TROPoe	Tropospheric Remotely Observed Profiling via Optimal Estimation
WFIP3	Wind Forecast Improvement Project-3

Nomenclature

α	angle of incidence of radiation
ΔE	energy jump associated with state transition
ε	instrumental noise or error
γ	underrelaxing factor in TROPoe
$\hat{}$	optimal estimate
$\kappa_{\tilde{\nu},\alpha}$	spectral optical thickness along a certain direction
ω	spatial angular frequency
ϕ	phase shift
$\Re$	spectral responsivity of the radiometry calibration
σ	standard deviation or standard deviation of the error
θ	potential temperature
$\tilde{\nu}$	spatial wavenumber
ξ	along-beam coordinate
A	A-kernel matrix
a	A-kernel (one-dimensional problem)
$a_{\tilde{\nu}}$	spectral absorption coefficient
B	spectral radiance (single-sided)
B_0	blackbody emission
B_{DS}	spectral radiance (double-sided)
c	speed of light
d	raw voltage at the detector
E	electric field
$e_{\tilde{\nu}}$	spectral emissivity
F	forward problem
f	temporal frequency
G	gain matrix
g	gain (one-dimensional problem)
h	Plank’s constant
$I(s)$	interferogram as a function of optical path difference
I_N	identity matrix with rank N
K	Jacobian of forward model
k	instrument sensitivity (one-dimensional problem)
k_B	Boltzmann’s constant
P	pressure
P_{rad}	monochromatic radiation power

R_{eff}	liquid water effective radius
S	covariance matrix
s	optical path difference
$s_{\max}$	maximum optical path difference
T	air temperature
t	time coordinate
W	molar concentration
x	state vector
x_a	prior mean
y	observation vector
z	height above the ground
p.d.f. or p	probability density function
r.u.	radiance units, or $\frac{\mathrm{mW}}{\mathrm{m}^2\ \mathrm{sr}\ \mathrm{cm}^{-1}}$

Executive Summary

The U.S. Department of Energy, in particular, the National Renewable Energy Laboratory (NREL), and the National Oceanic and Atmospheric Administration (NOAA) have recently invested significant financial and human resources to improve measurement capabilities for wind energy applications. One of the newest additions to NREL's fleet of atmospheric instruments is the ground-based infrared spectrometer called the Atmospheric Sounder Spectrometer by Infrared Spectral Technology (ASSIST-II by LR Tech Inc., referred to in the remainder of the report simply as an ASSIST). The raw measurements from an ASSIST carry information about the temperature and moisture profiles in the atmosphere, which can be estimated through the Tropospheric Remotely Observed Profiling via Optimal Estimation (TROPoe) algorithm. At the time this report was written, the NREL ASSIST+TROPoe data have already been used in several journal publications (Bodini et al. 2024; Moriarty et al. 2024; Abraham et al. 2024; Radünz et al. 2024), with many more expected in future decades from the American WAKE experimeNt (AWAKEN, Moriarty et al. 2024), the Wind Forecast Improvement Project-3 (WFIP3, NOAA Physical Sciences Laboratory 2024), and future field experiments.

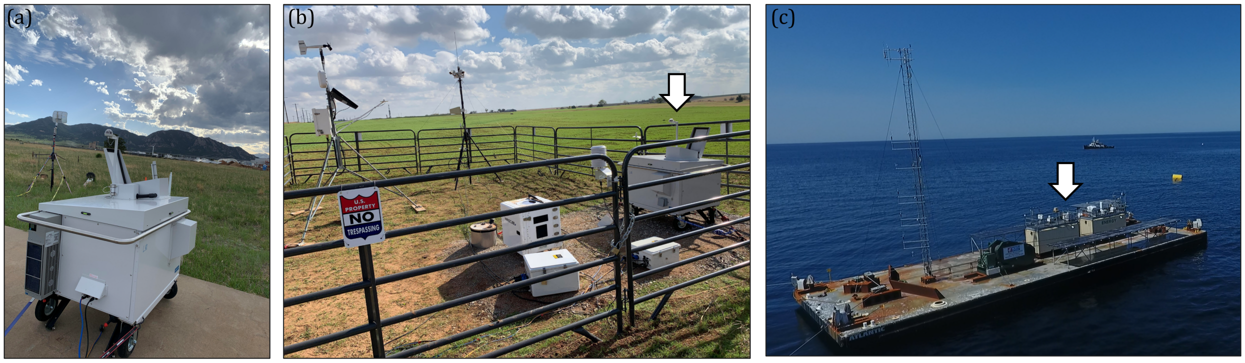

Figure A. Photos from different deployments of the ASSISTs: (a) at NREL's Flatirons Campus, 2022, before the AWAKEN campaign; (b) at site B for AWAKEN in Oklahoma, 2023; (c) on the WFIP3 barge, off the coast of Rhode Island, 2024.

The proper operation of the ASSIST+TROPoe and its outputs require a deep understanding of the underpinning theoretical principles, which touch upon several disciplines such as spectrometry, radiative transfer, and optimal estimation theory. The ASSIST+TROPoe system is also different and more complex than other sensors that wind energy and atmospheric scientists are generally accustomed to. This report represents a concise overview of the essential theoretical and technical background required for an informed and scientifically sound use of the hardware, software, and data associated with thermodynamic profiling.

This report includes four chapters that address the following topics:

1. Fundamentals of spectrometry, which describes the operation of the ASSIST and is intended to familiarize the reader with the concepts required to troubleshoot the instrument and ensure operation within specifications.
2. Fundamentals of radiative transfer through a participating medium, where we review the main phenomena affecting the spectral radiance sensed by the ASSIST and serves as the basis for the TROPoe retrievals.
3. Retrieval of thermodynamic profiles, which focus on the optimal estimation of the atmospheric state based on spectral observations, complemented by a succinct theoretical guide to TROPoe.
4. Case studies, which contain some useful examples of how thermodynamic profile data can be read, quality-controlled, and used.

For the reader who is unfamiliar with the topic and is merely interested in deploying the ASSIST or using TROPoe data, some discussions of theoretical concepts may sound redundant. However, we strongly recommend reading the report in its entirety, as all the topics covered have important implications for the resulting data usage. To facilitate navigation through the most technical sections, we singled out the main "takeaways" that carry the most relevant message.

Table of Contents

List of Figures

List of Tables

1 Introduction

Since the advent of the wind energy industry, researchers have understandably spent their greatest efforts in characterizing the wind velocity field, which is the most bankable quantity and the one driving the wind energy generation. This focus on wind velocity has resulted in great advances in wind sensors, from traditional meteorological masts to more modern wind lidars and radars, all of which are now profusely used to assess the wind resource.

However, the Earth's atmosphere is governed by the laws of mass, momentum, and energy conservation, and the velocity vector represents just one of several interdependent quantities. The wind field is indeed dynamically coupled with other physical variables, primarily temperature and moisture (the **thermodynamic profiles**), but also with cloud properties, gas concentrations, and airborne particulate and hydrometeor distributions, which cannot be easily disentangled.

These quantities become even more important as the focus of wind research moves towards more challenging topics such as the effect of stability on shear and wakes, wind plant blockage, flow in complex terrain, low-level jets, gravity waves, and the impact of wake-generated turbulence on climate. Therefore, improved knowledge of, for instance, temperature and moisture profiles in the planetary boundary layer could benefit wind research and atmospheric science in general through:

1. Better characterization of local atmospheric stability.
2. Improved boundary conditions for microscale numerical models.
3. More comprehensive data assimilation for numerical weather prediction tools.
4. Accurate estimation of air density for turbine power performance tests.
5. Better characterization of the impact of wind plants and their wakes on local climate and vegetation.
6. More comprehensive description of low-level jets, gravity waves, katabatic flows, and other thermally-driven atmospheric phenomena.

Based on the realization that measuring and modeling winds alone are not enough (Veers et al. 2019; Shaw et al. 2022), the U.S. Department of Energy has funded initiatives to incorporate instruments capable of measuring thermodynamic profiles into the suite of sensors used for wind energy applications.

To meet this objective, the National Renewable Energy Laboratory (NREL) and the National Oceanic and Atmospheric Administration (NOAA) have acquired ground-based infrared spectrometers (IRSs) such as the Atmospheric Sounder Spectrometer by Infrared Spectral Technology (ASSIST) (Michaud-belleau et al. 2025), which has been deployed at the American WAKE experimeNt (AWAKEN) and Wind Forecast Improvement Project-3 (WFIP3) sites. ASSISTs and IRSs in general sense the spectra of downwelling infrared radiation at ground level. The spectral radiance carries information about the thermodynamic profile above the instrument, mostly in terms of temperature, moisture content, and cloud properties. However, estimating the desired thermodynamic quantities and their uncertainty is a complex operation that requires in-depth knowledge of the physics of infrared radiation as it travels through the atmosphere. Furthermore, unlike other remote sensing devices like lidars and radars, the ASSIST is a passive instrument in the sense that it does not send any pulse towards the measuring volume. This makes the reconstruction of a profile—i.e., the distribution of the target quantity as a function of the range—less intuitive than it is for lidars, for example.

The ASSIST raw spectra can be converted into thermodynamic profiles by using a retrieval method, which in our case is Tropospheric Remotely Observed Profiling via Optimal Estimation (TROPoe). TROPoe is an optimal estimation (OE) algorithm that iteratively runs a radiative transfer model and adjusts its input (the thermodynamic profiles) until its output (the spectral radiance) matches the observations (Turner and Löhnert 2014; Turner and Blumberg 2019). For us, the observations are mainly downwelling infrared spectral radiances detected by the ASSIST in specific spectral bands. Ancillary observations such as cloud base height and surface meteorological data can also be provided to TROPoe to increase its accuracy. The retrieval is done in an OE framework (Rodgers 2000), which implies that (1) the final solution is the most likely given the current observations and historical climatology and (2) we can produce online quantification of the uncertainty (or at least part of it). Figure 1 summarizes the process for obtaining thermodynamic profiles based on radiance observations.

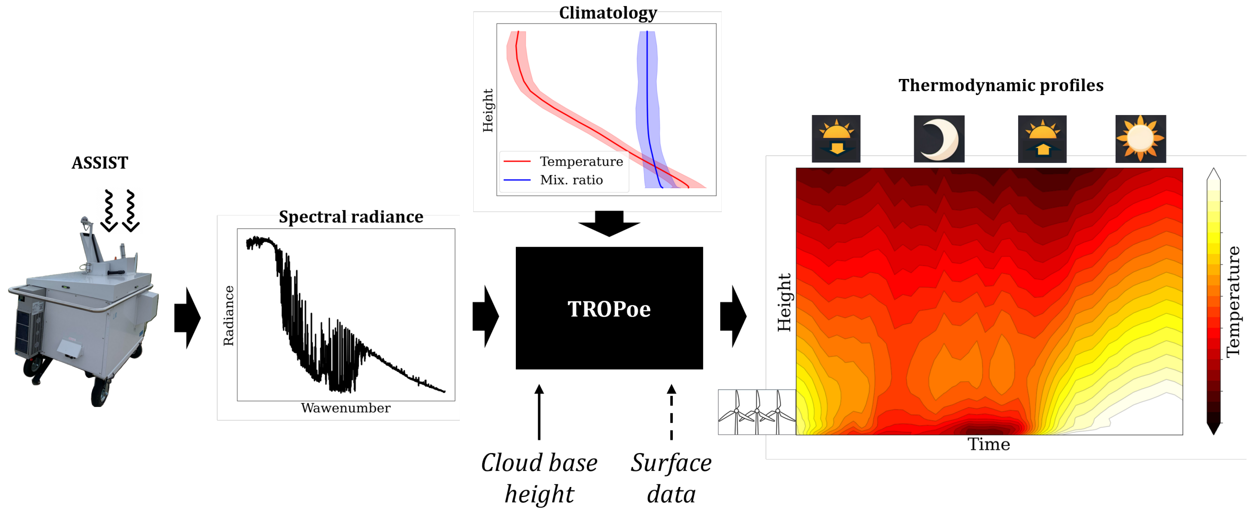

Figure 1. General workflow of thermodynamic profiling using ASSIST+TROPoe.

Already at this point, the complexity of the ASSIST+TROPoe environment becomes evident. It is even more obscure to the wind scientist unfamiliar with spectrometry and atmospheric sounding when dealing with the operation of the ASSIST and the TROPoe processing or data use. Exotic terms like "black bodies," "imaginary radiance," "prior and posterior," "averaging kernel," etc., become ubiquitous and oftentimes essential. This report is intended to be a user's manual for anyone involved with the deployment of an ASSIST, the use of TROPoe, and/or the analysis of its output. After reading through the whole report, the reader will be able to:

1. Understand the measuring principle and the calibration process of an ASSIST.
2. Read a calibration report from the ASSIST's manufacturer.
3. Understand the main sources of instrumental errors in an ASSIST and how to minimize their impact.
4. Understand the basics of the radiation transfer physics that TROPoe uses.
5. Understand the OE principles behind TROPoe.
6. Read the TROPoe output.
7. Carry out error analysis for thermodynamic profiles.

This report adds just minor novelties to the existing literature, and its main merit is to distill a large volume of scientific and technical knowledge from the recent past (published and unpublished) that makes accurate thermodynamic profiling possible today.

An important question to address before embarking on this short but dense journey is: Why did we use ASSIST+TROPoe as the preferred technology for thermodynamic profiling, specifically for wind energy? First, the use of remote sensing technology enables the detection of thermodynamic profiles with far greater temporal resolution and less logistical cost than in situ sensors such as radiosondes and instrumented aircraft. Moreover, IRSs have shown superior capabilities with respect to other remote sensing tools when it comes to detecting thermodynamic profiles (and in particular temperature) near the Earth's surface where turbines operate. In this regard, several studies compared the accuracy of thermodynamic profiles retrieved through IRSs such as the ASSIST and other technologies such as microwave radiometers (Blumberg et al. 2015; Turner and Lohnert 2021; Bianco et al. 2024), Raman lidars (Turner and Blumberg 2019), radio acoustic sounding systems (Bianco et al. 2024), and differential absorption lidars (Turner and Lohnert 2021). The main pros and cons of each of these profiling instruments compared to an IRS are summarized in Table 1. Although each instrument has its niche of application, it is evident how the IRS prevails when it comes to simultaneous profiling of temperature and water vapor, especially in the lowest few hundred meters above the surface of the Earth.

	Advantages vs. IRS	Disadvantages vs. IRS
Microwave radiometer	• Can sense better above clouds	• Calibration needs • Poorer vertical resolution • Smaller information content (see Section 4.1.3)
Raman lidar	• More accuracy in water vapor • Better vertical resolution	• No temperature
Differential absorption lidar	• More accuracy in water vapor • Better vertical resolution	• No temperature • Blind zone near surface • Smaller range • Calibration needs
Radio acoustic sounding system	• More accuracy in virtual temperature • Better vertical resolution	• No water vapor • More complex system • Blind zone near surface • Smaller range • Lower sampling rate • Noise • Data loss in high winds

Table 1. Comparison Between IRS and Other Similar Technologies.

The choice of an OE method like TROPoe, which is a physical retrieval, rather than a method based on machine learning or statistical regression of past data, is advantageous for at least four reasons:

1. The retrieved profiles will provide "closure" with the observed radiance spectra; i.e., the profiles are consistent with the measurement (this is not guaranteed with a statistical/machine learning method)
2. It provides the uncertainty of the retrieval.
3. It benefits from improvements of the embedded radiative transfer model.
4. It allows identification of the physical mechanisms driving the error more easily than a black-box/data-driven approach.
5. Its accuracy is less sensitive to the size and origin of the historical dataset provided than a purely data-driven method.

The disadvantages of TROPoe are its computational costs, the limited vertical resolution (inherent to the physics of radiation, see Section 3.1) compared to in-situ measurements, and the fact that the solution could converge to nonphysical profiles if the setup is inadequate. However, an informed use of ASSIST and TROPoe (which is what we hope to facilitate through this report) can guarantee the estimation of high-quality datasets of thermodynamic profiles with known uncertainties and representativeness, which is a powerful tool for wind and weather researchers to unlock new knowledge of the atmosphere and its dynamics.

Thermodynamic profiling through an IRS and a physical retrieval based on optimal estimation has a long track record. The bulk of the development and application of this technique occurred as a part of the Atmospheric Radiation Measurement (ARM) Program (Mlawer and Turner 2016). The instrument used in the ARM was the AERI (Atmospheric Emitted Radiance Interferometer) (Knuteson et al. 2004a, 2004b), while the retrieval algorithm was called AERIoe, which are the ancestors of current ASSISTs and TROPoe, respectively. AERI observations were instrumental in enhancing the fidelity of radiative transfer models of the atmosphere and ultimately improving the thermodynamic retrieval methods (Turner et al. 2004; Mlawer and Turner 2016). To date, ASSIST+TROPoe systems have been used to map spatial gradients of thermodynamic properties (Wagner et al. 2022), characterize precipitation in mountainous environments (Boer et al. 2023; Adler et al. 2023), investigate the genesis of tornadoes (Kosiba et al. 2024), characterize low level jets (Radünz et al. 2024), and study interaction of the atmospheric boundary layer with wind power plants (Abraham et al. 2024; Moriarty et al. 2024).

This report is organized as follows: In Section 2, we discuss the operating principle of the ASSIST and build the foundation for understanding the raw data used for thermodynamic profiling and their uncertainty; in Section 3, we review the physics of infrared radiation as it travels through the atmosphere and how spectral radiance carry thermodynamic information; in Section 4, we look into the theory of OE and then focus on TROPoe's equations, output, and limitations; and in Section 5 we provide basic examples of ASSIST and TROPoe data important for general use.

2 Fundamentals of Spectrometry

This chapter describes the hardware components and data processing inside the ASSIST. We start with the principles governing an idealized spectrometer and then progressively include the several limitations of the actual instrument and the associated mitigation strategies.

2.1 The Ideal Interferometer

Highly accurate observations of the spectra of downwelling infrared radiance are essential for thermodynamic profiling using TROPoe. Therefore, the core of the ASSIST is an interferometer of the Michelson type, an instrument that uses constructive and destructive interference of the radiance to characterize its spectral content. A generic Michelson interferometer is shown in Figure 2. The output of a scanning interferometer is the **interferogram** of the incoming radiation, namely, the power of the incoming radiation superposed to time-shifted versions of itself for different lags. The interferogram provides a direct measurement of the autocorrelation function of the signal—and ultimately its spectrum through a Fourier transform—by leveraging the Wiener-Khinchin theorem (Chatfield and Xing 2019). To show the principle of the Michelson interferometer, let us consider a monochromatic electromagnetic plane wave traveling through its aperture and perfectly parallel to its optical axis. In this case, using phasor notation, we have the following expression for the associated electric field:

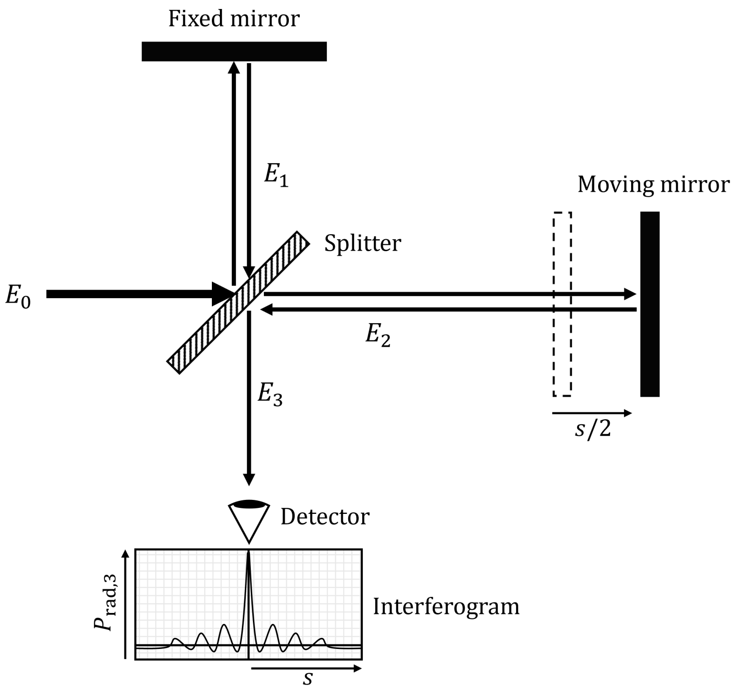

Figure 2. Schematic of a generic Michelson interferometer. The Es indicate electromagnetic waves at different stages and s is the optical path difference.

$$E_0 = |E| e^{j\omega(\xi - ct)}, \tag{2.1}$$

where $|E|$ is the amplitude, $j \equiv \sqrt{-1}$, ω is the spatial angular frequency, c is the speed of light, ξ is a location in space along the beam, and t is time. We recall that ω is related to the **wavenumber**, $\tilde{\nu}$, as $\omega = 2\pi\tilde{\nu}$, while the temporal frequency is $f = \tilde{\nu}c$.

In general, the radiation power, P_{rad}, is proportional to the mean of the real part of the signal over a cycle of duration $1/f$, also equal to the square of the amplitude divided by 2, according to:

$$P_{\mathrm{rad}} \propto f \int_0^{\frac{1}{f}} \mathrm{Real}(E)^2 dt = \frac{1}{2}|E|^2. \tag{2.2}$$

In the most idealized scenario of a 50:50 lossless (i.e., energy-conserving) splitter, half of the radiation power is reflected in one direction and fully reflected back by the fixed mirror as the signal E_1. The other half passes through the splitter and is fully reflected by the moving mirror as a signal E_2. The different travel path lengths induce a phase shift between the two waves, ϕ. Without loss of generality, we pose the phase of E_1 to 0. Then, both these waves are transmitted (E_1) or reflected (E_2) by the splitter and finally superposed at the detector as a signal of the form

$$E_3 = \frac{E_1}{\sqrt{2}} + \frac{E_2}{\sqrt{2}} = \frac{|E|}{2}\left(e^{j\omega(\xi-ct)} + e^{j[(\omega(\xi-ct)+\phi)]}\right) = E_0\,\frac{1+e^{j\phi}}{2}, \tag{2.3}$$

where factors $\sqrt{2}$ appear due to the 50% reduction in the wave power (or $\sqrt{2}$ reduction in amplitude according to Eq. 2.2) caused by the second pass through the splitter. We see how the original wave is now undergoing complex amplitude modulation, where we immediately recognize a perfect constructive interference for $\phi = 0$ and a perfect destructive interference for $\phi = \pi$.

To understand what happens in the general case, we need to do some calculations. The phase delay ϕ is proportional to the additional travel time required by light to cover twice the displacement of the moving mirror. It is convenient to define the **optical path difference**, that is, the difference in the distance traveled by E_2 and E_1 before recombination, as s. Therefore, we have $\phi = \omega s = 2\pi\tilde{\nu}s$. Finally, after some manipulations, we can prove that the power of the signal measured by the detector is:

$$P_{\text{rad},3} \propto \frac{|E|^2}{4}\left(1+\cos(2\pi\tilde{\nu}s)\right). \tag{2.4}$$

In words, Equation 2.4 reads "the variable part (i.e., without the constant DC component) of the power at the detector or interferogram, $I(s)$, of a monochromatic wave is a sinusoidal function of the optical path difference, s, with wavenumber $\tilde{\nu}$". From an energy conservation standpoint, for zero phase shift, $P_3 = P_0$, i.e., all the input power reaches the detector. Conversely, when the signals are in perfect counter-phase (e.g., $s = \pm 0.5\ \tilde{\nu}^{-1}$) the detector receives no power. Since we assumed a lossless interferometer, where did the input power go? The answer is simple: The missing power at the detector is reflected back to the input and does not contribute to the measurement.

In the more realistic case of a multi-chromatic radiation with power spectral density or spectral radiance $B(\tilde{\nu})$,

$$I(s) = \int_0^\infty B(\tilde{\nu})\cos(2\pi\tilde{\nu}s)\,d\tilde{\nu}, \tag{2.5}$$

where the proportionality constant has been dropped for simplicity. We can extend the integral to negative frequencies by using the double-sided spectral radiance B_{DS}. Finally, through a change of variable and recalling that B_{DS} is a real even function, we get a familiar expression:

$$I(s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} B_{\text{DS}}(\omega)e^{j\omega s}\,d\omega, \tag{2.6}$$

which says that the interferogram is the inverse Fourier transform of the spectral radiance. Thus, we can obtain the spectral radiance of the radiation by performing the Fourier transform of the interferogram:

$$B_{\text{DS}}(\omega) = \int_{-\infty}^{\infty} I(s)e^{-j\omega s}\,d\omega. \tag{2.7}$$

This analysis was done for an ideal interferometer. However, a real one is affected by several limitations that are discussed in the following sections.

2.2 Limitation 1: Radiometric Calibration

As is common in metrology, the measured quantity does not correspond directly to the quantity of interest. The ASSIST is no exception, and recording the interferogram requires two detectors that translate the radiation intensity of the interferometer (quantity of interest) into a voltage that is the actual measurement. The long-wave "channel A" detector is an HgCdTe sensor that senses wavenumbers from 525 to 1,825 cm^{-1}. The shortwave "channel B" detector is an InSb sensor and senses wavenumbers from 1,720 to 3,300 cm^{-1}. In the following, we assume a perfectly linear response of both detectors; we call the raw signal at the detector $d(s)$ and its Fourier transform $D(\tilde{\nu})$.

The raw reading from the detectors is converted online into an estimate of the spectral radiance through the so-called **radiometric calibration**. The radiometric calibration uses the known radiative properties of two **blackbodies** and the linearity of the detector response and the Fourier transform operator to build a linear regression between the detector output and the spectral energy at each wavenumber. One blackbody is kept at 60°C and is named hot blackbody (HBB); the other blackbody is not thermally controlled and is therefore named ambient blackbody (ABB). The emission of an ideal blackbody follows Planck's law at thermal equilibrium:

$$B_0(\tilde{\nu},T) = 2\cdot 10^{11} hc^2\tilde{\nu}^3 \left(e^{\frac{100\cdot hc\tilde{\nu}}{k_B T}} - 1\right)^{-1}, \tag{2.8}$$

where:

- B_0 is the blackbody (single-sided) spectral radiance in $\frac{\text{mW}}{\text{m}^2\ \text{sr cm}^{-1}}$ (power per unit surface, per unit solid angle, per unit wavenumber band), also called **radiance units** (r.u.), which is the unit used by the ASSIST
- $\tilde{\nu}$ is the wavenumber in cm^{-1}
- $h = 6.626070\cdot 10^{-34}$ J s is the Planck constant
- $c = 2.997924\cdot 10^{8}$ m s^{-1} is the speed of light
- $k_B = 1.380649\cdot 10^{-23}$ J K^{-1} is the Boltzmann constant
- T is the temperature in K.

The "real" blackbodies installed on the ASSIST do not exactly follow Eq. 2.8 because, due to their emissivity being $e_{\tilde{\nu}} < 1$, they also reflect a small fraction of energy from the surroundings structure at a uniform temperature T_r:

$$B_{\text{BB}}(\tilde{\nu}) = e_{\tilde{\nu}} B_0(\tilde{\nu}) + (1 - e_{\tilde{\nu}}) B_0(\tilde{\nu}, T_r). \tag{2.9}$$

The emissivity is generally > 0.99 in the normal direction sensed by the interferometer, due to the black coating and the cavity effect achieved by the special geometry of the blackbodies.

Now that we have an equation to estimate the spectra from the HBB and ABB and we can measure interferograms from HBB, ABB, and the sky scene, we have all the ingredients to formulate the radiometric calibration workflow, which includes these steps:

1. A rotating scene-selection mirror deflects the emission of the HBB into the interferometer. The detector collects the raw signal $d_{\text{HBB}}(s)$ in volts and calculates and stores its Fourier transform, $D_{\text{HBB}}(\tilde{\nu})$. The temperature of the blackbody, T_{HBB} is also recorded.
2. The rotating mirror deflects the emission of the ABB into the interferometer and stores $D_{\text{ABB}}(\tilde{\nu})$ and T_{ABB}.
3. The mirror moves to sky view and records the downwelling infrared radiation coming from the atmosphere in the form of $D_{\text{sky}}(\tilde{\nu})$.
4. Step 2 is repeated.
5. Step 1 is repeated.
6. Step 3 is repeated.

The aim of the sequence HBB-ABB-sky-ABB-HBB-sky is to make sure that each sky view is bracketed by views of both blackbodies. The sky spectral radiance is finally calibrated using the information from both bracketing views of the blackbodies.

Based on the scheme shown in Figure 3 we have:

$$B_{\text{DS}}(\tilde{\nu}) = \overbrace{\frac{B_{\text{DS,HBB}}(\tilde{\nu}) - B_{\text{DS,ABB}}(\tilde{\nu})}{D_{\text{HBB}}(\tilde{\nu}) - D_{\text{ABB}}(\tilde{\nu})}}^{\Re(\tilde{\nu})^{-1}} \left(D_{\text{sky}}(\tilde{\nu}) - D_{\text{ABB}}(\tilde{\nu})\right) + B_{\text{DS,ABB}}(\tilde{\nu}), \tag{2.10}$$

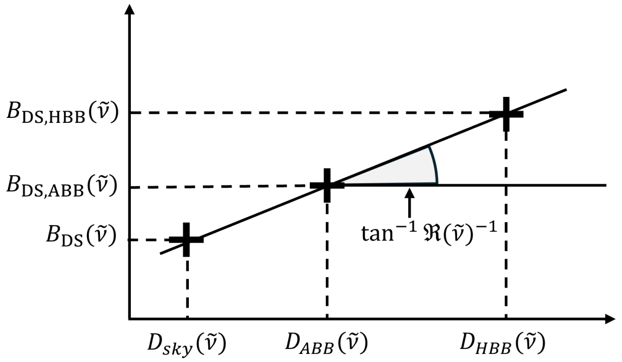

Figure 3. Schematic of the radiometric calibration.

where double-sided blackbody spectra are obtained by simply mirroring Eq. 2.8.

It is noteworthy that any internal emission and phase lag due to electrical and optical dispersion cancel out, provided that they are invariant between the three different views. This is ensured by a careful design of the geometry of the back bodies and by keeping the internal surfaces at a uniform temperature for the duration of the entire calibration cycle (i.e., HBB-ABB-sky-ABB-HBB-sky).

The inverse of the slope of the linear regression is called **responsivity**, $\Re(\tilde{\nu})$, which we will assume to be a real quantity for now, and it is of fundamental importance to evaluate the performance of the instrument. An insightful perspective is gained by applying the linear error propagation in $B_{\text{DS}}(\tilde{\nu})$ as a function of the uncertainty of all terms on the right-hand side of Eq. 2.10. By indicating the random error standard deviation with σ and defining for compactness $\Delta = \frac{D_{\text{sky}}(\tilde{\nu})-D_{\text{ABB}}(\tilde{\nu})}{D_{\text{HBB}}(\tilde{\nu})-D_{\text{ABB}}(\tilde{\nu})}$, we get the following equation, where dependence from $\tilde{\nu}$ is omitted for brevity:

$$\sigma(B_{\text{DS}}) = \sqrt{\Delta^2\sigma^2(B_{\text{DS,HBB}}) + (1-\Delta)^2\sigma^2(B_{\text{DS,ABB}}) + \frac{\Delta^2}{\Re^2}\sigma^2(D_{\text{HBB}}) + \left(\frac{1-\Delta}{\Re}\right)^2\sigma^2(D_{\text{ABB}}) + \frac{1}{\Re^2}\sigma^2(D_{\text{sky}})}. \tag{2.11}$$

Therefore:

- The impact of all errors on the Fourier transform of the interferogram (D_{HBB}, D_{ABB}, D_{sky}) increases for decreasing responsivity.
- Errors relative to the ABB (D_{ABB}, $B_{\text{DS,ABB}}$) always increase for decreasing Δ for the largely applicable condition $\Delta < 1$ (e.g. cold sky temperature or high ABB temperature), due to the relatively large degree of linear extrapolation (see Figure 3).
- Errors relative to the HBB ($D_{\text{DS,HBB}}$, $B_{\text{DS,HBB}}$) are zero for $\Delta = 0$ (or $D_{\text{sky}}(\tilde{\nu}) = D_{\text{ABB}}(\tilde{\nu})$, since there is virtually no extrapolation) and increase as $|\Delta|$ increases.

Takeaway 1: Higher radiometric calibration errors are to be expected in conditions where the responsivity of the instrument is low (e.g., due to dirty or damaged optics), the sky temperature is cold (e.g., most transparent bands on a clear dry day), or the ambient temperature is high.

2.3 Limitation 2: Finite Optical Path Difference

The maximum mirror displacement of a real interferometer, and therefore its maximum optical path difference, s_{max}, are limited by technological constraints. This results in a distortion of the measured spectral radiance compared to the real one. In particular, the reconstructed spectral radiance, $B_{\text{DS, real}}$, is smoothed by a sinc function convolution that widens the spectral lines and causes an actual reduction in spectral resolution that can be achieved. To show this, we can model the limited optical path of the interferogram by introducing a rectangular function into Eq. 2.7:

$$B_{\text{DS, real}}(\omega) = \int_{-\infty}^{\infty} I(s)h(s)e^{-j\omega s}d\omega, \tag{2.12}$$

where $h(s) = 1$ in the range $-s_{\max} < x < s_{\max}$ and 0 otherwise. The convolution theorem readily gives (Greenberg 1998)

$$\hat{B}_{\mathrm{DS}}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} B_{\mathrm{DS}}(\omega') H(\omega - \omega') d\omega'. \tag{2.13}$$

According to an established result, the FT of $h(s)$ is a sinc function of the form

$$H(\omega) = 2\frac{\sin(\omega s_{\max})}{\omega}. \tag{2.14}$$

We have shown how the spectral radiance derived from an interferogram is equal to the original spectral radiance convolved with a sinc function whose width increases as $s_{\max}$ decreases. An example of such effect for different values of $s_{\max}$ is shown in Figure 4. Here, a sample spectral radiance from the ASSIST (black line, top row) taken as the baseline is reprocessed with decreasing $s_{\max}$ to simulate interferometers with progressively smaller maximum optical path differences. The associated smoothed spectra are shown in red. The bottom plots provide the associated sinc smoothing function, both for the real ASSIST (black) and for the interferometers with reduced $s_{\max}$ (red).

To have a sense of the severity this "spectral leakage" in the ASSIST (black lines), we notice that the second lobes of the sinc function have an integral that is 23% of the central lobe, the third lobes 13%, the fourth lobes 1%, etc. An a posteriori correction for this effect is not possible, as it ensues from an irreversible information loss due to a fundamental limitation of the interferometer (i.e., it is mathematically impossible to de-convolute a signal). The maximum optical path difference of the ASSIST is $s_{\max} = 1.037$ cm, and the associated spectral resolution is generally adequate for carrying out thermodynamic retrievals. In fact, spectral lines of interest (see Section 3) measured at ground levels are affected by atmospheric broadening, which is generally more severe than the smoothing due to the limited optical path difference (see Section 3.1). Furthermore, some radiative models (like the one in TROPoe) include a correction that mimics the effect of the finite optical path of the interferometer.

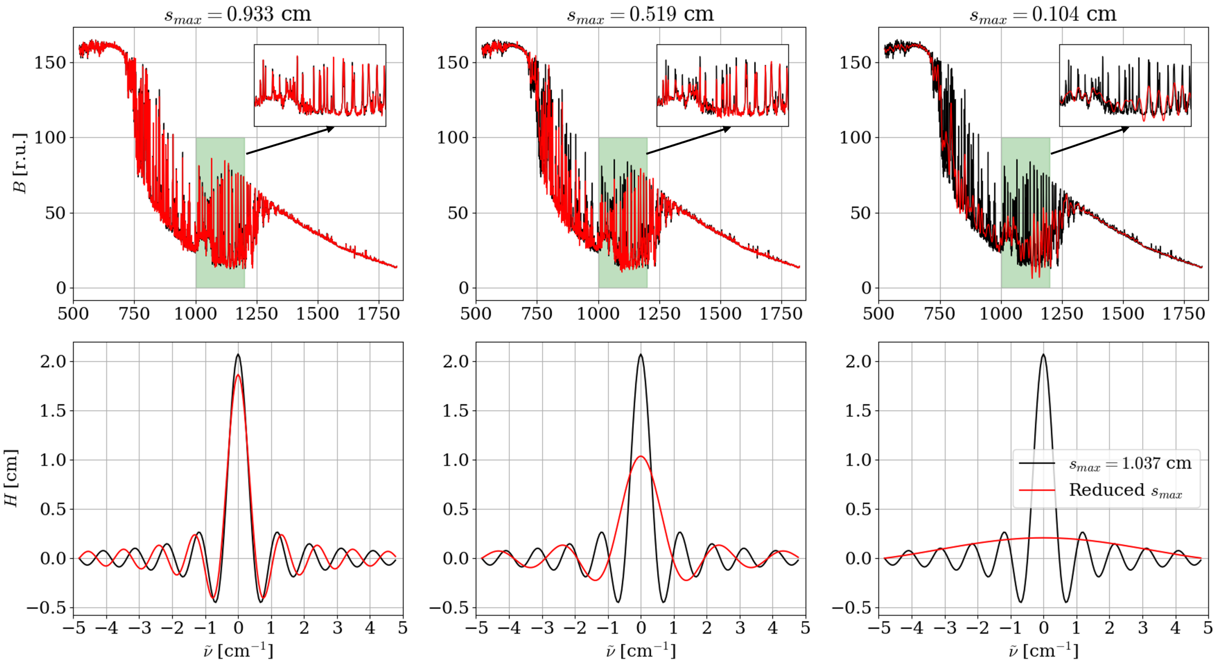

Figure 4. Effect of different maximum optical path differences on the spectral radiance. Top row: single-sided spectra; bottom row: sinc smoothing function from Eq. 2.14.

Takeaway 2: The finite maximum displacement of the moving mirror results in a smoothing of the spectral radiance that limits the spectral resolution. However, the extent of such smoothing in the ASSIST is small enough that it should not significantly impact the thermodynamic profiles.

2.4 Limitation 3: Finite Sampling Rate

A real interferometer also necessarily has a finite sampling frequency. To make sure that the interferogram is sampled at equally spaced mirror displacements, the detectors of the ASSIST are triggered by a quasi-monochromatic meteorology laser with wavenumber $\tilde{\nu}_0 = 15798.05\ \mathrm{cm}^{-1}$ that goes through the same optical path as the measured radiation. Since monochromatic radiation produces an interferogram that is purely sinusoidal, the detector samples the radiation simultaneously to the occurrence of the rising-edge zero-crossings of the interferogram generated by the laser. This process ensures a highly accurate sampling rate.

The discrete nature of the interferogram also implies that the Fourier transform in Eq. 2.7 can be performed as a discrete Fourier transform (DFT). The DFT provides the Fourier transform values for a discrete signal of length N at N equally spaced sample points (Oppenheim 1999). In particular, for even N, like in the ASSIST:

$$\begin{cases} B_{\mathrm{DS},k}(\tilde{\nu}_k) = \sum_{n=-N/2}^{N/2-1} I(s_n) e^{-j\frac{2\pi}{N}kn} \\ \tilde{\nu}_k = \frac{k}{N\Delta s} \end{cases} \qquad \forall\ k = -N/2, \ldots, N/2-1, \tag{2.15}$$

where $\Delta s = \tilde{\nu}_0^{-1}$ is the resolution in optical path difference, or twice the movement of the mirror between two sampling locations.

The first implication is that for an ASSIST with a moving mirror generating optical path differences from $-s_{\max}$ to $+s_{\max}$ and recording a double-sided interferogram at N points, the spectral resolution is equal to $\Delta\tilde{\nu} = \frac{1}{N\Delta s} = \frac{1}{2s_{\max}}$. An ASSIST has an $s_{\max} = 1.037$ cm, resulting in a discrete spectral resolution of $\Delta\tilde{\nu} = 0.482\ \mathrm{cm}^{-1}$.

A second consequence of the finite sampling of the interferometer is that the energy contained at wavenumbers larger than the Nyquist wavenumber will be mirrored on the resolved part of spectral radiance, creating aliasing. The Nyquist wavenumber for the interferometer is $(2\Delta s)^{-1} \sim 7,900\ \mathrm{cm}^{-1}$. Plank's law indicates that the energy content above the ASSIST's Nyquist wavenumber exceeds 1% for a black body temperature of 1,150°C, which is far beyond the maximum temperature detected in the atmosphere. Aliasing could occur if sunlight were to enter either directly or indirectly entering the detector. In this instance, the antialiasing filter embedded in the ASSIST (mostly active to prevent noise aliasing, see Section 2.7) would still provide protection.

Takeaway 3: The ASSIST is minimally affected by aliasing when observing atmospheric spectra.

2.5 Limitation 4: Nonlinearity of Channel A Detector

The radiometric calibration requires a linear response of each detector to apply the linear extrapolation of the blackbody spectra on the sky view (Figure 3). The InSb channel B detector is inherently linear, so Eq. 2.10 can be applied directly at long wavenumbers. However, the HgCdTe channel A detector is nonlinear enough that a direct application of the linear radiometric calibration would throw the estimated spectra off specification due to a scene-dependent responsivity.

Therefore, the ASSIST applies a nonlinearity correction to the DFT of the interferogram of channel A before the radiometric calibration. The correction uses a quadratic expression of the generic response of the detector:

$$D(\tilde{\nu}) = a_0 + \Re(\tilde{\nu}) B_{\mathrm{DS}}(\tilde{\nu}) - a_2 B_{\mathrm{DS}}^2(\tilde{\nu}) \rightarrow D_{\mathrm{linear}}(\tilde{\nu}) = D(\tilde{\nu}) + a_2 \mathrm{DFT}(I(s) + V_{\mathrm{DC}}), \tag{2.16}$$

where V_{DC} is a DC offset present in the detector that is modeled internally by the ASSIST (Knuteson et al. 2004a). V_{DC} cancels out in the radiometric calibration since it is constant but must be considered in the nonlinearity correction. Operatively, the parameter a_2, which is a single scalar, is obtained by installing a source with known spectral radiance on the sky window, such as a liquid nitrogen bath. In this way, it is possible to adjust the coefficient a_2 to minimize the difference in the responsivity between the HBB and the ABB and the difference in the responsivity between the ABB and the known cold source. Constant responsivity when staring at different scenes implies linearity of $D_{\mathrm{linear}}(\tilde{\nu})$ with respect to $B_{\mathrm{DS}}(\tilde{\nu})$. During field deployments, the validity of the nonlinearity correction (among other processes) can be checked using the so-called **third blackbody calibration**, where an isothermal auxiliary blackbody is installed on the sky aperture, thus producing a known spectral emission to be compared to the ASSIST estimate.

Takeaway 4: It is important to carry out a third blackbody calibration before and after each deployment to ensure that the nonlinearity of channel A is adequately corrected.

Takeaway 5: Uncertainty in the nonlinearity correction coefficient a_2 has a negligible impact for when the observed radiance is close to that of the ABB (e.g., in more opaque channels), but the impact increases as the scene temperature decreases (e.g., in atmospheric window channels).

2.6 Limitation 5: Finite Field of View and Laser Alignment

The fundamental equation of the interferogram (Eq. 2.6) was obtained under the assumption of a radiation beam perfectly parallel to the optical axis of the inlet and the detector. However, a real interferometer will admit energy coming from a finite solid angle to acquire information with an acceptable signal-to-noise ratio. However, this finite field of view (FFOV) causes changes in (1) the shape of the measured spectral radiance and (2) the effective sampling frequency dictated by the metrology laser.

It is easy to prove that incident radiation at an angle α from the optical axis will recombine at the detector with an optical path difference reduced by a factor of $\cos\alpha$ (Genest and Tremblay 1999). Integration over a small half-cone angle α_{max} leads to the following expression for the so-called self-apodized interferogram, i.e., the interferogram produced by an interferometer with an FFOV:

$$I(s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} B_{\text{DS}}(\omega)\frac{\sin(\omega s\alpha_{\text{max}}^2/4)}{\omega s\alpha_{\text{max}}^2/4}e^{j\omega s}d\omega. \tag{2.17}$$

The half-cone angle of ASSIST is $\alpha_{\text{max}} = 23$ mrad and is so small that the sinc function is very close to the Dirac delta for a wide set of wavenumbers. Additionally, the ASSIST applies an analytical correction based on Eq. 2.17 on the spectra, which is generally less than 1 r.u.

Another effect of the FFOV is the slight shift in the actual sampling frequency of the detector triggered by the metrology laser. Specifically, the actual sampling wavenumber is increased by a factor $\alpha_{\text{max}}^2/4$ (in radians), which is about 100–200 ppm of the nominal wavenumber for the ASSIST, and it is corrected by internal processing. Additional and more problematic shifts in the sampling frequency of the laser occur for any small misalignments of the laser with respect to the nominal optical path inside the interferometer. The specification of the ASSIST for wavenumber knowledge is better than 0.01 cm^{-1}, so even a small deviation in sampling frequency can result in an unacceptable error on $\Delta s = \tilde{\nu}_0^{-1}$ and therefore $\tilde{\nu}$ (see Eq. 2.15), with a consequent stretching/contraction of the spectral radiance on the wavenumber axis. This can be detrimental to the retrieval because it shifts important atmospheric absorption lines. The current method to correct for this error is based on the calibration of the water vapor absorption lines that are always detectable in the spectral radiance. The third blackbody calibration can also help in identifying spectral errors due to FFOV and laser misalignments.

Takeaway 6: Finite field-of-view effects have an analytical correction that is automatically applied by the ASSIST, but wavenumber shifts due to misalignments of the laser need to be checked and corrected after each relocation.

2.7 Limitation 6: Noise

So far, we have considered an instrument without noise in the acquisition chain. However, random noise is introduced in the ASSIST from different internal and external sources. An inevitable and fortuitously measurable consequence of the random noise in the interferogram is the presence of nonsymmetrical components. These components translate in the frequency domain into an imaginary part added to the real spectral radiance obtained in Eq. 2.10, which then makes it possible to perform online noise quantification (Revercomb et al. 1988). The health of the instrument can be checked by monitoring the imaginary radiance and making sure that it does not significantly increase over time.

Takeaway 7: The health status of the ASSIST should be assessed by monitoring the magnitude of the imaginary radiance. The imaginary radiance is also useful to estimate the noise level needed by TROPoe.

3 Fundamentals of Radiative Transfer Through a Participating Medium

We have reviewed the complex and interesting science that goes beyond the measurement of spectral radiance. However, in many applications, the spectral information itself is not readily usable, for example, to estimate atmospheric stability or validate atmospheric models. Through a thermodynamic retrieval (such as TROPoe), we can extract the information of the atmospheric state concealed in the spectral radiance. This chapter provides a quick overview of the most important physical mechanisms that govern the relationships between the spectral radiance observed at the ground and the thermodynamic state of the gases and water molecules in the atmosphere. The description will be mostly qualitative in that the main mathematical result is the law of radiative transfer, which is the backbone of TROPoe. Yet, even this brief overview will equip the ASSIST user with a useful skill: Understanding the overall atmospheric conditions by just looking at the raw spectra and quickly spotting nonphysical artifacts that require immediate attention.

3.1 General Concepts

An electromagnetic wave traveling through a medium undergoes a series of interactions that are strongly coupled with the molecular and physical properties of the medium itself. In the treatment of radiative heat transfer through a participating medium, it is convenient to describe the radiation as a packet of photons. The photon energy is proportional to its wavenumber and equal to $h\tilde{\nu}$, where h is Planck's constant. When a photon interacts with a particle, it can be absorbed or scattered. In the first case, the photon contributes to increase the energy of the medium. In the second case, the same photon is deflected without (elastic scattering) or with (inelastic scattering) a loss or gain of energy. The medium itself also contributes to the radiative flux by emitting photons based on its local thermodynamic state through a process that is the mirror version of the absorption. An illustration of these four interaction modes between medium and photons (absorption, emission, elastic scattering, and inelastic scattering) is shown in Figure 5 .

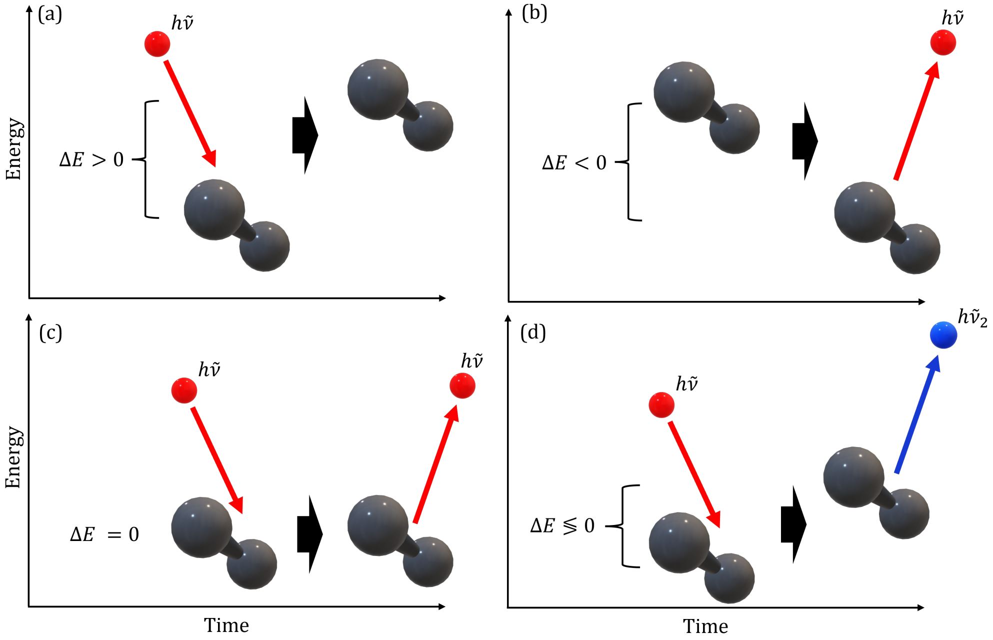

Figure 5. Schematic representation of the type of photon-medium interactions: (a) absorption; (b) emission; (c) elastic scattering; (d) inelastic scattering.

Scattering can happen in the form of reflection, refraction, or diffraction and be subject to different physical mechanisms, mostly based on the size of the particle compared to the wavelength of the incident radiation. The treatment of scattering is significantly more complex than absorption and emission. Fortunately, it represents a negligible fraction of the infrared energy of interest for thermodynamic profiling, with some exceptions in the presence of clouds (Turner and Löhnert 2014). Therefore, the remainder of the report will focus on **absorption** and **emission** only.

Absorption and emission in a medium occur at specific energy levels associated with state transitions. A photon can be absorbed only if the difference in two of the energy states of the medium, $\Delta E > 0$, is equal to its energy $h\tilde{\nu}$ or, equivalently, the photon has a wavenumber such that $\tilde{\nu} = \Delta E/h$. In reverse, a negative energy transition results in the emission of a photon with energy $h\tilde{\nu} = -\Delta E$ and wavenumber $\tilde{\nu} = -\Delta E/h$. Therefore, the linear relationship between the energy and the wavenumber of a photon implies that the absorptive and emissive properties of a medium are inherently wavenumber-dependent. In other words, the spectral properties of the radiation are correlated to the energy state of the medium. The retrieval of thermodynamic profiles in the atmosphere from the spectral radiance takes advantage of this fundamental physical property.

There is a qualitative difference between the spectral absorption and emission of non-ionized and ionized media. A non-ionized medium includes molecules and atoms that have energy states associated with specific rotational and vibrational modes and electronic states. In this case, energy transitions are called bound-bound and occur as jumps between discrete energy levels. This means that only photons with specific energy levels (or, equivalently, wavenumbers) can be absorbed or emitted. This explains why spectrum of the radiation traveling through a non-ionized medium is characterized by narrow absorption/emission bands at specific frequencies.

In contrast, a medium that contains ions and free electrons is characterized by a continuum absorption/emission spectrum because these types of particles can have any energy above a threshold called ionization energy. The contribution of ions in the atmospheric processes is negligible, and we will therefore focus only on the non-ionized medium in the following discussion.

In reality, however, as a result of so-called **line broadening**, even a non-ionized medium adsorbs/emits over a finite range of wavenumbers around ideal spectral lines predicted by quantum theory, $\Delta E/h$. The shape of the broadened spectral lines is referred to as **lineshape**. For atmospheric applications, broadening can occur mainly because of two mechanisms:

- Doppler effect due to the relative motion of the absorbing/emitting particle induced by thermal agitations: In this case, the observed lineshape is Gaussian, and its amplitude is proportional to $T^{1/2}$ (where T is temperature); thus, it is more relevant at high temperatures where molecules move more vigorously.
- Collision between particles of the same species (self-broadening) or different species (foreign broadening): The observed lineshape is an approximate Lorentzian function with slowly decaying tails or wings and amplitude proportional to $P/T^{1/2}$ (where P is pressure), becoming important at high pressures and low temperatures.

Broadening is important in the context of thermodynamic profiling because it eases the spectral resolution requirements of the spectrometer and impacts the accuracy of physical radiative transfer models (see Section 4.2.3).

Takeaway 8: The atmosphere can be treated as an absorptive-emissive, non-ionized medium in the infrared region of interest for thermodynamic profiling. Absorption/emission occurs around specific spectral lines that are associated with the discrete energy states of the different atmospheric layers.

3.2 The Law of Radiative Transfer

Energy conservation along the travel path for radiation in nonscattering media takes the form of the so-called **law of radiative transfer** and plays a crucial role in thermodynamic profiling. It basically represents the balance between the absorbed and emitted spectral radiance along the path traveled by the radiation. The path is defined by the direction in three-dimensional space, α, and the line-of-sight distance from an arbitrary origin, ξ.

The absorption of energy is described by the following differential equation:

$$\left.\frac{\partial B(\tilde{\nu},\xi,\alpha)}{\partial \xi}\right|_{\text{absorption}} = -a_{\tilde{\nu}}(\xi,\alpha)B(\tilde{\nu},\xi,\alpha), \tag{3.1}$$

where $B(\tilde{\nu},\xi,\alpha)$ is the spectral radiance at wavenumber $\tilde{\nu}$, along-beam location ξ, and in direction α. The absorption coefficient, $a_{\tilde{\nu}}(\xi,\alpha)$, represents the fraction of energy absorbed in a unit thickness layer at the specific wavenumber. The derivative includes only radiation attenuation as a result of absorption.

If each parcel of the medium is in local thermodynamic equilibrium (which largely applies to the atmospheric case when the air pressure is significantly greater than zero), the emission is only a function of the local thermodynamic state. Based on fundamental thermodynamic considerations (Siegel 1971), it can be shown that the emission spectrum has a fairly simple expression: It is equal to the blackbody radiation at the local temperature, $B_0(\tilde{\nu}, T(\xi, \alpha))$, times the absorption coefficient, and it is conveniently isotropic. So, the local increase in radiation due to emission along direction α is:

$$\left.\frac{\partial B(\tilde{\nu}, \xi, \alpha)}{\partial \xi}\right|_{\text{emission}} = +a_{\tilde{\nu}}(\xi, \alpha) B_0(\tilde{\nu}, T(\xi, \alpha)). \tag{3.2}$$

These two basic equations allow us to derive the law of radiative transfer from the total variation of energy due to absorption and emission along the travel path. The law of radiative transfer of the electromagnetic radiation in a nonscattering participating medium can be recast as:

$$B(\tilde{\nu}, \kappa_{\tilde{\nu},\alpha}) = B(\tilde{\nu}, 0, \alpha) e^{-\kappa_{\tilde{\nu},\alpha}} + \int_0^{\kappa_{\tilde{\nu},\alpha}} B_0(\tilde{\nu}, T(\kappa'_{\tilde{\nu},\alpha})) e^{-(\kappa_{\tilde{\nu},\alpha} - \kappa'_{\tilde{\nu},\alpha})} d\kappa'_{\tilde{\nu},\alpha}, \tag{3.3}$$

where α is the direction and $\kappa_{\tilde{\nu},\alpha}$ is a special spatial coordinate called **optical thickness** defined as:

$$\kappa_{\tilde{\nu},\alpha} = \int_0^{\xi} a_{\tilde{\nu}}(\xi', \alpha) d\xi'. \tag{3.4}$$

Equation 3.3 shows the spectral radiation at a given point in space coming from a specific direction as the sum of two terms:

- The radiation at the boundary of the domain $B(\tilde{\nu}, 0, \alpha)$ attenuated by $e^{-\kappa_{\tilde{\nu},\alpha}}$ as a consequence of absorption.
- The sum of the emission of the medium along the specific direction, which is proportional to all the blackbody spectral energy encountered along the path, $B_0(\tilde{\nu}, \kappa_{\tilde{\nu},\alpha})$, convolved with an exponential weighting function which represents the loss due to absorption along the path.

For illustrative purposes, we can derive the law of radiative transfer for an ideal medium with homogeneous absorption, the optical thickness is simply $\kappa_{\tilde{\nu}} = a_{\tilde{\nu}}\xi$. Equation 3.3 then becomes

$$B(\tilde{\nu}, \xi, \alpha) = B(\tilde{\nu}, 0, \alpha) e^{-a_{\tilde{\nu}}\xi} + a_{\tilde{\nu}} \int_0^{\xi} B_0(\tilde{\nu}, T(\xi', \alpha)) e^{-a_{\tilde{\nu}}(\xi - \xi')} d\xi'. \tag{3.5}$$

To get a sense of how thermodynamic profiling from a ground-based instrument works, we can apply Eq. 3.5 to an observer located at the Earth's surface and looking up along the zenith direction. In this case, we can neglect the radiation coming from outer space, so that the law of radiative transfer reads

$$B(\tilde{\nu}) = a_{\tilde{\nu}} \int_0^{\infty} B_0(\tilde{\nu}, T(z)) e^{-a_{\tilde{\nu}} z} dz, \tag{3.6}$$

where the spatial coordinates have been dropped for simplicity. The spectral radiance at the ground at a specific wavenumber is the integral of the blackbody radiance emitted by the columns of gases overhead multiplied by the transmission of the atmosphere from the level of emission to the instrument, which takes an exponential form. This vertical product is called the **weighting function** in the context of atmospheric sounding. Moreover, the stronger the absorption, the faster the decay of the weighting function. Equation 3.6 is qualitatively shown in Figure 6 for low and high absorption coefficients. The centroid of the weighting function, located at an altitude of $z = a_{\tilde{\nu}}^{-1}$, is also indicated in red.

Clearly, a less absorptive or "transparent" medium weights more contributions from higher layers compared to a more absorptive or "opaque" medium, which is mainly affected by emission close to the ground. We can translate this conclusion into a useful property for the spectra measured at the ground by an ASSIST. Spectral regions with large $a_{\tilde{\nu}}$ will appear more opaque and allow only radiation close to the surface to reach the instrument. Spectral regions with small $a_{\tilde{\nu}}$ will be more transparent and allow radiation from higher altitudes to reach the instrument.

In this context, a convenient variable (that is also provided by the ASSIST at specific wavenumbers) is the brightness temperature, or the temperature of a blackbody emitting with a certain energy at the specific wavenumber, which reads (in °C):

$$T_b(\tilde{\nu}, B_0(\tilde{\nu})) = \frac{100 \cdot hc\tilde{\nu}}{k_B} \left[\log\left(\frac{2 \cdot 10^{11} \cdot hc^2 \tilde{\nu}^3}{B_0(\tilde{\nu})} + 1 \right) \right]^{-1} - 273.15. \tag{3.7}$$

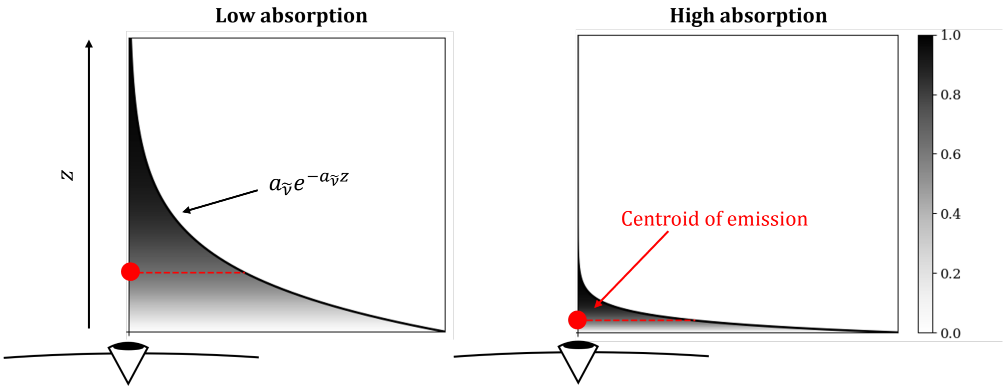

Figure 6. Qualitative behavior of weighting functions of the law of radiative transfer for a ground-based, zenith-looking observer and a homogeneous absorption coefficient from Eq. 3.6. The area under the function is colored by its cumulative integral.

In light of this example, a very rudimentary thermodynamic profiling method could then assume that the brightness temperature derived from the spectral value at a certain wavenumber is a proxy for the temperature at the centroid of the emission layer, i.e., $z = a_{\tilde{\nu}}^{-1}$. Therefore, the temperature profile across different altitudes could be estimated from observations over a set of wavenumbers with large variations of $a_{\tilde{\nu}}$, for instance, close to a steep absorption line. A real-world spectrum is clearly more complex with contributions from the emission from multiple layers that overlap on the same spectral bands, inhomogeneous absorption coefficients, and boundary effects in very transparent bands. This calls for a more sophisticated retrieval method to be used to estimate the atmospheric state, i.e., TROPoe.

Takeaway 9: The spectral absorption and emission of infrared radiance depend strongly on the thermodynamic state of the medium (i.e., the atmosphere), which is the basis of TROPoe and thermodynamic profiling in general.

3.3 Typical Atmospheric Spectra

A simple visual inspection of the ASSIST spectra can provide important information on the atmospheric state and physics. Figure 7 shows three spectra collected at AWAKEN site B (Letizia 2023) in August 2023 and the associated brightness temperature. The ASSIST was collocated with a meteorological station (Goldberg 2023) that provided surface temperature (T_s) and relative humidity (RH_s) measurements at 2 m above the ground. A laser ceilometer (Hamilton 2023) was also installed at site A1 (about 20 km northeast) and measured the cloud base height. The first and second spectra (Figure 7a,b) were collected on 2023-08-24 during clear-sky conditions at nighttime (06:00 UTC, 01:00 local time) and daytime (15:00 UTC, 10:00 local time), respectively. They are qualitatively similar. We can easily identify wavenumbers in the regions $500 < \tilde{\nu} < 750$ cm^{-1}, $1,250 < \tilde{\nu} < 2,000$ cm^{-1}, and $2,250 < \tilde{\nu} < 2,380$ cm^{-1} where the radiance closely follows Planck's law evaluated at surface temperature (black line in the figure). These opaque bands represent the spectral regions where the atmosphere is generally optically very thick or opaque, so that the instrument "sees" mostly the warm emission very close to the ground. In fact, the brightness temperature in these bands is close to the surface value. Conversely, the regions $700 < \tilde{\nu} < 1,250$ cm^{-1}, $2,000 < \tilde{\nu} < 2,250$ cm^{-1}, $2,380 < \tilde{\nu} < 3,000$ cm^{-1} show significantly lower energy, meaning that these are more transparent bands if the sky is clear. In this region, the ASSIST is "seeing" radiation coming from far above the Earth surface. In this region of the atmosphere temperatures are generally lower than the surface values, so the spectral energy is below Planck's function at ambient temperature and the brightness temperature also drops to -40°C or below. Interestingly, some spectral region show brightness temperatures above the ambient blackbody curve. This is due to the presence of very opaque gases inside the instrument that can obscure the view of the blackbodies and invalidate the radiometric calibration. Finally, the region close to the transition between Channel A and B (not used in TROPoe), shows significant spectral noise.

The third spectrum (Figure 7c) was collected during an overcast event on 2023-08-27 17:15 UTC (12:15 local time). Here, the spectrum shows a significant difference from those recorded during clear-sky conditions, especially in the transparent band. The brightness temperature stays close to or even above the surface value. This higher brightness temperature is because the clouds block the infrared radiation from the higher altitudes and emit at temperatures closer

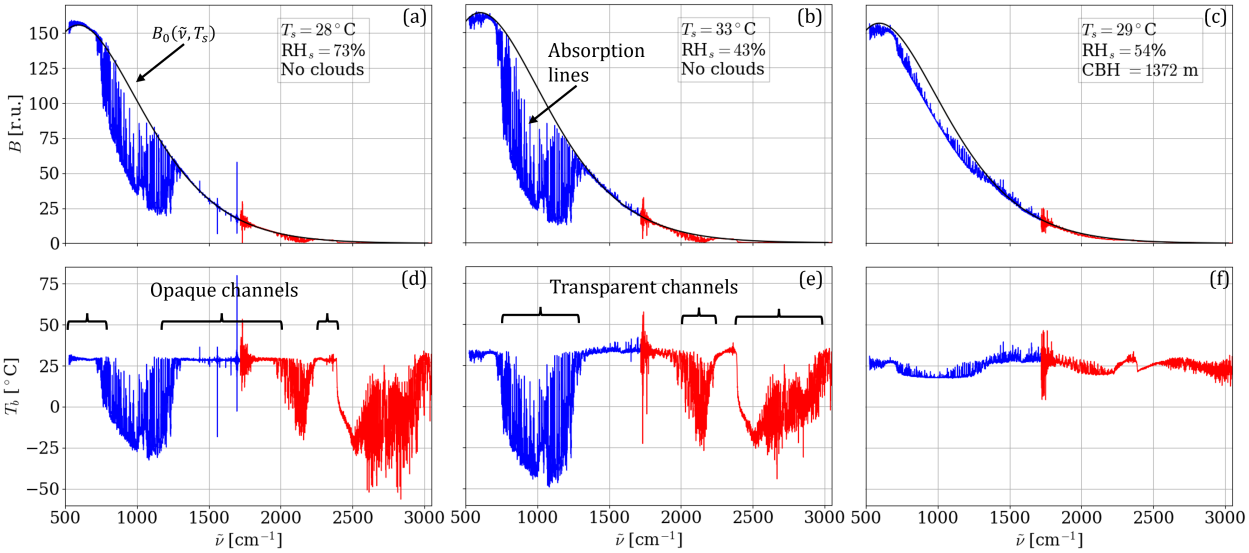

Figure 7. Three sample spectra (top) and brightness temperatures (T_b, bottom) from the ASSIST installed at AWAKEN site B: (a,d) nighttime, clear sky; (b,e) daytime, clear sky; (c,f) daytime, cloudy. The blue and red lines indicate data from channel A and channel B, respectively. The black line is the Planck function evaluated at T_s.

to the surface values. In general, the spectral information from above cannot penetrate dense cloud layers; therefore, any physical retrieval above cloud base height is generally affected by large uncertainty (see Section 4.2.4).

Takeaway 10: Different regions of the spectral radiance carry information about different layers of the atmosphere based on the associated optical thickness. Clouds generally block most of the information above the cloud layer, and emit significant energy to the surface.

Zooming into the spectra, it is possible to identify absorption lines from the different gases that absorb and emit at specific wavenumbers based on their molecular structure. The most important species with active absorption/emission within the spectral band detected by the ASSIST are H_2O, CO_2, CH_4, N_2O, and O_3. These known spectral dependencies are leveraged to carry out the optimal retrieval of the temperature and humidity profiles, as described next.

4 Retrieval of Thermodynamic Profiles

We have discussed so far how the spectral radiance is affected by the thermodynamic state of the atmosphere, in the form of temperature and spectral absorption coefficient, and how it is measured at the ground level. The final missing piece in understanding how thermodynamic profiling works is to describe the way a the retrieval algorithm (for us, TROPoe) inverts the **forward model** represented by the law of radiative transfer to estimate the thermodynamic profiles. A quick inspection of the law of radiative transfer (Eq. 3.3) reveals the complexity of performing its inversion. In fact, when the boundary radiation is neglected, it takes the form of a convolution of blackbody spectra with an exponential function of the optical thickness. The inverse of a convolutions, just like all purely integral equations, is inherently ill-posed, meaning that it admits several solutions. In other words, the **state** that we want to estimate (temperature, moisture, cloud properties, etc.) has more pieces of information than the **observations** (spectral radiance at the ground). Said differently, there are frequently more than one atmospheric state conditions that could have yielded the observed infrared spectrum. This is equivalent to trying to guess the shape of a three-dimensional object by looking at its shadow.

Moreover, the **nonlinearity** of the blackbody radiation and absorption coefficient in the thermodynamic state variables and the instrumental **noise** both contribute to make the inversion of Eq. 3.3 even more complicated.

TROPoe tackles this challenging task from a both a physical and statistical point of view by identifying the solution that has the highest likelihood given the set of available observations and based on the physics of the radiation. To do so in a numerically stable way, it also uses information from the site climatology, auxiliary instruments, and an estimate of the noise level in the observations. In addition to providing a statistically sound solution, TROPoe is capable of producing and estimating the uncertainty of the retrieval, which is invaluable for data analysis and assimilation.

4.1 Fundamentals of Optimal Estimation

As previously mentioned, the estimation of thermodynamic profiles based on observed spectral radiance at the ground level requires the solution of an ill-posed inverse problem. A statistically robust approach to solving an inverse problem is OE (Rodgers 2000). Formally, given an unknown state vector that we want to estimate, x, a **prior** probability density function (p.d.f.) describing the overall variability of x, and a set of observations, y, affected by measurement **noise** with known p.d.f., OE theory provides a way to estimate the most likely x compatible with both prior statistics and current y. This optimal solution is indicated by $\hat{x}$. A big advantage of the OE is that it also provides the uncertainty of $\hat{x}$ due to measurement noise and missing information in the forward model in the form of the so-called **posterior** p.d.f.

OE methods are often based on Bayes's theorem, which states that the posterior p.d.f. (formalized as the conditional p.d.f. of x given y) can be obtained as the product of the p.d.f. of the measurement noise, the prior p.d.f. (based on historical data, for example) and a p.d.f. of the observations which act as a normalization factor. In formula:

$$\overbrace{p(x|y)}^{\text{posterior}} = \left(\overbrace{p(y|x)}^{\text{noise}} \overbrace{p(x)}^{\text{prior}} \overbrace{p(y)^{-1}}^{\text{normalization}} \right). \tag{4.1}$$

In our case, x contains the thermodynamic properties of the atmosphere at selected heights, y includes the spectral radiance and auxiliary observations, $p(y|x)$ contains the forward model for atmospheric absorption and emission, plus instrumental noise of the ASSIST, and $p(x)$ represents the climatology. In this framework, the problem of thermodynamic profiling is multidimensional and nonlinear. However, we first consider the much simpler scalar, linear problem, which already provides an insightful look into OE theory. Then, we move to more complex scenarios and finally to the general case.

4.1.1 Linear, One-Dimensional, Gaussian Problem

We want to optimally estimate a scalar variable x that is normally distributed with a prior mean x_a and a standard deviation σ_a based on an observation y made with an instrument with a linear calibration curve and Gaussian noise ε with standard deviation σ_ε. The forward model is simply

$$y = kx + \varepsilon, \tag{4.2}$$

where the slope k can be seen as the sensitivity of the instrument. Bayes' (Eq. 4.1) theorem in this case reads

$$\overbrace{p(x|y)}^{\text{posterior}} = \overbrace{\left(\frac{1}{\sqrt{2\pi}\sigma_\varepsilon} e^{-\frac{(y-kx)^2}{2\sigma_\varepsilon^2}}\right)}^{\text{noise}} \overbrace{\left(\frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{(x-x_a)^2}{2\sigma_a^2}}\right)}^{\text{prior}} \overbrace{p(y)^{-1}}^{\text{normalization}}, \tag{4.3}$$

where the $p(y)$ is deliberately left unexpanded for compactness. The best solution in an OE sense is the $\hat{x}$ corresponding to the maximum of $p(x|y)$, i.e., the most probable state conditional on the observation. This can be found by maximizing the left-hand side of Eq. 4.3, which is equivalent to

$$\frac{\partial}{\partial x}\left[\frac{(y-kx)^2}{2\sigma_\varepsilon^2} + \frac{(x-x_a)^2}{2\sigma_a^2}\right] = 0. \tag{4.4}$$

The solution to this simple optimization problem is customarily recast as

$$\hat{x} = x_a + \left(\frac{k^2}{\sigma_\varepsilon^2} + \frac{1}{\sigma_a^2}\right)^{-1} \frac{k}{\sigma_\varepsilon^2}(y - kx_a), \tag{4.5}$$

or, equivalently

$$\hat{x} = \frac{y\, k\sigma_a^2 + x_a\, \sigma_\varepsilon^2}{k^2\sigma_a^2 + \sigma_\varepsilon^2}. \tag{4.6}$$

Therefore, the optimal state estimate is a weighted mean of the observation and the prior mean based on noise, prior variability, and k. Figure 8 shows a graphical representation of the process. If $k\sigma_a \gg \sigma_\varepsilon$ (Figure 8a), i.e., the noise is significantly smaller than the variability in the prior times the sensitivity of the instrument, most of the information comes from the observation, and we approach the naive solution $x \sim y/k$. If $k\sigma_a \sim \sigma_\varepsilon$ (Figure 8b), the OE mean uses the prior to partially constrain the observation; if $k\sigma_a < \sigma_\varepsilon$ (Figure 8c), the observation is too noisy or the instrument sensitivity is too low, and the solution is significantly constrained by the prior. In the worst-case scenario of $k\sigma_a \ll \sigma_\varepsilon$, we get the uninformative solution $x = x_a$. The three cases qualitatively represent the archetypes of ideal, realistic, and bad retrieval, respectively. The purple OE mean line is observed to lie at the loci of the maxima of the joint p.d.f. $p(x,y)$ (shown as a heatmap), showing that OE represents the most likely state of x for a given y.

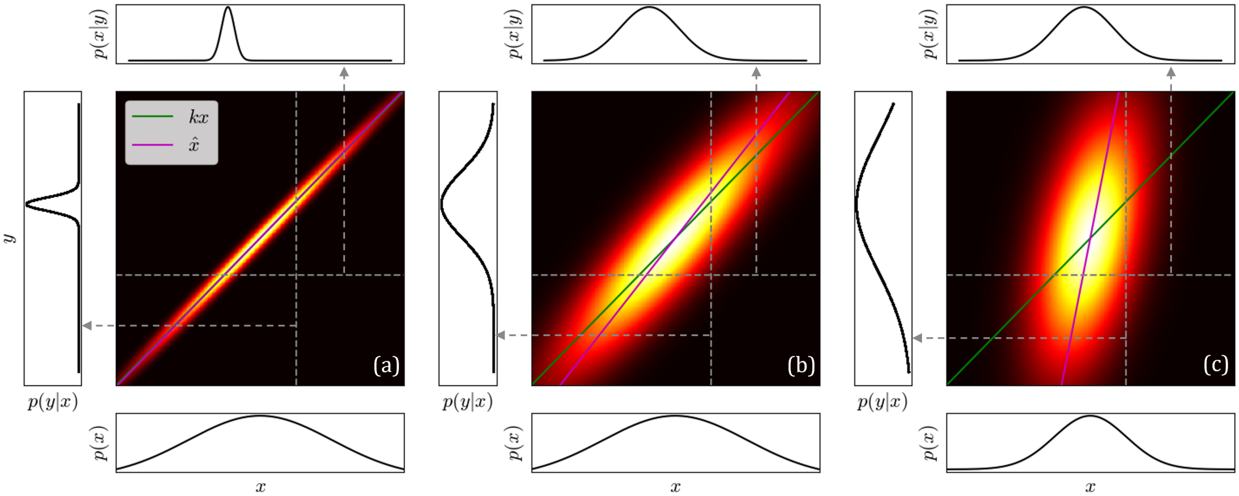

Figure 8. Graphical depiction of a scalar linear OE for $k = 1$: (a) ideal case, where $\sigma_a = 10$, $\sigma_\varepsilon = 1$; (b) realistic case, where $\sigma_a = 10$, $\sigma_\varepsilon = 5$, (c) bad case, where $\sigma_a = 5$, $\sigma_\varepsilon = 10$. The heatmap indicates the joint p.d.f. $p(x,y)$.

Looking again at Equation 4.3 and recalling that the product of two Gaussian functions of x is itself a Gaussian, we can readily derive the standard deviation of the posterior p.d.f. by simply equating terms with x^2. The result is

$$\hat{\sigma} = \left(\frac{k^2}{\sigma_\varepsilon^2} + \frac{1}{\sigma_a^2}\right)^{-\frac{1}{2}}. \tag{4.7}$$

The $\hat{\sigma}$ represents the overall uncertainty of $\hat{x}$, but can be further broken down. Equations 4.5 and 4.2 can be combined to obtain the total error of the retrieval, ε_t, as

$$\overbrace{\hat{x}-x}^{\varepsilon_t} = \overbrace{(a-1)(x-x_a)}^{\varepsilon_s} + \overbrace{g\varepsilon}^{\varepsilon_n}, \tag{4.8}$$

where the so-called **gain**, g is

$$g = \left(\frac{k^2}{\sigma_\varepsilon^2} + \frac{1}{\sigma_a^2}\right)^{-1} \frac{k}{\sigma_\varepsilon^2}, \tag{4.9}$$

while the **averaging kernel** (or A kernel) is defined as $a = gk$. Interestingly, the gain is also the partial derivative of the solution with respect to the observation (i.e., $g = \frac{\partial \hat{x}}{\partial y}$), or the inverse of the slope of the purple "optimal" line in Figure 8. The averaging kernel is also the partial derivative of the solution with respect to the real (unknown) value of the state vector (i.e., $a = \frac{\partial \hat{x}}{\partial x}$).

The total error is then the sum of the so-called **smoothing** error, ε_s, and the error due to noise, ε_n (which is not the same as ε). A graphical representation of how these errors contribute to the overall uncertainty is given in Figure 9. The smoothing error is due to the use of prior information that is needed to mitigate the effect of measurement uncertainty and constrain the ill-posed problem to a realistic solution. This results in the solution being attracted to the prior mean proportionally to $x - x_a$, meaning that, for a given system, large smoothing errors are expected for conditions that significantly depart from the prior mean, i.e., extreme events. The word "smoothing" associated with this term will become clearer when considering multidimensional retrievals (Section 4.1.3). The error due to noise is instead a direct consequence of the measurement error propagated into the retrieval solution.

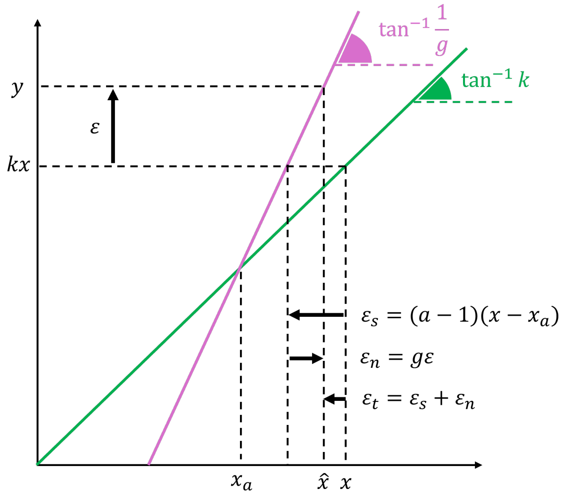

Figure 9. Graphical representation of the total error of the retrieval as the sum of smoothing and noise.

The retrieval uncertainty, $\hat{\sigma}$, includes contributions from smoothing and noise, as it can be easily proven:

$$\sigma^2(\varepsilon_t) = \sigma^2(\varepsilon_s) + \sigma^2(\varepsilon_n) = (a-1)^2\sigma_a^2 + g^2\sigma_\varepsilon^2 = \hat{\sigma}^2. \tag{4.10}$$

However, there are other sources of uncertainty that are not included in the OE output. The modeling error occurs as the result of the imperfect representation of reality by the forward model (in our simplified case, Eq. 4.2). Furthermore, the instrument could have persistent biases that are not well represented by random noise. Additional errors may arise from an imperfect estimation of the prior and/or noise statistics. These effects are much less amenable to general treatment and are harder to predict for real applications, as will be extensively discussed in Section 4.2.4.

Takeaway 11: The solution of the retrieval is the most likely state given the prior and the current observations. The posterior p.d.f. includes information on the uncertainty of the retrieval due to smoothing and measurement noise.

4.1.2 Nonlinear, One-Dimensional, Gaussian Problem

We now add an additional layer of complexity by considering a nonlinear (but still one-dimensional) forward problem defined by:

$$y = F(x) + \varepsilon, \tag{4.11}$$

where again x and ε are both Gaussian. Bayes' theorem in the form of Eq. 4.4 can still be applied, but it looks slightly different from Eq. 4.3:

$$p(x|y) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} e^{-\frac{(y-F(x))^2}{2\sigma_\varepsilon^2}} \frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{(x-x_a)^2}{2\sigma_a^2}} p(y)^{-1}. \tag{4.12}$$

This form is not amenable to a general solution for the maximum $\hat{x}$ as its linear version. However, if the nonlinearity within the most likely region of the sample space of $p(x|y)$ (that is, few $\hat{\sigma}$ from $\hat{x}$) is small, the forward model could be approximated as $F(x) \sim F(x_l) + k(x_l)(x_l - x)$, provided that the linearization point x_l lies within the proximity of $\hat{x}$. This type of problem is called **moderately nonlinear**. The OE equation is then

$$\frac{\partial}{\partial x}\left[\frac{(y - F(x_l) - k(x_l)(x_l - x))^2}{2\sigma_\varepsilon^2} + \frac{(x - x_a)^2}{2\sigma_a^2}\right] = 0. \tag{4.13}$$

However, since $\hat{x}$ is indeed the solution of the problem, x_l cannot be chosen a priori but can be approximated with progressively good accuracy in an iterative fashion. The solution scheme to Eq. 4.13 is then

$$x^{(n+1)} = x_a + \left. \frac{k^{(n)2}}{\sigma_\varepsilon^2} + \frac{1}{\sigma_a^2}\right)^{-1} \frac{k^{(n)}}{\sigma_\varepsilon^2}\left[y - F(x^{(n)}) + k^{(n)}(x^{(n)} - x_a)\right] \Bigg(\tag{4.14}$$

where the superscript (n) is the iteration counter. This solution scheme, valid for a moderately nonlinear problem, is referred to as the **Gauss-Newton** method. It is very similar to the one implemented in TROPoe, which in fact assumes that the forward radiative transfer model is only moderately nonlinear (Maahn et al. 2020). The iterative solution algorithm is responsible for most of its computational cost, as the forward model is performing complex spectral radiative transfer calculation.

So far, we have seen the computational advantages of dealing with a moderately linear problem. Another important conclusion can be made in terms of the retrieval uncertainty. In fact, when $F(x)$ is approximated through a linear function, Eq. 4.12 is not different from that of a linear problem. We then infer that the posterior p.d.f., $p(x|y)$, is also Gaussian with a standard deviation of

$$\hat{\sigma} = \left. \frac{\hat{k}^2}{\sigma_\varepsilon^2} + \frac{1}{\sigma_a^2}\right)^{-\frac{1}{2}}, \tag{4.15}$$

where the ˆ symbol in this context indicates the value at convergence. Additionally, the results regarding the smoothing error and the error due to noise seen for the linear problem apply. Simply, in this case, the gain and the A kernel will correspond to their values at convergence, or, using our nomenclature, $\hat{g}$ and $\hat{a}$, respectively.

Takeaway 12: Problems having a forward model with small nonlinearity within the most likely region of the posterior are called "moderately nonlinear," and the uncertainty of the solution can be estimated as in a strictly linear case.

To provide an example of how a nonlinear problem is solved, we apply the Gauss-Newton method to invert a forward parabolic problem ($y = kx + bx^2$) with different quadratic terms, b. Figure 10 shows the resulting optimal solutions $\hat{x}$ for different values of y and is superimposed on the full two-dimensional p.d.f., $p(x,y)$, obtained using a Monte Carlo method with 10^7 samples. In particular, Figure 10a refers to a linear retrieval, showing that $\hat{x}$ perfectly tracks the maxima of $p(x,y)$ (the heatmap). Figure 10b is what could be called a moderately nonlinear problem, where realistic solutions for $\hat{x}$ are achieved in the whole domain, and the posterior p.d.f., $p(x|y)$, assumed as Gaussian with mean $\hat{x}$ and standard deviation $\hat{\sigma}$ (top), is close to the real p.d.f. approximated through Monte Carlo sampling. Figure 10c, which corresponds to a higher quadratic term, shows a well-behaved solution in only part of the domain of y and a significant departure of the posterior from Gaussianity. It is therefore a grossly nonlinear problem for which the Gauss-Newton algorithm and the linear uncertainty theory may fail.

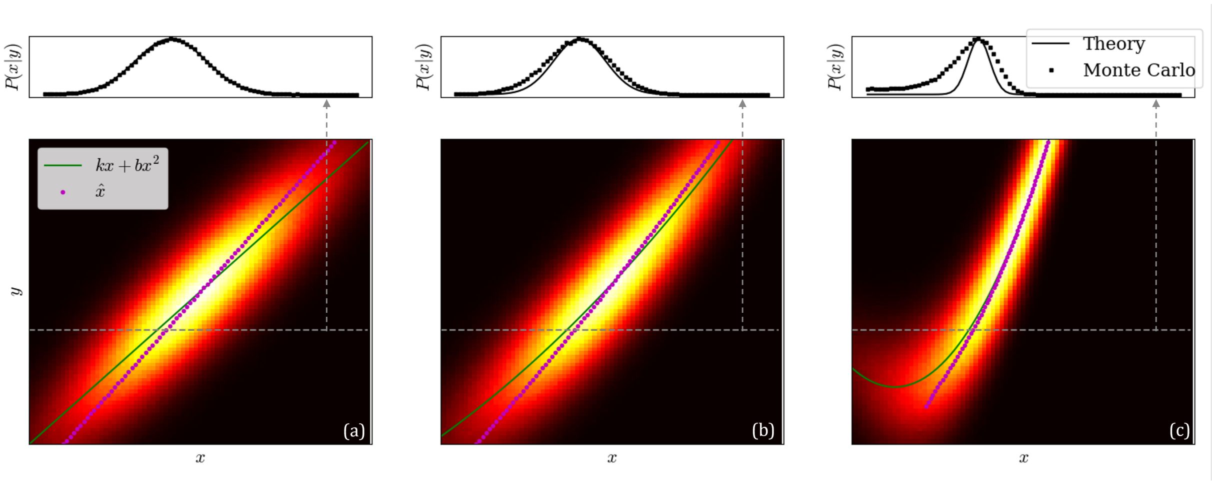

Figure 10. Graphical depiction of a scalar nonlinear OE for $k = 1$, $\sigma_a = 10$, $\sigma_\varepsilon = 5$: **(a)** $b = 0$; **(b)** $b = 0.01$; **(c)** $b = 0.1$.

4.1.3 Linear, Multidimensional, Gaussian Problem

The multidimensional linear problem described here is a more realistic approximation of the real TROPoe setup. The main difference between the one-dimensional problems seen before and the one presented in this section is that now the scalar and linear forward model is replaced by a linear system. From a purely mathematical standpoint, extending the scalar OE to multiple dimensions implies that x, x_a, y, and ε now represent vectors of the state (e.g., temperature profiles), prior mean profile, observations (e.g., radiance at different wavenumbers) and noise, respectively. The forward model then is

$$y = Kx + \varepsilon, \tag{4.16}$$

where K is an $M \times N$ matrix mapping the N states evaluated at specific heights or grid points into M observations. By adopting the multidimensional version of the derivation in Section 4.1.1, it can be proven (Rodgers 2000) that the optimal solution in this case is

$$\hat{x} = x_a + \left(K^T S_\varepsilon^{-1} K + S_a^{-1}\right)^{-1} K^T S_\varepsilon^{-1} (y - Kx_a), \tag{4.17}$$

where S_ε is the covariance matrix of the noise and S_a is the prior covariance matrix. Equation 4.17 shares several similarities with its one-dimensional counterpart, Eq. 4.5, with scalars replaced by vectors, k by K, and variances by covariance matrices. Analogously, the posterior covariance, which is the same as the covariance of the total retrieval error, $\hat{x} - x$, is

$$\hat{S} = \left(K^T S_\varepsilon^{-1} K + S_a^{-1}\right)^{-1}. \tag{4.18}$$

The gain is now a matrix that quantifies the sensitivity of the solution to the observation (and the noise):

$$G = \frac{\partial \hat{x}}{\partial y} = \left(K^T S_\varepsilon^{-1} K + S_a^{-1}\right)^{-1} K^T S_\varepsilon^{-1}. \tag{4.19}$$

The covariance of the error due to noise is then

$$S_n = G S_\varepsilon G^T. \tag{4.20}$$

The averaging kernel, $A = GK$, is also a matrix equal to the sensitivity of the solution to the true state, $\frac{\partial \hat{x}}{\partial x}$. The matrix A contains a lot of useful information, in particular:

- It can be used to calculate the smoothing error as

$$S_s = (A - I_n) S_a (A - I_n)^T, \tag{4.21}$$

  where I_N is the identity matrix of rank N, which implies that $A = I_N$ would represent an ideal case with no smoothing.

- The trace of A is the number of degrees of freedom of the signal (DFS), namely, the number of independent pieces of information of the state available in the observations at each grid point; for an ideal retrieval, $A = I_N$, so DFS $= N$, but in general DFS $< N$.
- The i-th rows of A are the linear coefficients indicating how the whole state vector affects the i-th element of the solution. In general, the diagonal elements are less than 1, indicating that the solution is partly attracted by the prior, as seen in the one-dimensional case. Also, the nondiagonal elements of A are generally not 0, indicating some contribution to the solution at i-th level from elements of the state and prior at $j \neq i$. Both effects result in the smoothing error.
- The full-width half-maximum of the rows of A provides a metric for the vertical resolution of the solution.

Takeaway 13: The severity of smoothing is quantified by the departure of the A kernel matrix from the identity, and it comes from two effects: (1) the use of prior information at the same level ($A_{ii} < 1$) and (2) contribution from different layers ($A_{ij} \neq 0$).

To illustrate how a multidimensional retrieval works, we use the simplified framework defined in Section 3.2 for a ground-based spectrometer observing downwelling spectral radiance radiation along a vertical line of sight (e.g., an ASSIST). By assuming a vertically homogeneous absorption coefficient, $a_{\tilde{\nu}}$, we can use Eq. 3.6 to build the forward model. In order to build a linear problem, we select $N = 80$ grid points equally spaced by $\Delta z = 1$ km to discretize the integral. We also focus on $M = 8$ spectral bands of the spectrometer in a small wavenumber region where large variations of $a_{\tilde{\nu}}$ occur, but we can neglect variations of the blackbody emission as a function of wavenumber, that is, $B_0(\tilde{\nu}, z) \sim B_0(z)$. To make the problem linear in the temperature space, we will neglect the nonlinearity of the Planck function as done by Rodgers (2000), so $B_0(z) \propto T(z)$, with the proportionality constant omitted for simplicity. Our linear and discrete forward model for the observed radiance is then

$$B_i = \Delta z\, a_i \sum_{j=1}^{N} T_j e^{-a_i z_j} + \varepsilon_i \qquad \forall\, i = 1, 2, \ldots, M, \tag{4.22}$$

which is the form of Eq. 4.16. Now, x is the vector of T_j, y is the vector of B_i, and the rows of K are the weighting functions, in this case $\Delta z\; a_i e^{-a_i z_j}$. If we select centroids of the weighting function, a_m^{-1}, which vary linearly between 1 and 20 km, we build a matrix K with rows equal to the sequence of functions shown in Figure 11. We see that band $j = 1$ (purple line) is the most opaque and it is mainly sensing temperatures close to the ground, while the most transparent band $j = 8$ (yellow line) also weighs higher altitudes. These weighting functions resemble qualitatively those used by TROPoe with AERI or ASSIST (e.g., Figure 1A in Blumberg et al. 2015).

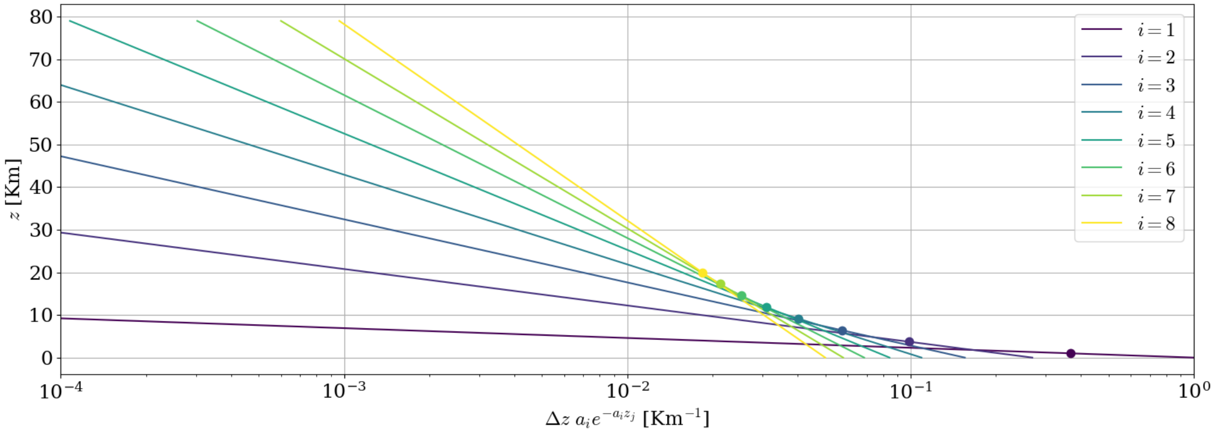

Figure 11. Rows of K matrix or weighting functions for the eight spectral bands of the spectrometer. The dots represent the centroids, equal to $z = a_j^{-1}$.

The missing ingredients for the retrieval are the prior mean, x_a, priori covariance, S_a, and noise covariance, S_ε. The prior mean is taken as the mean temperature profile from the U.S. Standard Atmosphere benchmark (NOAA 1976), and a synthetic prior covariance is given by Eq. 2.82 in Rodgers 2000. The noise covariance is assumed to be diagonal with a constant variance of 0.25 K. Both prior and noise are assumed to follow Gaussian distributions.

It is now possible to apply the retrieval (viz. Eq. 4.17) to our synthetic problem; this is done in a Monte Carlo fashion by generating 10,000 temperature profiles with statistics following the prior. Figure 12 shows the three sample temperature profiles with minimum (a), median (b), and maximum (c) height-averaged error. Three aspects of the OE theory become evident: (1) profiles cannot retrieve small-scale features due to smoothing; (2) small error is associated with cases that fall close to the prior (thanks to the small smoothing error); and (3) the error increases with height, due to the reduced information content farther away from the observer, which also causes the solution to rely more on the prior higher up.

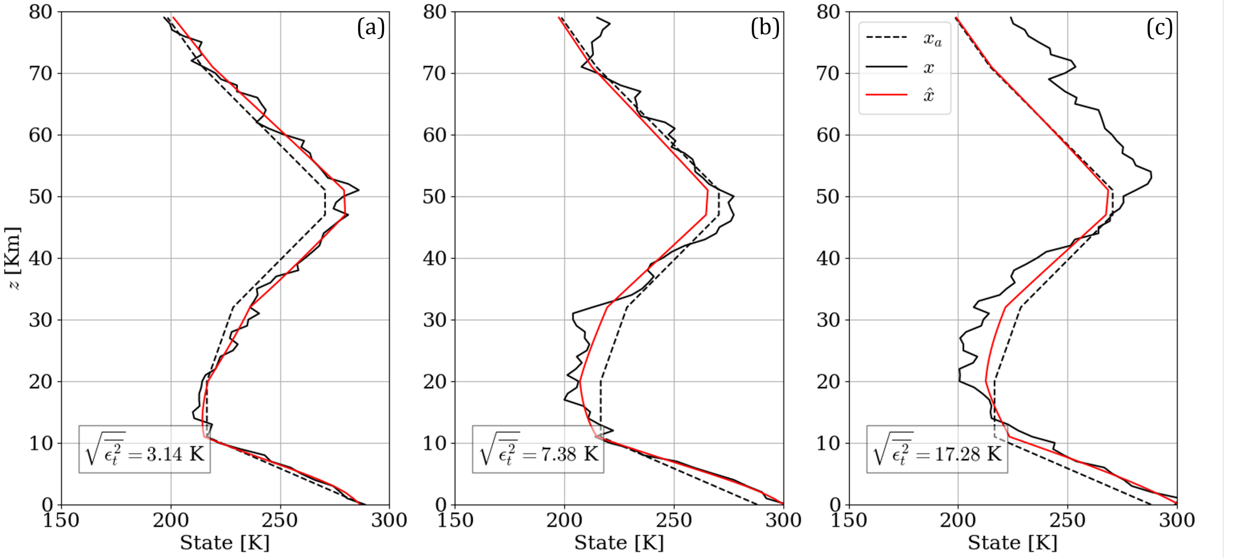

Figure 12. Retrieved profiles with (a) minimum, (b) median, and (c) maximum height-averaged error.

Further insight is provided by inspecting the posterior covariance matrix, $\hat{S}$, starting from its diagonal elements that represent the uncertainty of the retrieval. Figure 13 shows the overall posterior standard deviation (solid black line), its breakdown by smoothing error (dashed black line), and error due to noise (dotted-dashed black line). The constant prior standard deviation is also shown in red for comparison. This visualization confirms that the retrieval error increases with altitude and is dominated by the smoothing part. The sample prior and posterior standard deviation from the 10,000 Monte Carlo draws is also superposed as a sanity check.

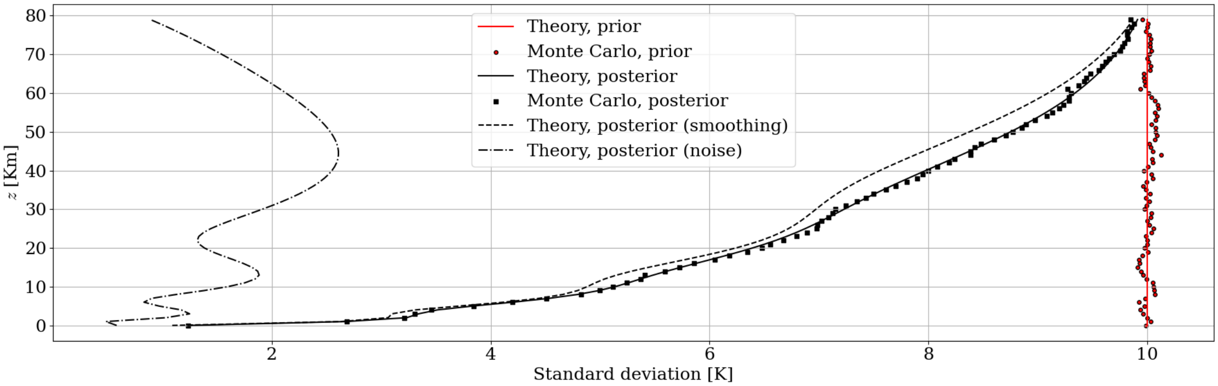

Figure 13. Standard deviation of prior and posterior.

With further examination of the prior and posterior covariance, we visualize the off-diagonal elements of $\hat{S}$ in Figure 14 from both theory (Figure 14a) and Monte Carlo (Figure 14b, just as a verification). As the diagonal value of posterior covariance increases with height, so does the depth of the layer in which a significant correlation of layer-to-layer error occurs. Since the noise was independent of z, this is entirely connected to the larger smoothing that is needed at high altitudes to compensate for the smaller information content away from the surface. This is also the algorithm acknowledging that the temperature profiles are more smoothed out by the weighting functions in the more transparent channels.

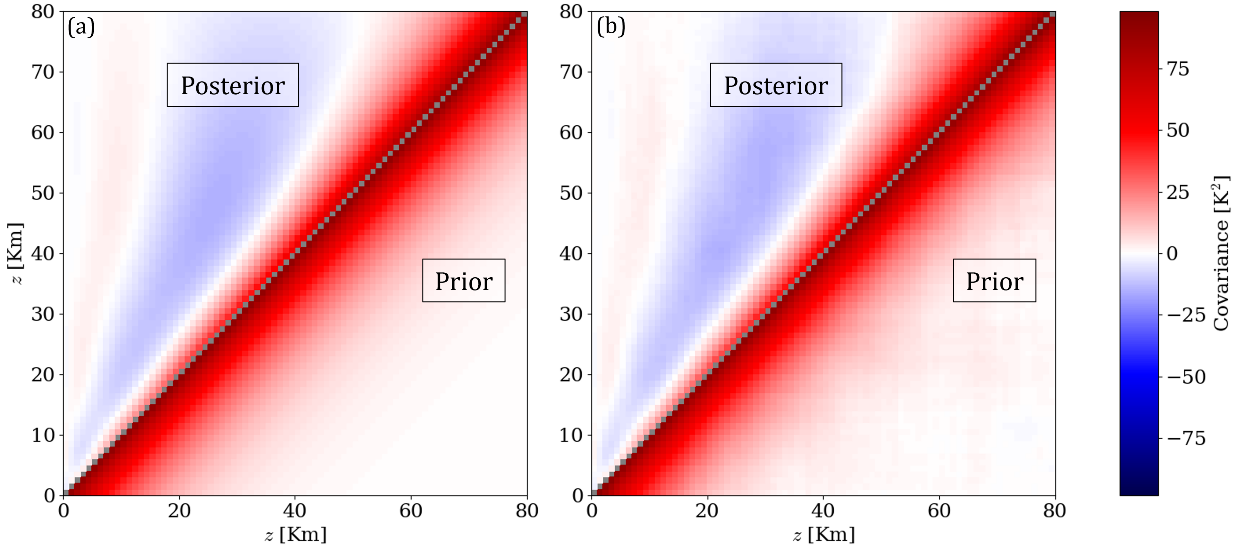

Figure 14. Prior and posterior covariances from (a) OE theory and (b) Monte Carlo.

The fact that the information content decreases with height is supported by the A kernel matrix (Figure 15). The rows of A show how layers closer to the surface have a more peaked shape, meaning that minimum smoothing occurs, and most of the information comes from observations at the target height. Moving up, the peak in the rows of A spreads out, leading to poorer vertical resolution (Figure 15 b) and lower local DFS (Figure 15c). In this simplified example, the vertical resolution, which is very important when interpreting retrieved data, is approximately equal to the height from the ground; a similar behavior is observed in TROPoe for temperature data based on ASSIST spectra.

On a final note, it is quite interesting how this toy problem produced results that are qualitatively similar to those of TROPoe (see Section 5.2), despite being significantly less complex but arguably more insightful and manageable than a full thermodynamic retrieval.

Takeaway 14: In ground-based retrievals, such as TROPoe for ASSIST, the information content decreases with height, resulting in poorer vertical resolution and greater smoothing error at higher altitudes.

4.2 TROPoe

4.2.1 General Architecture

The theoretical concepts and simplified applications discussed so far have laid the foundation for a description of TROPoe. TROPoe is a state-of-the-art retrieval algorithm for thermodynamic profiles that applies optimal estimation theory to ground-based remote sensing observations. TROPoe evolved from the earlier AERIoe, an optimal retrieval method applied to spectral observations from the AERI (Knuteson et al. 2004a), of which the ASSIST is a modern version. In particular, TROPoe was designed to be able to work with a range of observations, including passive spectral observations from ground-based infrared spectrometers and microwave radiometers, active remote sensing observations such as water vapor lidars and radio acoustic sounding systems, and in-situ observations like radiosondes, tower observations, and unmanned aerial systems (Turner and Blumberg 2019). TROPoe is available on Docker (Turner 2025) and is regularly updated.

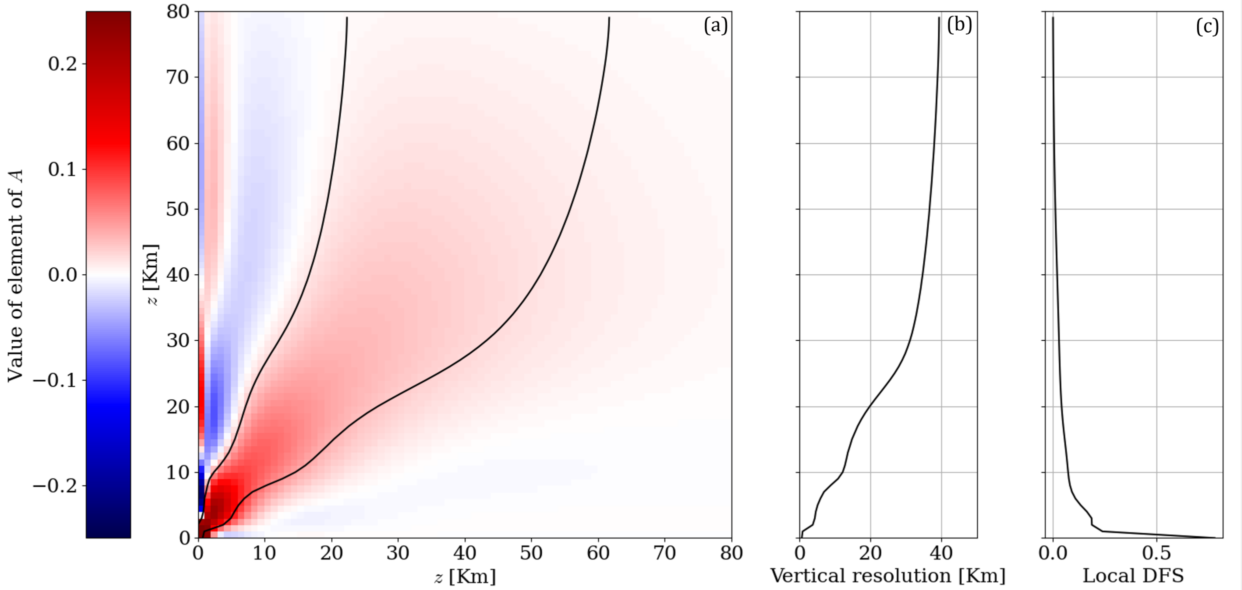

Figure 15. Value of all the elements of the (a) A kernel, (b) vertical resolution based on full-width half-maximum of the rows of A, and (c) local DFS based on the diagonal elements of A. The black lines in (a) represent the boundaries used to calculate the vertical resolution.

The goal of TROPoe is to iteratively find an optimal estimation of the state vector, x, which is a composite of different thermodynamic information, specifically:

$$x = [T_1, T_2, ..., T_N, r_1, r_2, ..., r_N, \text{LWP}, R_{\text{eff}}, \text{others}]^T, \tag{4.23}$$

where T_i, r_i are temperature and water vapor mixing ratio values at N vertically stacked grid points (typically 55), LWP is the liquid water path, and R_{eff} is the liquid water effective radius. The other outputs are relative to the ice clouds and the gas concentrations and are not discussed here.

The general workflow is described in Figure 16, including definitions that are important in the TROPoe realm for both data producers and users. The main input to TROPoe is typically the downwelling spectral radiance observed at the ground level that is used to build the observation vector y. The signal-to-noise ratio of the measured spectra can be increased by processing the radiance through a principal component analysis (PCA) filter (Turner et al. 2006). Only M out of all spectral bands detected by the ASSIST are selected based on their sensitivity to different quantities of interest (i.e., what they measure) in the state vector and the associated optical thickness of the atmosphere (i.e., where they measure). Indeed, spectral regions where other trace gases (e.g., methane, nitrous oxide, ozone, etc.) contribute markedly to the absorption and hence downwelling radiance should be avoided. The spectral bands of the current TROPoe version are reported in Table 2; these were chosen because they are relatively free from the contributions of other gases (Smith et al. 1999).

The accuracy of retrieval is significantly increased by also adding auxiliary inputs to the observation vector, such as the cloud base height (typically from a ceilometer or a lidar), and surface temperature, pressure, and relative humidity. TROPoe can also utilize partial profiles of temperature and/or water vapor from active remote sensors to help constrain the retrieved profiles (Turner and Löhnert 2014; Turner and Lohnert 2021; Djalalova et al. 2022). The need for cloud base height ensues from the inherent challenges that the optically thick clouds cause to the retrieval (see Section 4.2.4), while surface observations help in mitigating biases and re-centering the prior to the observed climate.

TROPoe requires a prior (in the form of mean x_a and covariance S_a, of the state vector) derived from previous long-term measured profiles of temperature and water vapor mixing ratio, r, to constrain the solution to physically realistic profiles. Radiosondes are the most common instruments used to build the prior. Prior statistics were compiled on a monthly basis for the AWAKEN and WFIP3 experiments to better capture seasonal variability. The configuration of the solver is stored in the VIP file and includes, among others, the data paths, the instrument types, and the time resolution.

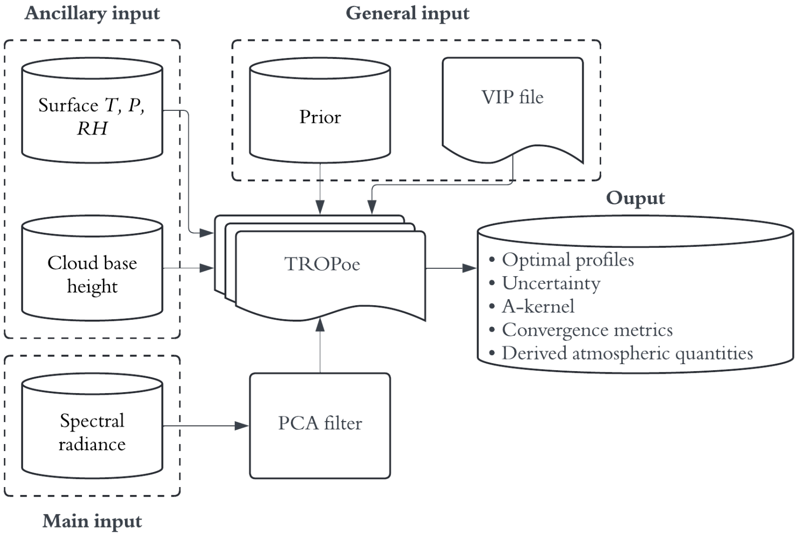

Figure 16. Flowchart of TROPoe.

Table 2. Spectral Bands Used by TROPoe-0.12 0.11.7 With Default Settings (for this version of TROPoe, the length of the observation vector is $M = 347$)

Wavenumber Band [cm^{-1}]	Main Target in State Vector
612.0–618.0	T
624.0–660.0	T
674.0–713.0	T
713.0–722.0	T
538.0–588.0	r
793.0–804.0	r
860.1–864.0	LWP, R_{eff}
872.2–877.5	LWP, R_{eff}
898.2–905.4	LWP, R_{eff}

The TROPoe output contains a great deal of information beyond the optimal solution that requires some degree of familiarity with the theory summarized in this report to be correctly interpreted. For postprocessing and quality control, the full posterior covariance, the A kernel, the vertical resolution, and several convergence metrics are also available. A particularly useful output parameter is the root-mean-square error of the algorithm (RMSA) that measures the difference between the observation vector and the final forward calculation:

$$\text{RMSA} = \sqrt{\sum_{j=1}^{M} \left(\frac{F_j(\hat{x}) - y_j}{y_j} \right)^2}. \tag{4.24}$$

The RMSA provides a nondimensional quantification of the agreement between observations and forward model applied to the retrieval. Large RMSA indicates imperfect convergence or incomplete/flawed physical description of the observed profile. A similar variable in the output files, RMSR, computes the root-mean-square error of only the passive spectral radiance part of the observation vector (i.e., the infrared radiance and/or microwave brightness temperatures) using the same formula.

4.2.2 The OE Algorithm

The inverse problem solved by TROPoe belongs to the multidimensional, moderately nonlinear type. An OE of the state vector can then generally be achieved through the Newton-Gauss method. In the actual implementation, the solution is calculated iteratively by a multidimensional version of Eq. 4.13 as follows:

$$x^{(n+1)} = x_a + \left(K^{(n)T} S_\varepsilon^{-1} K^{(n)} + \gamma^{(n)} S_a^{-1}\right)^{-1} K^{(n)T} S_\varepsilon^{-1} \left[y - F(x^{(n)}) + K^{(n)}(x^{(n)} - x_a)\right] \tag{4.25}$$

K is the Jacobian matrix of the forward problem resulting from the linearization, or

$$K_{ij}^{(n)} = \frac{\partial y_i}{\partial x_j}_{x^{(n)}}. \tag{4.26}$$

The parameter $\gamma^{(n)}$ is a sort of relaxation term that improves the numerical stability and is reduced from a large value (e.g., 1000) to (ideally) 1 as the solution approaches convergence. Convergence criteria are

$$\begin{cases} \left(x^{(n)} - x^{(n+1)}\right) S^{(n)} \left(x^{(n)} - x^{(n+1)}\right)^T < N \\ n < 10, \end{cases} \tag{4.27}$$

where N is the dimension of the state vector, n is the iteration counter. The retrieval is considered converged when one of the two criteria is met, and $S^{(n)}$ is the latest update of the posterior covariance described next.

Since linear error analysis applies only in the neighborhood of the solution, it will suffice to define the uncertainty metrics at convergence. The final posterior covariance is

$$\hat{S} = \left(\hat{K}^T S_\varepsilon^{-1} \hat{K} + \hat{\gamma} S_a^{-1}\right) \left(\hat{K}^T S_\varepsilon^{-1} \hat{K} + \hat{\gamma}^2 S_a^{-1}\right)^{-1} \left(\hat{K}^T S_\varepsilon^{-1} \hat{K} + \hat{\gamma} S_a^{-1}\right) \tag{4.28}$$

which simplifies to the more compact result obtained in the linear case for $\hat{\gamma} = 1$ (see Eq. 4.18). The gain matrix is

$$\hat{G} = \left(\hat{K}^T S_\varepsilon^{-1} \hat{K} + \hat{\gamma} S_a^{-1}\right)^{-1} \hat{K}^T S_\varepsilon^{-1}, \tag{4.29}$$

and the smoothing A kernel is still

$$\hat{A} = \hat{G}\hat{K}. \tag{4.30}$$

Both $\hat{G}$ and $\hat{A}$ can be used to calculate the error due to noise and the smoothing error, as seen in Equations 4.20 and 4.21, respectively.

4.2.3 The Forward Model

The forward model in TROPoe maps the state vector into a ground-based downwelling spectrum of infrared radiance. The core of the model is the Line-By-Line Radiative Transfer Model (LBLRTM, Clough and Iacono 1995), which calculates the radiation spectrum applying the law of radiative transfer (Eq. 3.3) in a discretized fashion across the different atmospheric layers.

One of the most complex aspects of LBLRTM is the estimation of the local absorption coefficient, $a_{\tilde{\nu}}$, in the form of optical depth, $\kappa_{\tilde{\nu}}$. The optical depth is a function of the following local state variables (Clough et al. 1981):

$$\kappa_{\tilde{\nu}} = \tilde{\nu}\, \frac{1 - e^{-\frac{hc\tilde{\nu}}{k_B T}}}{1 + e^{-\frac{hc\tilde{\nu}}{k_B T}}} \sum_i W_i\, \Delta E_i(T) f(\tilde{\nu} - \tilde{\nu}_i), \tag{4.31}$$

where

- W_i is the molar concentration of the i-th interacting species (H_2O, CO_2, N_2, etc.)
- ΔE_i is the energy associated with the energy transition that absorbs a photon with wavenumbers $\tilde{\nu}_i$.

The model parameters (transition energies, line locations, and line broadening) are extracted from the well-established HITRAN database (Rothman et al. 2005; Gordon et al. 2017).

The key point is that the optical depth (and thus the forward model) is a function of the local temperature, density (i.e., pressure), and species concentration. Speaking in terms of OE theory, a targeted choice of the spectral bands to include in the forward model, $F(x)$, allows us to relate the observed spectra (the observations, y) to the temperature and species concentration profiles, including water vapor, and cloud properties (that is, the state x) in different regions of the atmosphere. The dependency of y on the temperature is also present through Planck's function, which regulates the emission.

Another important aspect of Eq. 4.31 concerns the so-called lineshape function, $f(\tilde{\nu} - \tilde{\nu}_i)$. The lineshape function essentially spreads the otherwise infinitely narrow absorption/emission lines to take into account the Doppler and collision broadening (recall Section 3.1). The combination of the Gaussian lineshape of the Doppler broadening with the Lorentzian lineshape of the collision broadening results in a Voigt function. Figure 17 shows three illustrative lineshapes for the limiting cases of a purely Doppler and purely collision type of broadening, as well as a mixed one.

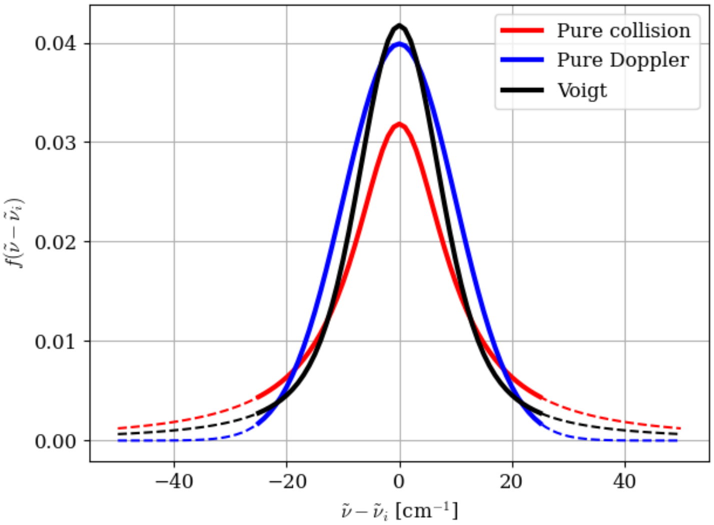

Figure 17. Examples of Voigt lineshape functions for line broadening. The thin dashed lines mark the region beyond the cutoff in the LBLRTM.

The numerical discretization of these lineshape functions requires the Voigt function to be truncated beyond the ±25 cm^{-1} cutoff from the line center. However, contributions to absorption from the tails or "wings" of the lineshape are still significant in the atmosphere. Moreover, the Voigt model is not a perfect description of the real lineshape functions observed in atmospheric applications. Therefore, LBLRTM applies a correction to Eq. 4.31 in the form of a "continuum absorption" (just a corrective term, not to be mistaken for the "real" continuum absorption spectrum observed for ionized gases). The definition of this correction term has been the subject of intense research (Clough, Iacono, and Moncet 1992; Turner et al. 2004) and its contribution to the solution should not be underestimated.

In general, LBLRTM in TROPoe is expected to have an algorithmic error of less than 0.5% and is mostly due to line parameters (i.e., ΔE_i, $\tilde{\nu}_i$). Newer versions also include a module that simulates the discussed nonideal effects of a Fourier transform spectrometer (like the ASSIST) on the observations, further increasing its accuracy.

4.2.4 Sources of Uncertainty

To conclude this chapter, we discuss the different sources of error in TROPoe and possible mitigation strategies. TROPoe provides an embedded uncertainty quantification, which is one of its main advantages compared to other retrievals. However, as hinted earlier, not all errors can be predicted analytically, and we discourage data users from blindly using the uncertainty provided by TROPoe as the only metrics for data quality.

Known Uncertainties

An important advantage of a physical OE method like TROPoe is the estimation of the uncertainty of the solution in the form of the posterior p.d.f., $p(x|y)$. As discussed previously, this comprises the uncertainty caused by the smoothing and noise of the instrument. Regarding the latter, TROPoe assumes that the error in the observed spectral radiance is unbiased and Gaussian, and the band-to-band correlation is negligible. This noise behavior in the ASSIST is generally confirmed by the third blackbody calibrations, where the 3-σ error between the measured and theoretical spectrum of the third blackbody should not exceed 1% of the blackbody radiance at ambient temperature, and the band-to-band correlation coefficients should not exceed 0.35 (Turner and Blumberg 2019). Improved estimates that include correlated errors between different bands have also been attempted but remain significantly more complex (Maahn et al. 2020). Operatively, the standard deviation of the noise, which is sufficient to fully describe the noise p.d.f., $p(y|x)$, with the present assumptions is estimated based on the imaginary radiance, as described in Section 2.7.

On the other hand, the smoothing error can be calculated using Eq. 4.21. Another way to assess the impact of smoothing error is to compare TROPoe profiles with reference observations such as radiosondes. Provided that the reference instruments represent the ground truth (which, in the case of radiosondes measuring water vapor, can be questionable (Turner et al. 2004)), the total difference between TROPoe and the reference profiles is a good estimate of the total error. By applying a TROPoe-like smoothing to the reference and reassessing the difference with the TROPoe profiles, one can get an estimate of the contribution of the smoothing only to the global error. We provide an example of this validation in Section 5.3. Using this method, Turner and Löhnert (2014) observed a slight reduction in RMS error of retrieved profiles with radiosondes mainly due to the reduced vertical sampling error in the radiosondes. In contrast, Blumberg et al. (2017) reported a drastic reduction in mean and total error when applying smoothing to the radiosonde profiles, indicating the limited spatial resolution of the retrieval as a major error driver. However, both authors highlight the presence of residual biases in their retrieval, indicating that other physical effects can play a significant role in determining the accuracy of the retrievals. These effects, which are described next, must be considered carefully since their contribution to the uncertainty may not be fully captured by the posterior p.d.f. calculated by TROPoe.

Instrumental bias

Bias in the detected spectra can propagate into the retrieval and increase the overall uncertainty but not necessarily be captured by the posterior p.d.f. This makes instrumental bias particularly insidious and has spurred important advances in the field of spectroscopy. Any bias in the radiance invalidates the assumption of unbiased noise at the basis of OE theory. When working with measured spectra, one should consider not only the error in the magnitude of the radiance, but also in the wavenumber. In fact, errors can arise because of an imperfect spectral calibration resulting in absorption lines being shifted from their theoretical wavenumbers. In general, instrumental bias in radiance and wavenumber can arise in the infrared spectrometer data from any of the mechanisms described in detail in Section 2, which justifies the efforts made to mitigate such errors. The high accuracy of spectrometers such as the ASSIST has been achieved mainly thanks to online radiometric calibration (Section 2.2). In this regard, a clever way to estimate the residual error due to radiometric calibration is to linearly expand Eq. 2.10 and assess the contribution of the uncertainty of blackbody emissivities and temperatures. Knuteson et al. (2004b) and Turner et al. (2004) that applied this method to the AERI and confirmed that the radiometric error was within specifications.

The requirements of ASSIST are a 3-σ uncertainty in the radiance less than 1% of the blackbody radiance at ambient temperature and a wavenumber error less than < 0.01 cm^{-1}. These specifications are tested in every third blackbody calibration. Therefore, instrumental biases from ASSIST measurement can generally be neglected compared to other sources of uncertainty, provided that the third blackbody calibrations, ideally before and after each deployment, indicate the radiometric and spectral specifications are met. It is also essential to routinely monitor the health of the instrument, following the example reported in Section 5.1. Finally, it is important to ensure that the field of view of the third blackbody during the calibration is the same as that of the sky scene. In fact, even a small obscuration (for example, a tree branch) can lead to severe biases (Mlawer and Turner 2016).

Clouds

Thermodynamic retrievals from a passive IRS such as the ASSIST face additional challenges in the presence of overhead clouds compared to clear-sky conditions. First, clouds become rapidly opaque in the infrared regions probed by the ASSIST as the LWP increases. At LWP > 60 g m^{-1}, clouds can be considered completely opaque to infrared radiation (Turner et al. 2007). From a physical point of view, this opacity translates into the weighting functions (which, as we recall, describes how much information from a certain height at a certain wavenumber reaches the surface) dropping practically to 0 above the cloud. From a mathematical point of view, this implies that the ground-based spectrometer cannot access any information of the thermodynamic profiles above the clouds, and the spectral bands become rapidly saturated with radiation emitted by the cloud base itself, which overshadows information coming from the layers below it.

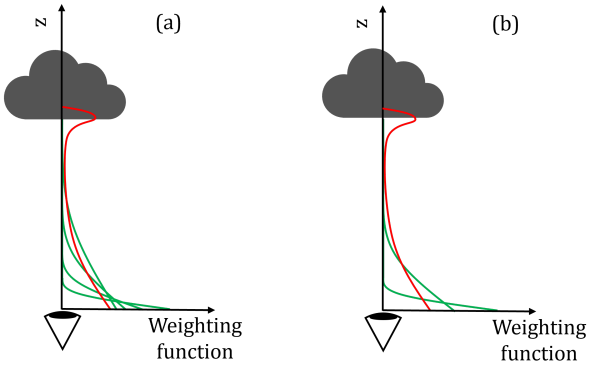

Figure 18. Ideal representation of the weighting functions for (a) temperature and (b) water vapor mixing ratio for an infrared spectrometer in the presence of a cloud. Green lines are weighting functions not including contributions from the clouds, red lines are those with significant values within the cloud layer.

A natural question is: How well does TROPoe perform below the clouds? Temperature profiles, which close to the surface rely on channels that are relatively opaque even in clear-sky conditions, have been shown to be just as accurate as those obtained in clear-sky conditions at heights "sufficiently below the cloud" base (Blumberg et al. 2015; Turner and Löhnert 2014; Turner and Blumberg 2019). Turner and Blumberg (2019) clarify that "sufficiently below the cloud" means heights where there is not a significant amount of information coming from the cloud base as described by the local smoothing, i.e., the rows of the A kernel matrix or the vertical resolution. These opaque channels have peaked weighting functions that decay before reaching the cloud layer, which means that at those wavenumbers the atmosphere is so optically thick that radiation from clouds cannot be seen from the ground. The same concept applied to water vapor, which uses flatter weighting functions than the temperature, implies that the accuracy of the water vapor estimate deteriorates faster with height below the clouds, even a few meters above the surface. This concept is shown schematically in Figure 18, where the temperature (Figure 18a) has more channels with weighting functions that do not include the contribution of the clouds than the water vapor (Figure 18b). In general, a rule of thumb when dealing with cloudy scenes is to exclude profiles above the first cloud base height (see Section 5.2).

A big advantage of TROPoe is the ability to also detect cloud properties, such as LWP and effective droplet radius, R_{eff}, which are important in several contexts. A big caveat is that the optical thickness of the cloud, which is the main property that can be inferred from the ASSIST, is a function of both LWP and R_{eff}. A blind comparison of several sophisticated retrievals that are predecessors of TROPoe showed unacceptably large scattering in the estimates of these properties (Turner et al. 2007). However, awareness of cloudy scenes in TROPoe helps limit errors below the cloud base that would arise if the clear sky was assumed all the time (Turner and Löhnert 2014). Residual errors below the clouds have been attributed to the increase of atmospheric scattering due to clouds, which is not modeled in TROPoe, as discussed next.

Takeaway 15: Thermodynamic profiles above the cloud have high uncertainty as a result of the lack of information penetrating the optically thick cloud layers. Below the clouds, the temperature profiles are generally more reliable than water vapor profiles.

Prior

The prior can be a source of bias if it is not representative of the climatology during the observation period (Maahn et al. 2020). The prior could be statistically inaccurate in at least two ways:

- If the prior statistics are not Gaussian; in this case, a clever transformation in the state vector space could fix the issue
- If the mean, x_a and/or covariance, S_a (including off-diagonal terms) are not a good measure of the state vector observed during the experiment.

In general, the prior should be derived from as many observations as possible and from a climatologically equivalent location, preferably on a monthly basis. A prognostic way to check whether the prior used is amenable for TROPoe retrievals is a Gaussianity check on the prior dataset. A diagnostic way to check whether the prior is adequate to represent the observation is to perform two χ^2 tests, the first applied to the difference between the observation vector and the prior transformed into the observation state, $y - y_a$, and the second to the difference between the optimal solution and the prior mean, $\hat{x} - x_a$ (see Chapter 12 in Rodgers 2000 for more details). However, the version of TROPoe at the time this report is written does not provide all the necessary outputs to carry out these statistical tests.

Forward Model

The LBLRTM is the backbone of TROPoe, and any inaccuracies associated with it can have important consequences. Most importantly, biases in the LBLRTM are not included in the posterior covariance because the Bayesian OE framework assumes a perfect forward model (Maahn et al. 2020). In theory, incorporating uncertainties of the forward model in the posterior is possible, although very complex and computationally heavy, so TROPoe assumes no LBLRTM error and offers the possibility of inflating the instrumental noise to compensate for this missing error component (Turner and Blumberg 2019).

The relevance of the radiative transfer model explains the extensive efforts to improve and validate the LBLRTM, for instance, in the Atmospheric Radiation Measurement (ARM) program, which have led to the maximum error on atmospheric spectra to go from ~ 30 r.u. in 1990 to less than ~ 5 r.u. in 2010 (Mlawer and Turner 2016). Quantifying the error due to the forward model for real applications remains an open research question. Thus, in the remainder of this section, we will just describe the main sources of error without attempting to provide a solution.

A limitation of LBLRTM is that it does not account for atmospheric scattering. Scattering occurs due to collision between photons and particles in the atmosphere and has markedly different behaviors based on the wavelength of the incident radiation and the size and shape of the scattering object (Siegel 1971). What is relevant to this discussion is that scattering results in a contribution to radiation sensed by the ASSIST from all directions. This multi-directionality means that efficient use of the law of radiative transfer (Eq. 3.3) along the vertical direction alone is insufficient. Complexity and computational efficiency are increased by several orders of magnitude if scattering is included. Fortunately, at the wavenumbers used by TROPoe emission and absorption are the dominant phenomena, so scattering can generally be neglected. However, there are instances where the scattering can play a role. Turner and Löhnert (2014) showed that during cloudy conditions, scattering can represent up to 12% of the spectral radiance on specific bands, which were indeed removed from the latest TROPoe version. Other special conditions where scattering becomes important are dust storms (Mlawer and Turner 2016).

Another modeling aspect that has received much attention is the parameterization of the continuum absorption/emission, especially from the water vapor (Turner et al. 2004; Mlawer and Turner 2016). As seen in Section 4.2.3, the line absorption/emission model requires a correction to address the inaccuracies of the lineshape functions far from the line center. Improvements in the continuum absorption/emission have been pivotal in improving the LBLRTM; this was achieved mostly thanks to the AERI observations. Recent research has demonstrated that there needed to be an adjustment made to the strength of the water vapor continuum absorption in the infrared window (i.e., between 8 and 12 μm) (Mlawer et al. 2024); current work demonstrates that this change to the radiative transfer model only has an impact when the precipitable water vapor in the column is larger than approximately 3.5 cm (Turner, personal communication, 2025).

In general, to rule out the presence of major biases due to LBLRTM, we recommend comparing TROPoe results with other nearby sensors that use different measurement principles, the golden standards being radiosondes and meteorological masts. This is especially true for applications requiring highly accurate uncertainty quantification of thermodynamic profiles. However, users should take care not to believe that either radiosondes or in-situ sensors on masts are free of systematic errors, and thus any differences between TROPoe retrievals and other observations needs to be investigated carefully.

Numerical Inaccuracy

It may seem odd to discuss numerical errors in the context of an experimental topic, but we have already seen how computationally intensive the estimation of thermodynamic profiles in TROPoe is. Most of the CPU time is spent during the iterative calculations of the Jacobian of the forward model, the LBLRTM. Computational errors (besides the obvious truncation errors that depend on the machine) can arise, for instance, as a consequence of the discretization of the law of radiative transfer in the LBLRTM. In this regard, Clough et al. (2005) report an algorithmic accuracy of 0.5%, which in any case much lower than the uncertainty on the line parameters. Therefore, the numerical error in the forward model itself can generally be ruled out. Another possible source of error is the incomplete convergence of TROPoe. Classifying this type of inaccuracy as a "numerical error" in a strict sense may not be fully correct, since troublesome convergences are generally associated with physical inconsistencies between the observed noise profile and the physics described by the forward model. However, the example in Section 5.2 shows a simple method to flag profiles that did not converge satisfactorily to a solution.

5 Case Studies

This section includes a collection of hands-on case studies in which we apply the fundamentals of spectroscopy, radiation physics, and OE reviewed in the previous sections. The selected examples are generally useful to ASSIST and TROPoe users, alongside the list of TROPoe output variables provided in Appendix A.

5.1 Routine Health Checks for ASSISTs

Here, we provide step-by-step actionable guidelines to monitor the operation of an ASSIST and to ensure correct functioning. The ASSIST is an extremely accurate but delicate spectrometer. In fact, for the radiometric and spectral specifications to be met and the estimated thermodynamic profiles to be accurate, all data processing modules and the hardware components described in Section 2 need to operate within strict margins of their nominal configuration.

In addition to the recommendation to perform a third blackbody calibration after and before each deployment, which was reiterated several times, there are routine checks that can help identify issues in the data acquisition chain in a timely manner.

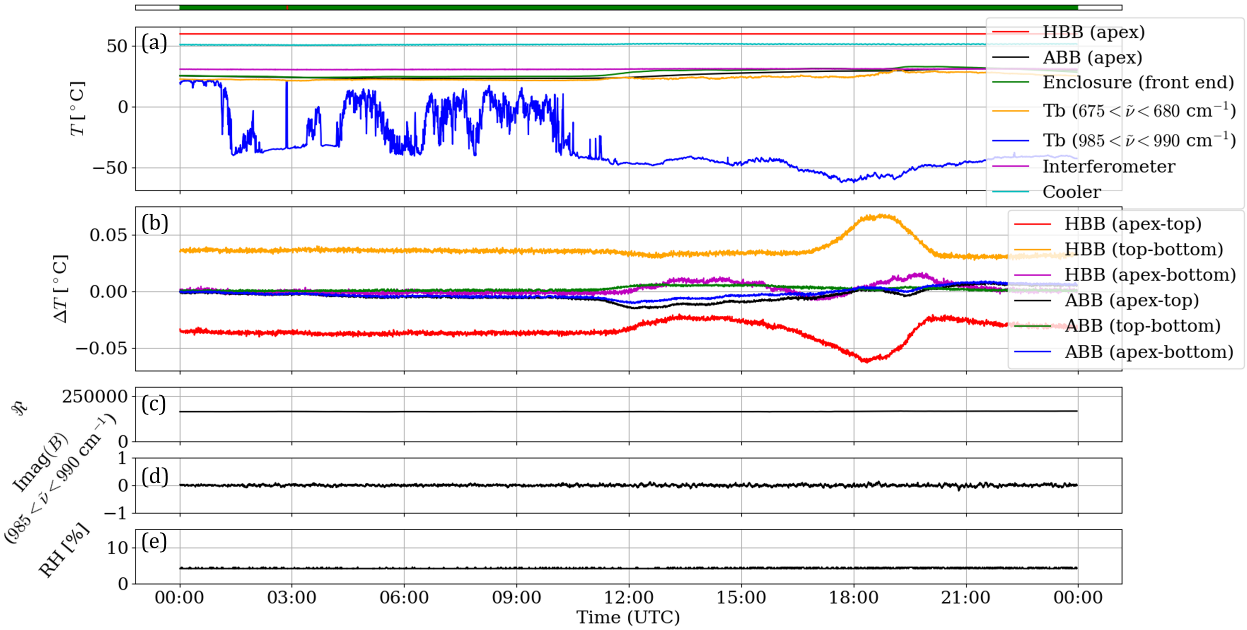

Figure 19. Daily health check summary for the NREL ASSISTs deployed at the WFIP3 Rhode Island site on 2024-07-14: (a) temperature from different channels, T; (b) temperature differences across black bodies, ΔT; (c) mean responsivity in channel A, $\Re$; (d) mean imaginary radiance, Imag(B), between 985 and 990 cm $^{-1}$; (e) relative humidity in the interferometer, RH. The bar on top indicates the hatch status (green=open, red=closed).

Figure 19 is a summary of the automated daily health check for one of the NREL ASSISTs in WFIP3. It includes several channels directly available in the "assistsummary" daily file. The meaning of each quantity is explained in Table 3. The user should also check for inconsistencies in the sampling rate that can be indicative of hardware problems. Should any of those parameters exhibit behavior different from the expectations, we recommend contacting technical support and flagging data accordingly.

5.2 Example of Daily Thermodynamic Profiles

Let us now take a closer look at the TROPoe output based on one day of infrared spectra collected by one of the NREL ASSISTs deployed at the WFIP3 Rhode Island site during the summer of 2024 (Letizia 2024). The selected UTC day is 2024-07-14 and is characterized by cloudy nighttime conditions ($\sim$ 00:00-09:00 UTC) and clear skies during the daytime ($\sim$ 09:00-24:00 UTC), local time being UTC-4.

Table 3. List of Health Parameters for an ASSIST.

Label	Meaning	Expected behavior
hbb_*_temperature	Temperature of HBB at different points	Stay within 60 ± 0.1°C
abb_*_temperature	Temperature of ABB at different points	Be close to ambient temperature and within ± 0.1°C from each other
front_end_temperature	Temperature of front end of enclosure	Track ambient temperature, with less thermal inertia than ABB
mean_tb_675_680	Brightness temperature in 675–680 cm^{-1} band	Track the ambient temperature (e.g.,front end)
mean_tb_985_990 cm^{-1}	Brightness temperature in 985–990 cm^{-1} band	• If hatch is closed, be equal to T_b in 675–680 cm^{-1} band • If hatch is open, sky is clear, be much colder than ambient • If hatch is open, sky is cloudy, be slightly colder than the ambient
interferometer_temperature	Temperature of the interferometer	Be constant
cooler_block_temperature	Temperature of the detector's cooler	Be constant
lw_responsivity	Mean responsivity of channel A	Be constant ($\sim 10^5$, units not relevant)
mean_imaginary_rad_985_990	Mean imaginary radiance in 985–990 cm^{-1} band	Be much smaller than 1 r.u. and non autocorrelated
interferometer_humidity	Relative humidity of interferometer	Stay below 10%

The retrieval used TROPoe-0.12 0.11.7 to process "instantaneous" spectra from the ASSIST complemented by the cloud base height from a Halo XR scanning lidar (Bodini 2024) using a CNR-based cloud detection (see Newsom et al. (2019)) and surface pressure, temperature, and relative humidity from a collocated met mast (Pekour 2024). The full pipeline is available at Letizia, Adler, and Turner 2024.

It is always recommended to quality control (QC) the retrieval by rejecting points that meet the following criteria:

- Data above the detected cloud base height and corresponding to LWP above a certain threshold (recommended value is 5 g m^{-2}); this excludes region beyond optically deep cloud layers that provide little to no information to the spectrometer. An even safer approach is to remove data below cloud height half the local vertical resolution (Turner and Blumberg 2019).
- Profiles with $\hat{\gamma}$ above a certain threshold (the strictest value is 1), to get rid of profiles affected by troublesome convergence of the Gauss-Newton algorithm.
- Profiles with RMSA greater that a certain threshold (recommended value is 5), to exclude occurrences where the spectral radiance are poorly described by the LBLRTM applied to the retrieved state.

The above QC can be conveniently applied to the TROPoe output open through *xarray* in Python as follows:

```
Data['cbh'][Data['lwp']<min_lwp]=Data['height'].max()#remove clouds with low lwp
Data['temperature_qc']=Data['temperature'].where(Data['gamma']<=max_gamma)\
    .where(Data['rmsa']<=max_rmsa).where(Data['height']<=Data['cbh'])#QC temperature
Data['waterVapor_qc']=  Data['waterVapor'].where(Data['gamma']<=max_gamma)\
    .where(Data['rmsa']<=max_rmsa).where(Data['height']<=Data['cbh'])#QC mix. ratio
```

The quality-controlled temperature and water vapor mixing ratio are visualized in Figure 20b and Figure 20d, respectively. The daytime heating of the ground is well captured, as is the negative temperature and humidity gradient as a function of height above 0.5 km. Several data points lying above clouds with LWP > 5 g m^{-2} are excluded, as are 6

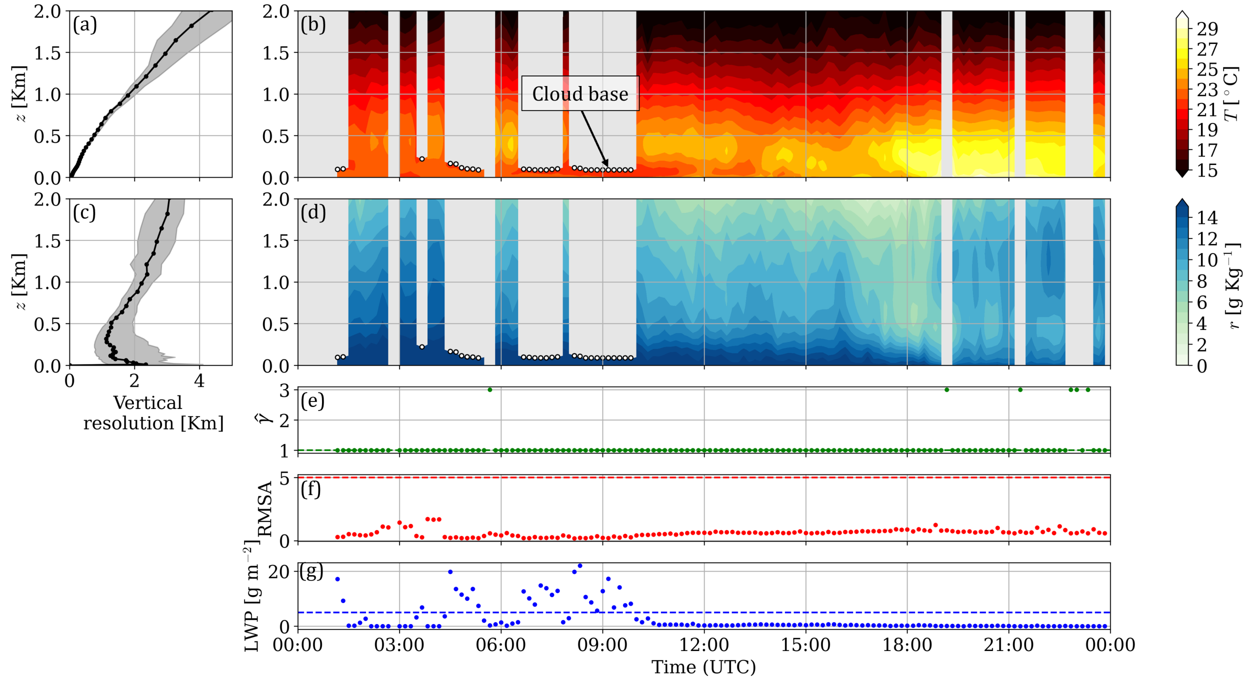

Figure 20. TROPoe retrieval from the WFIP3 Rhode Island (rhod) site on 2024-07-14: (a) median vertical resolution of temperature profiles, where the shaded area marks the interquartile range; (b) temperature profiles; (c) median vertical resolution of water vapor mixing ratio; (d) 10-minute profiles of water vapor mixing ratio; (e) time series of $\hat{\gamma}$; (f) time series of RMSA; (g) time series of LWP. The dashed lines in (e,f,g) represent the upper (for $\hat{\gamma}$, RMSA) or lower (for LWP) thresholds used in the quality control.

full profiles that have $\hat{\gamma} = 3$. The RMSA never exceeded the upper limit of 5. The same figure also shows the median vertical resolution for temperature (20a) and mixing ratio (20b). The vertical resolution of temperature increases roughly linearly with height and at a rate comparable to the height itself, being the probable cause of the lack of a sharp capping inversion at the top of the boundary layer. The resolution of the water vapor mixing ratio is instead always close to 2.5 km, except for the very first grid point that comes from the met station. TROPoe developers are continuing to work on extracting more information, especially on water vapor, from the infrared spectra which would improve the vertical resolution.

The 1-σ posterior uncertainty maps of the profiles are shown in Figure 21 and provide a great deal of essential information. The uncertainty generally increases with height and larger uncertainties are located above the thicker clouds, as expected.

Examining the data further, we can visualize the A kernel at convergence for a cloudy and cloud-free profile (Figure 22). The matrix has been split into (1) the part describing the retrieved temperature as a function of real temperature (rows and columns 1–55 or quadrant III); (2) the part describing the retrieved mixing ratio as function of real mixing ratio (rows and columns 56–110 or quadrant I); (3) the part describing the retrieved mixing ratio as a function of real temperature (rows 56–110 and columns 1–55 or quadrant II); and (4) the part describing the retrieved temperature as a function of real mixing ratio (rows 1–55 and columns 56-110 or quadrant IV).

When clouds hover above ASSIST, the A kernel has larger values for those columns corresponding to temperature and mixing ratio below the cloud base as a result of the larger amount of information they contain compared to the blind regions above the clouds. There is also a high value at the height corresponding to the cloud base as a result of the strong emission from the clouds. The dark red in the first layer in quadrant I indicates the high weight given to the water vapor information from the met station. In addition, recalling how the smoothing error is proportional to the departure of $\hat{A}$ from the identity matrix, we identify the temperature profiles below the clouds as the most trustworthy region of the retrieval (bottom-left corner of Figure 22a). Outside of this narrow region, the A kernel is arguably very different from I_N. Figure 22, which corresponds to clear-sky conditions, tells a completely different story. Here, the A kernel for both temperature and mixing ratio resembles I_N, with larger values on the diagonal. Many off-diagonal

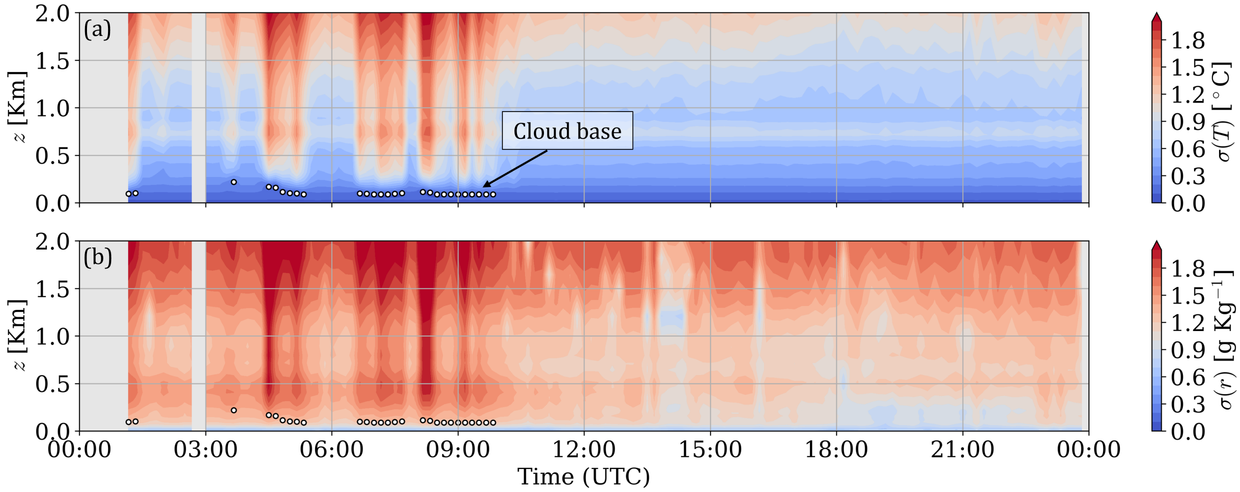

Figure 21. 1-σ uncertainty of TROPoe retrieval from the WFIP3 Rhode Island site on 2024-07-14: (a) uncertainty of temperature; (b) uncertainty of water vapor mixing ratio.

elements are still not negligible as a consequence of the smoothing, and there is also information shared between T and r, as the nonzero values in quadrants II and IV highlight. The analysis of the A kernel is more advanced and requires a deeper knowledge of TROPoe and the physics of radiation but is still insightful and sometimes unavoidable, as shown in the next example.

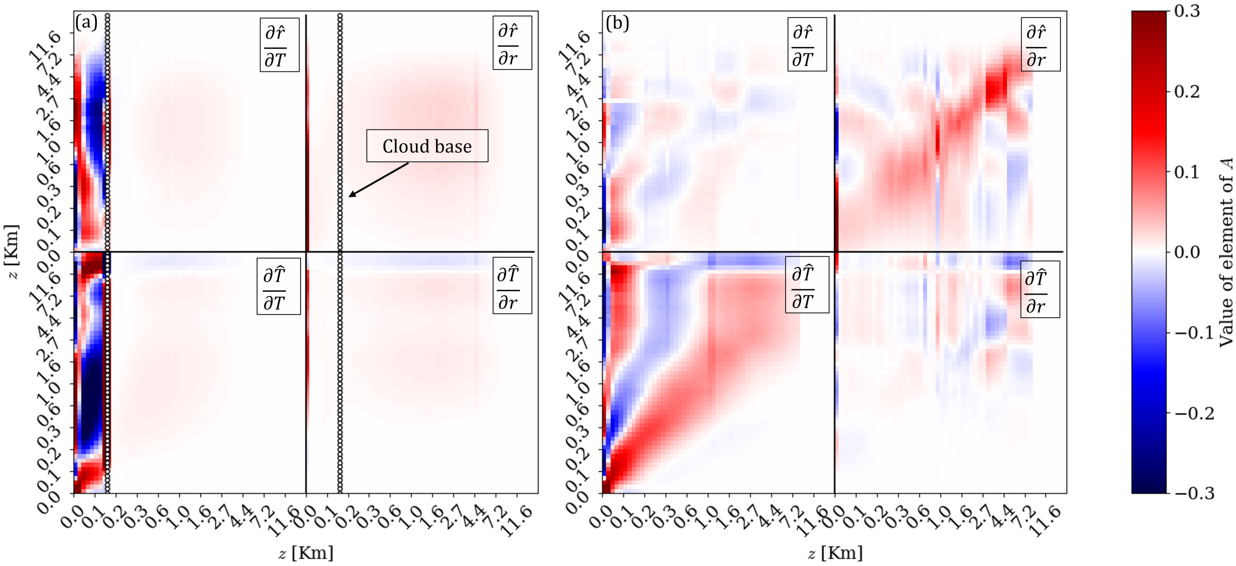

Figure 22. A kernel of TROPoe retrieval from the WFIP3 Rhode Island site on 2024-07-14: (a) A kernel at 08:20 UTC, when low clouds were present; (b) A kernel at 17:50 UTC, during clear-sky conditions.

5.3 Comparing With Radiosonde Measurements Using the Averaging Kernel

As part of AWAKEN, three ASSISTs were deployed near wind plants in northern Oklahoma alongside a suite of other instruments, including radiosondes (Moriarty et al. 2024). This campaign provides an opportunity to compare ASSIST+TROPoe retrievals with measurements obtained by the radiosonde launched from a nearby site. The comparison is done using both the unsmoothed radiosondes at the native resolution and the smoothed radiosonde profiles as ground truth. We smooth the radiosonde measurements using the averaging kernel, $\hat{A}$, introduced in Section 4.1.3 as

$$x_{\text{smooth}} = x_a + \hat{A}(x - x_a), \tag{5.1}$$

where x is the observation from the radiosonde interpolated to the 55-level-height grid used for ASSIST retrievals and x_{smooth} is the smoothed radiosonde profile. Note that all variables include temperature and water vapor mixing ratio, appended together, to account for cross-talk between temperature and humidity. A similar process is conducted by Blumberg et al. (2017).

Example temperature and mixing ratio profiles, taken on 2023-08-11 at 02:23 UTC (2023-09-10 21:23 local time), are presented in Figure 23. In this example, several features are apparent. In general, the ASSIST retrieval closely matches the smoothed radiosonde profile, indicating that the primary difference between the two instruments lies in the vertical smoothing. The temperature profiles show better agreement than the mixing ratio, as humidity profiles are generally more difficult to retrieve from the spectra than temperature. The dependence of vertical resolution on height is also apparent. The strong temperature inversion near the surface is captured by both the ASSIST and the radiosonde, whereas the temperature inversion measured by the radiosonde 1.5 km above the ground is smoothed over in the ASSIST retrieval.

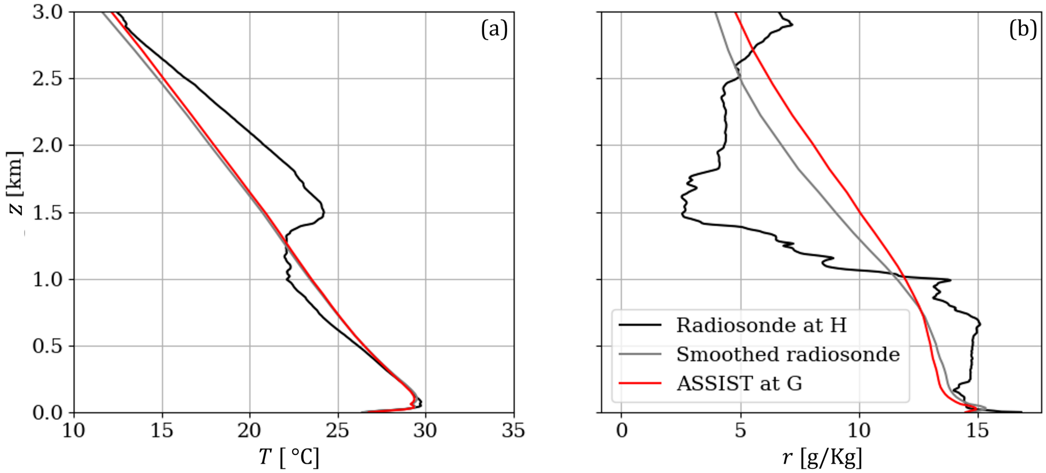

Figure 23. (a) Temperature and (b) water vapor mixing ratio profiles retrieved at AWAKEN on 2023-08-11 at 02:23 UTC from the ASSIST at site G, along with the unsmoothed and smoothed profiles from radiosondes launched from the nearby site H.

Now, how can we decide whether the agreement between TROPoe profiles and the reference profile satisfactory? In other words, what is the error we should expect for the TROPoe retrievals in the best-case scenario? For unsmoothed profiles, the answer is straightforward: The total error, assuming naively that the radiosondes represent the truth, should have a covariance equal to the posterior $\hat{S}$. Any additional error could be due to all the other reasons outlined in Section 4.2.4 and/or spatial and temporal heterogeneity between ASSIST and the radiosonde.

For smoothed profiles, the answer is different. In fact, what Eq. 5.1 really intends to do is apply to the reference profiles the same level of smoothing as TROPoe. In theory, all we are left with is the error due to noise. In fact, as we have seen for the linear case, but also applicable to the moderately nonlinear problems, the total error of TROPoe can be expressed as a sum of smoothing error and error due to noise:

$$\hat{x} - x = (\hat{A} - I_N)(x - x_a) + \hat{G}\varepsilon. \tag{5.2}$$

Equation 5.1 also can also be reformulated as

$$x_{\text{smooth}} - x = (\hat{A} - I_N)(x - x_a). \tag{5.3}$$

By subtracting the former two equations, we immediately find that the error between the retrieval and smoothed profiles includes only the noise part, that is, $\hat{x} - x_{\text{smooth}} = \hat{G}\varepsilon$. The comparison with and without smoothing is therefore useful mainly to assess the impact of smoothing on the total error. If the residual error is seen to depart significantly from the estimated error due to noise, either in magnitude or because it is systematic rather than random, then one should look for other causes of inaccuracy.

Takeaway 16: The error between TROPoe retrievals and unsmoothed radiosonde profiles will have a covariance equal to the posterior covariance if no other errors are present. The error between TROPoe retrievals and smoothed radiosondes will include only the noise part if no other errors are present.

5.4 Definition of Planetary Boundary Layer Height

The TROPoe algorithm also outputs the planetary boundary layer height (PBLH), derived from the retrieved profiles. The method used to calculate PBLH is based on the parcel method, which identifies the height at which a parcel of air from the surface would be in equilibrium with its environment (Holzworth 1964). Unlike other implementations of this method (e.g., Coniglio et al. 2013; Duncan Jr. et al. 2022), TROPoe uses the potential temperature profile (θ) rather than the virtual potential temperature in order to eliminate the additional uncertainty introduced by humidity. In the TROPoe implementation, PBLH is defined as the first height where $\theta = \theta_{\text{surf}} + \sigma(T) + 0.5\text{K}$, where θ_{surf} is the potential temperature of the lowest level of the retrieval (0 m above ground level). A PBLH floor is set to 300 m, as the parcel method is most effective under convective conditions; a surface inversion-based method for stable conditions has not yet been implemented.

A sample day from the ASSIST at site G of the AWAKEN campaign is presented in Figure 24, along with an example θ profile showing the method used to calculate PBLH. Between 01:00 and 14:00 UTC (20:00 and 09:00 local time), PBLH is set to the minimum allowable value of 300 m. After 14:00 UTC, the increase in the PBLH characteristic of daytime convective conditions is observed. The lapse rate above $z \sim 300$ m is always positive because TROPoe forces the potential temperature profile to increase monotonically with height above a fixed level above the ground. This realistic constraint proved useful to improve overall accuracy and enhance convergence (Turner and Blumberg 2019), but the user should not be surprised to find no occurrences of negative lapse rate far from the surface in the TROPoe results.

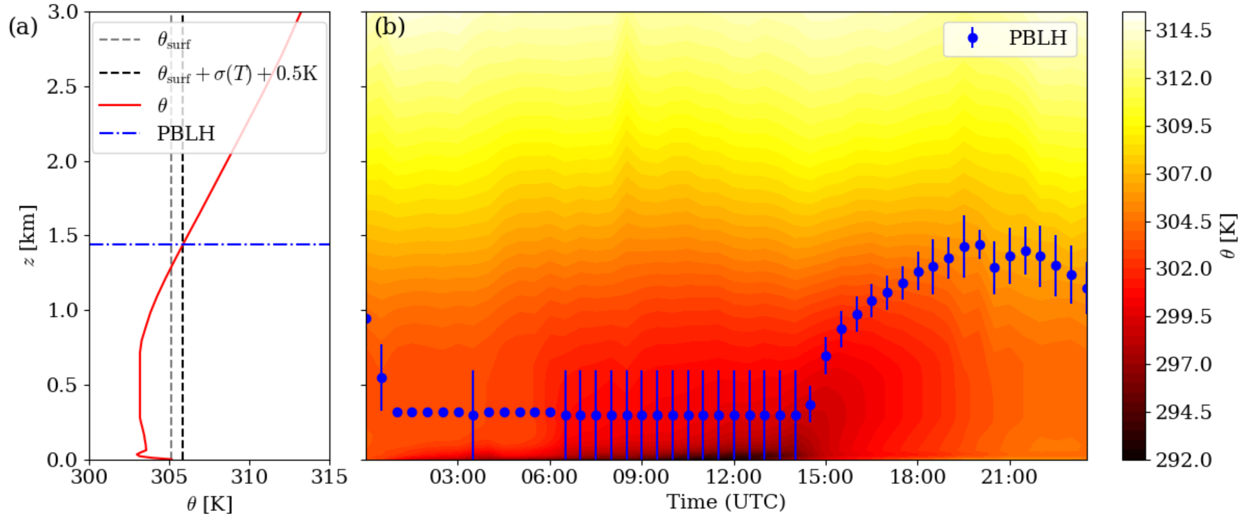

Figure 24. Potential temperature, θ, and PBLH throughout a diurnal cycle (2023-08-15) at the AWAKEN site, including a sample θ profile from 20:00 UTC (15:00 local time). The sample profile illustrates the method used to compute PBLH.

Takeaway 17: TROPoe profiles can be used to accurately assess the local static stability of the boundary layer, especially near the surface. The main limitations are the smoothing of capping inversions that occur far from the surface, the hard-coded stable lapse rate above a fixed height.

6 Conclusions

This report provides a practical guide to the theoretical framework behind the thermodynamic profiling process using ASSIST infrared spectrometer data and TROPoe retrievals. To answer the question “How do we measure spectra?” we reviewed how the interferometer inside the ASSIST is capable of measuring the spectrum of infrared radiation and the signal processing techniques required to mitigate several technological limitations. To address the question “Why do we measure spectra?” we discussed the fundamental physics of infrared radiation in the atmosphere and how the thermodynamic profiles affect the measured spectra. Finally, we tackled the question “How do we estimate thermodynamic profiles?” by describing the optimal estimation method used by TROPoe to translate the infrared spectra into useful profiles of temperature, humidity, and other atmospheric quantities.

This work is part of a larger effort to provide detailed and rigorous documentation of experimental data for wind energy applications. We deem that the access to reliable and comprehensible knowledge of experimental techniques is a non-negotiable requirement for the proper use of field data by the research community. Field experimental data have often been used as the indisputable ground truth for subsequent theoretical and numerical analysis, without too much attention paid to the uncertainties and limitations that quantities measured in the field inevitably possess. This poses a serious risk of experimental errors propagating into physical models and calls for a closer collaboration between the field researchers that collect the data and their peers who use it. This is now truer than ever, as both experimental and modeling techniques in the wind energy research space are advancing at an unprecedented pace to investigate subtler and more complex physical phenomena.

To conclude, this document is not intended to replace the great body of literature on this topic, which is referenced throughout the report, and that the reader is welcome to refer to for more in-depth information. For the future, we hope that this report and the reference herein will be the cornerstone for high-quality analysis of ASSIST+TROPoe data from the AWAKEN, WFIP3, and other campaigns.

References

Abraham, A., M. Puccioni, A. Jordan, E. Maric, N. Bodini, N. Hamilton, S. Letizia, et al. 2024. “Operational wind plants increase planetary boundary layer height: An observational study.” *Wind Energy Science Discussions* 2024:1–34. https://doi.org/10.5194/wes-2024-148. https://wes.copernicus.org/preprints/wes-2024-148/.

Adler, B., J. M. Wilczak, L. Bianco, L. Bariteau, C. J. Cox, G. de Boer, I. V. Djalalova, et al. 2023. “Impact of Seasonal Snow-Cover Change on the Observed and Simulated State of the Atmospheric Boundary Layer in a High-Altitude Mountain Valley.” *Journal of Geophysical Research: Atmospheres* 128 (12). ISSN: 21698996. https://doi.org/10.1029/2023JD038497.

Bianco, L., B. Adler, L. Bariteau, I. V. Djalalova, T. Myers, S. Pezoa, D. D. Turner, and J. M. Wilczak. 2024. “Sensitivity of thermodynamic profiles retrieved from ground-based microwave and infrared observations to additional input data from active remote sensing instruments and numerical weather prediction models.” *Atmospheric Measurement Techniques* 17 (13): 3933–3948. ISSN: 18678548. https://doi.org/10.5194/amt-17-3933-2024.

Blumberg, W. G., D. D. Turner, U. Löhnert, and S. Castleberry. 2015. “Ground-based temperature and humidity profiling using spectral infrared and microwave observations. Part II: Actual retrieval performance in clear-sky and cloudy conditions.” *Journal of Applied Meteorology and Climatology* 54 (11): 2305–2319. ISSN: 15588432. https://doi.org/10.1175/JAMC-D-15-0005.1.

Blumberg, W. G., T. J. Wagner, D. D. Turner, and J. Correia. 2017. “Quantifying the accuracy and uncertainty of diurnal thermodynamic profiles and convection indices derived from the atmospheric emitted radiance interferometer.” *Journal of Applied Meteorology and Climatology* 56 (10): 2747–2766. ISSN: 15588432. https://doi.org/10.1175/JAMC-D-17-0036.1.

Bodini, N., A. Abraham, P. Doubrawa, S. Letizia, R. Thedin, N. Agarwal, B. Carmo, et al. 2024. “An international benchmark for wind plant wakes from the American WAKE ExperimeNt (AWAKEN).” *Journal of Physics: Conference Series* 2767 (9). ISSN: 17426596. https://doi.org/10.1088/1742-6596/2767/9/092034.

Bodini, N. 2024. *WFIP3 - RHOD site - NREL Scanning lidar (upgraded Galion SN:60) / Processed Data.* https://wdh.energy.gov/ds/wfip3/rhod.lidar.z01.a0. Accessed: 2025-04-22. Wind Data Hub for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. https://doi.org/10.21947/2228755.

Boer, G. de, A. White, R. Cifelli, J. Intrieri, M. Hughes, K. Mahoney, T. Meyers, et al. 2023. “Supporting Advancement in Weather and Water Prediction in the Upper Colorado River Basin The SPLASH Campaign.” *Bulletin of the American Meteorological Society* 104 (10): E1853–E1874. ISSN: 15200477. https://doi.org/10.1175/BAMS-D-22-0147.1.

Chatfield, C., and H. Xing. 2019. *The analysis of time series: an introduction with R.* Chapman / hall/CRC.

Clough, S. A., F. X. Kneizys, L. S. Rothman, and W. O. Gallery. 1981. “Atmospheric transmittance and radiance: FASCOD1B.” *Proc Soc Photo Opt Instrum Eng* 277:152–166.

Clough, S. A., and M. J. Iacono. 1995. “Line-by-line calculation of atmospheric fluxes and cooling rates: 2. Application to carbon dioxide, ozone, methane, nitrous oxide and the halocarbons.” *Journal of Geophysical Research* 100 (D8): 16519–16535. https://doi.org/10.1029/95JD01386.

Clough, S. A., M. J. Iacono, and J. Moncet. 1992. “Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor.” *Journal of Geophysical Research* 97:15761–15785.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown. 2005. “Atmospheric radiative transfer modeling: A summary of the AER codes.” *Journal of Quantitative Spectroscopy and Radiative Transfer* 91 (2): 233–244. ISSN: 00224073. https://doi.org/10.1016/j.jqsrt.2004.05.058.

Coniglio, M. C., J. Correia, P. T. Marsh, and F. Kong. 2013. “Verification of Convection-Allowing WRF Model Forecasts of the Planetary Boundary Layer Using Sounding Observations” [in en]. *Weather and Forecasting* 28, no. 3 (June): 842–862. ISSN: 0882-8156, 1520-0434, accessed January 28, 2024. https://doi.org/10.1175/WAF-D-12-00103.1. https://journals.ametsoc.org/doi/10.1175/WAF-D-12-00103.1.

Djalalova, I. V., D. D. Turner, L. Bianco, J. M. Wilczak, J. Duncan, B. Adler, and D. Gottas. 2022. “Improving thermodynamic profile retrievals from microwave radiometers by including radio acoustic sounding system (RASS) observations.” Published 31 January 2022; DOI: 10.5194/amt-15-521-2022. *Atmospheric Measurement Techniques* 15 (2): 521–537. https://doi.org/10.5194/amt-15-521-2022.

Duncan Jr., J. B., L. Bianco, B. Adler, T. Bell, I. V. Djalalova, L. Riihimaki, J. Sedlar, et al. 2022. “Evaluating convective planetary boundary layer height estimations resolved by both active and passive remote sensing instruments during the CHEESEHEAD19 field campaign” [in en]. *Atmos. Meas. Tech.* 15, no. 8 (April): 2479–2502. ISSN: 1867-8548, accessed August 1, 2023. https://doi.org/10.5194/amt-15-2479-2022. https://amt.copernicus.org/articles/15/2479/2022/.

Feltz, W. F., W. L. Smith, R. O. Knuteson, H. E. Revercomb, H. M. Woolf, and H. B. Howell. 1998. “Meteorological applications of temperature and water vapor retrievals from the ground-based Atmospheric Emitted Radiance Interferometer (AERI).” *Journal of Applied Meteorology* 37 (9): 857–875. ISSN: 08948763. https://doi.org/10.1175/1520-0450(1998)037<0857:MAOTAW>2.0.CO;2.

Gallery, W. O., and S. A. Clough. 1993. *FFTSCAN: a program for spectral smoothing using Fourier transforms.* Technical report. Phillips Laboratory.

Genest, J., and P. Tremblay. 1999. “General analytic solutions for the instrument line shape of non-uniformly illuminated, off-axis detectors.” *Optics InfoBase Conference Papers* 38 (25): 58–60. ISSN: 21622701. https://doi.org/10.1364/fts.1999.fwb3.

Goldberg, L. 2023. *Site B PNNL Surface Met Station / Reviewed Data.* https://wdh.energy.gov/ds/awaken/sb.assist.z01.00. Accessed: 2025-04-21. Wind Data Hub for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. https://doi.org/10.21947/1959700.

Gordon, I. E., L. S. Rothman, C. Hill, R. V. Kochanov, Y. Tan, P. F. Bernath, M. Birk, et al. 2017. “The HITRAN2016 molecular spectroscopic database.” *Journal of Quantitative Spectroscopy and Radiative Transfer* 203:3–69. https://doi.org/10.1016/j.jqsrt.2017.06.038.

Greenberg, M. 1998. *Advanced Engineering Mathematics.* Pearson Education. ISBN: 9788177585469. https://books.google.com/books?id=2xpmgza88tAC.

Hamilton, N. 2023. *AWAKEN Site A1 - NREL Ceilometer (Vaisala CL51) / Derived Data.* https://wdh.energy.gov/ds/awaken/sa1.ceil.z01.c0. Accessed: 2025-04-21. Wind Data Hub for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. https://doi.org/10.21947/2221789.

Hogg, R. V., J. W. McKean, A. T. Craig, et al. 2013. *Introduction to mathematical statistics.* Pearson Education India.

Holzworth, G. C. 1964. “Estimates of mean maximum mixing depths in the contiguous United States.” *Monthly Weather Review* 92 (5): 235–242.

Klein, P., T. A. Bonin, J. F. Newman, D. D. Turner, P. B. Chilson, C. E. Wainwright, W. G. Blumberg, et al. 2015. “LABLE: A multi-institutional, student-led, atmospheric boundary layer experiment.” *Bulletin of the American Meteorological Society* 96 (10): 1743–1764. ISSN: 00030007. https://doi.org/10.1175/BAMS-D-13-00267.1.

Knuteson, R. O., H. E. Revercomb, F. A. Best, N. C. Ciganovich, R. G. Dedecker, T. P. Dirkx, S. C. Ellington, et al. 2004a. “Atmospheric Emitted Radiance Interferometer. Part I: Instrument design.” *Journal of Atmospheric and Oceanic Technology* 21 (12): 1763–1776. ISSN: 07390572. https://doi.org/10.1175/JTECH-1662.1.

———. 2004b. “Atmospheric Emitted Radiance Interferometer. Part II: Instrument performance.” *Journal of Atmospheric and Oceanic Technology* 21 (12): 1777–1789. ISSN: 07390572. https://doi.org/10.1175/JTECH-1663.1.

Kosiba, K. A., A. W. Lyza, R. J. Trapp, E. N. Rasmussen, M. Parker, M. I. Biggerstaff, S. W. Nesbitt, et al. 2024. “The Propagation, Evolution, and Rotation in Linear Storms (PERiLS) Project.” *Bulletin of the American Meteorological Society,* 1768–1799. ISSN: 0003-0007. https://doi.org/10.1175/bams-d-22-0064.1.

Letizia, S. 2023. *Site B - NREL Thermodynamic profiler (Assist II-11) / Raw Data.* https://wdh.energy.gov/ds/awaken/sb.assist.z01.00. Accessed: 2025-04-21. Wind Data Hub for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. https://doi.org/10.21947/1865061.

———. 2024. *NREL ASSIST Rhode Island / Thermodynamic retrievals TROPoe.* https://wdh.energy.gov/ds/wfip3/rhod.assist.tropoe.z01.c0. Accessed: 2025-04-22. Wind Data Hub for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

Letizia, S., B. Adler, and D. D. Turner. 2024. *TROPoe_processor: wrapper for thermodynamic retrievals.* GitHub repository, accessed on November 25, 2024. https://github.com/StefanoWind/TROPoe_processor.

Maahn, M., D. D. Turner, U. Löhnert, D. J. Posselt, K. Ebell, G. G. Mace, and J. M. Comstock. 2020. "Optimal estimation retrievals and their uncertainties." *Bulletin of the American Meteorological Society* 101 (9): E1512–E1523. ISSN: 00030007. https://doi.org/10.1175/BAMS-D-19-0027.1.

Michaud-belleau, V., M. Gaudreau, J. Lacoursière, É. Boisvert, L. Ravelomanantsoa, D. D. Turner, and L. Rochette. 2025. "The Atmospheric Sounder Spectrometer by Infrared Spectral Technology (ASSIST): Instrument design and signal processing." *Atmospheric Measurement Techniques Discussions* 20 (February): 1–39. https://doi.org/10.5194/egusphere-2024-3617.

Mlawer, E. J., and D. D. Turner. 2016. "Spectral Radiation Measurements and Analysis in the ARM Program." *Meteorological Monographs* 57:14.1–14.17. ISSN: 0065-9401. https://doi.org/10.1175/amsmonographs-d-15-0027.1.

Mlawer, E. J., J. Mascio, D. D. Turner, V. H. Payne, C. J. Flynn, and R. Pincus. 2024. "A More Transparent Infrared Window." *Journal of Geophysical Research: Atmospheres* 129 (22). https://doi.org/10.1029/2024JD041366.

Moriarty, P., N. Bodini, and L. Cheung. 2023. "Overview of recent observations and simulations from the American WAKE experimeNt (AWAKEN) field campaign Overview of recent observations and simulations from the American WAKE experimeNt (AWAKEN) field campaign." *Journal of Physics: Conference Series* 2505 (012049). https://doi.org/10.1088/1742-6596/2505/1/012049.

Moriarty, P., N. Bodini, S. Letizia, A. Abraham, T. Ashley, K. B. Bärfuss, R. J. Barthelmie, A. Brewer, P. Brugger, T. Feuerle, et al. 2024. "Overview of preparation for the American Wake Experiment (AWAKEN)." *Journal of Renewable and Sustainable Energy* 16 (5).

Newsom, R. K., C. Sivaraman, T. R. Shippert, and L. D. Riihimaki. 2019. "Doppler Lidar Vertical Velocity Statistics Value-Added Product," no. April, https://www.osti.gov/biblio/1238068%0Ahttps://www.osti.gov/servlets/purl/1238068.

NOAA. 1976. *U.S. Standard Atmosphere.* Technical report. Washingth DC, US: National Oceanic and Atmopheric Administration.

NOAA Physical Sciences Laboratory. 2024. *WFIP3: Wind Forecast Improvement Project 3.* https://psl.noaa.gov/renewable_energy/wfip3/. Accessed: 2024-11-18.

Oppenheim, A. V. 1999. *Discrete-time signal processing.* Pearson Education India.

Pekour, M. 2024. *WFIP3 - RHOD Site - PNNL Surface Meteorological Suit / Raw Data.* https://wdh.energy.gov/ds/wfip3/rhod.met.z01.00. Accessed: 2025-04-22. Wind Data Hub for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. https://doi.org/10.21947/2352807.

Radünz, W., B. Souza Carmo, J. K. Lundquist, S. Letizia, A. Abraham, A. S. Wise, M. Sanchez Gomez, N. Hamilton, R. K. Rai, and P. S. Peixoto. 2024. "Influence of simple terrain on the spatial variability of a low-level jet and wind farm performance in the AWAKEN field campaign." *Wind Energy Science Discussions.*

Revercomb, H. E., D. D. Turner, D. C. Tobin, R. O. Knuteson, W. F. Feltz, J. Barnard, J. Bösenberg, et al. 2003. "The ARM program's water vapor intensive observation periods." *Bulletin of the American Meteorological Society* 84 (2): 217–236. ISSN: 00030007. https://doi.org/10.1175/BAMS-84-2-217.

Revercomb, H. E., H. Buijs, H. B. Howell, D. D. LaPorte, W. L. Smith, and L. A. Sromovsky. 1988. “Radiometric calibration of IR Fourier transform spectrometers: solution to a problem with the High-Resolution Interferometer Sounder.” *Applied Optics* 27 (15): 3210. ISSN: 0003-6935. https://doi.org/10.1364/ao.27.003210.

Rodgers, C. D. 2000. *Inverse methods for atmospheric sounding: theory and practice.* Vol. 2. World scientific.

Rothman, L. S., D. Jacquemart, A. Barbe, D. C. Benner, M. Birk, L. R. Brown, M. R. Carleer, et al. 2005. “The HITRAN 2004 molecular spectroscopic database.” *Journal of Quantitative Spectroscopy and Radiative Transfer* 96 (2 SPEC. ISS.): 139–204. ISSN: 00224073. https://doi.org/10.1016/j.jqsrt.2004.10.008.

Shaw, W. J., L. K. Berg, M. Debnath, G. Deskos, C. Draxl, V. P. Ghate, C. B. Hasager, et al. 2022. “Scientific challenges to characterizing the wind resource in the marine atmospheric boundary layer.” *Wind Energy Science* 7 (6): 2307–2334. ISSN: 23667451. https://doi.org/10.5194/wes-7-2307-2022.

Siegel, R. 1971. *Thermal Radiation Heat Transfer: Vol. 3: Radiation Transfer with Absorbing, Emitting, and Scattering Media.* Vol. 164. Scientific / Technical Information Division, National Aeronautics / Space Administration.

Smith, W. L., W. F. Feltz, R. O. Knuteson, H. E. Revercomb, H. M. Woolf, and H. B. Howell. 1999. “The retrieval of planetary boundary layer structure using ground-based infrared spectral radiance measurements.” *Journal of Atmospheric and Oceanic Technology* 16 (2-3): 323–333. ISSN: 07390572. https://doi.org/10.1175/1520-0426(1999)016<0323:tropbl>2.0.co;2.

Tobin, D. C., F. A. Best, P. D. Brown, S. A. Clough, R. G. Dedecker, R. G. Ellingson, R. K. Garcia, et al. 1999. “Downwelling spectral radiance observations at the SHEBA ice station: Water vapor continuum measurements from 17 to 26μm.” *Journal of Geophysical Research Atmospheres* 104 (D2): 2081–2092. ISSN: 01480227. https://doi.org/10.1029/1998JD200057.

Turner, D. D., W. F. Feltz, and R. A. Ferrare. 2000. “Continuous water vapor profiles from operational ground-based active and passive remote sensors.” *Bulletin of the American Meteorological Society* 81 (6): 1301–1317. ISSN: 00030007. https://doi.org/10.1175/1520-0477(2000)081<1301:CWBPFO>2.3.CO;2.

Turner, D. D., and U. Löhnert. 2014. “Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI).” *Journal of Applied Meteorology and Climatology* 53 (3): 752–771. ISSN: 15588432. https://doi.org/10.1175/JAMC-D-13-0126.1.

Turner, D. 2025. *TROPOe Docker Image.* https://hub.docker.com/r/davidturner53/tropoe. Accessed: 2025-04-25.

Turner, D. D., and W. G. Blumberg. 2019. “Improvements to the AERIoe thermodynamic profile retrieval algorithm.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12 (5): 1339–1354. ISSN: 21511535. https://doi.org/10.1109/JSTARS.2018.2874968.

Turner, D. D., R. O. Knuteson, H. E. Revercomb, C. Lo, and R. G. Dedecker. 2006. “Noise reduction of atmospheric emitted radiance interferometer (AERI) observations using principal component analysis.” *Journal of Atmospheric and Oceanic Technology* 23 (9): 1223–1238. ISSN: 07390572. https://doi.org/10.1175/JTECH1906.1.

Turner, D. D., and U. Lohnert. 2021. “Ground-based temperature and humidity profiling: Combining active and passive remote sensors.” *Atmospheric Measurement Techniques* 14 (4): 3033–3048. ISSN: 18678548. https://doi.org/10.5194/amt-14-3033-2021.

Turner, D. D., D. C. Tobin, S. A. Clough, P. D. Brown, R. G. Ellingson, E. J. Mlawer, R. O. Knuteson, et al. 2004. “The QME AERI LBLRTM: A closure experiment for downwelling high spectral resolution infrared radiance.” *Journal of the Atmospheric Sciences* 61 (22): 2657–2675. ISSN: 00224928. https://doi.org/10.1175/JAS3300.1.

Turner, D. D., A. M. Vogelmann, R. T. Austin, J. C. Barnard, K. Cady-Pereira, J. C. Chiu, S. A. Clough, et al. 2007. “Thin liquid water clouds: Their importance and our challenge.” *Bulletin of the American Meteorological Society* 88 (2): 177–190. ISSN: 00030007. https://doi.org/10.1175/BAMS-88-2-177.

Veers, P., K. Dykes, E. Lantz, S. Barth, C. L. Bottasso, O. Carlson, A. Clifton, et al. 2019. “Grand challenges in the science of wind energy.” *Science* 366 (6464): eaau2027. https://doi.org/10.1126/science.aau2027. https://www.science.org/doi/10.1126/science.aau2027.

Wagner, T. J., D. D. Turner, T. Heus, and W. G. Blumberg. 2022. “Observing Profiles of Derived Kinematic Field Quantities Using a Network of Profiling Sites.” *Journal of Atmospheric and Oceanic Technology* 39 (3): 335–351. ISSN: 15200426. https://doi.org/10.1175/JTECH-D-21-0061.1.

Appendix A. TROPoe Output Variables

Table A.1. List of TROPoe-0.12 0.11.7 Output Variables and Their Descriptions.

Variable name	Description	Unit
time	Time of each retrieval, corresponding to middle of interval	UTC
height	Retrieval height above ground level	km
qc_flag	Quality control flag, with 0 indicating acceptable quality	-
temperature	Air temperature	°C
waterVapor	Water vapor mixing ratio	$g\ kg^{-1}$
lwp	Liquid water path	$g\ m^{-2}$
lReff	Liquid water effective radius	μm
iTau	Ice cloud optical depth	-
iReff	Ice effective radius	μm
co2	Parameterized profile of CO_2	ppm
ch4	Parameterized profile of CH_4	ppm
n2o	Parameterized profile of N_2O	ppm
sigma_temperature	1-σ uncertainty in air temperature	°C
sigma_waterVapor	1-σ uncertainty in water vapor mixing ratio	$g\ kg^{-1}$
sigma_lwp	1-σ uncertainty in liquid water path	g/m^2
sigma_lReff	1-σ uncertainty in liquid water effective radius	μm
sigma_iTau	1-σ uncertainty in ice cloud optical depth	-
sigma_iReff	1-σ uncertainty in ice effective radius	μm
sigma_co2	1-σ uncertainty in carbon dioxide concentration	ppm
sigma_ch4	1-σ uncertainty in methane concentration	ppm
sigma_n2o	1-σ uncertainty in nitrous oxide concentration	ppm
converged_flag	Algorithm convergence flag, with 0 indicating no convergence	-
gamma	Underrelaxing factor (see $\hat{\gamma}$ in Section 4.2.2)	-
n_iter	Number of iterations performed	-
rmsr	Root-mean-square error between observation vector and forward model calculation, using only IRS radiance	-
rmsa	Root-mean-square error between observation vector and forward model calculation, using entire vector (see Eq. 4.24)	-
rmsp	Root-mean-square error between prior and retrieved profiles	-
chi2	Chi-squared statistic comparing observation vector and forward model calculation	-
dfs	Degrees of freedom (DFS), including total DFS, then DFS for each retrieved variable	-
sic	Shannon information content	-
vres_temperature	Vertical resolution of temperature profile	km
vres_waterVapor	Vertical resolution of water vapor profile	km
cdfs_temperature	Vertical profile of cumulative degrees of freedom for temperature	-
cdfs_waterVapor	Vertical profile of cumulative degrees of freedom for water vapor	-
hatchOpen	Flag indicating if ASSIST hatch was open (1) or closed (0)	-
cbh	Cloud base height above ground level	km
cbh_flag	Flag indicating source of cloud base height	-
pressure	Air pressure, estimated from hypsometric equation	mb
theta	Potential temperature, derived from temperature and pressure	K
thetae	Equivalent potential temperature, derived from temperature, pressure, and water vapor mixing ratio	K
rh	Relative humidity, derived from temperature and water vapor mixing ratio	%
dewpt	Dew point temperature, derived from temperature and water vapor mixing ratio	°C
pwv	Precipitable water vapor, derived from temperature and water vapor mixing ratio	cm
pblh	Planetary boundary layer height, computed as described in Section 5.4	km

sbih	Surface-based inversion height	km
sbim	Surface-based inversion magnitude	°C
sbLCL	Lifted condensation level for a surface-based parcel	km
sbCAPE	Convective available potential energy for a surface-based parcel	J kg^{-1}
sbCIN	Convective inhibition for a surface-based parcel	J kg^{-1}
mlLCL	Lifted condensation level for a mixed-layer parcel	km
mlCAPE	Convective available potential energy for a mixed-layer parcel	J kg^{-1}
mlCIN	Convective inhibition for a mixed-layer parcel	J kg^{-1}
sigma_pwv	1-σ uncertainty in precipitable water vapor	cm
sigma_pblh	1-σ uncertainty in planetary boundary layer height	km
sigma_sbih	1-σ in surface-based inversion height	km
sigma_sbim	1-σ in surface-based inversion magnitude	°C
sigma_sbLCL	1-σ in lifted condensation level for a surface-based parcel	km
sigma_sbCAPE	1-σ in convective available potential energy for a surface-based parcel	J kg^{-1}
sigma_sbCIN	1-σ in convective inhibition for a surface-based parcel	J kg^{-1}
sigma_mlLCL	1-σ in lifted condensation level for a mixed-layer parcel	km
sigma_mlCAPE	1-σ in convective available potential energy for a mixed-layer parcel	J kg^{-1}
sigma_mlCIN	1-σ in convective inhibition for a mixed-layer parcel	J kg^{-1}
obs_vector	Observation vector (see x in Section 4.1.3)	mixed
obs_vector_uncertainty	1-σ uncertainty in observation vector	mixed
forward_calc	Forward model calculation from state vector (i.e., $F(\hat{x})$)	Several
Xop	Optimal solution, including all retrieved variables	mixed
Sop	Posterior covariance (see $\hat{S}$ in Section 4.2.2)	mixed
Akernal	Averaging kernel (see $\hat{A}$ in Section 4.2.2)	mixed
Xa	Mean state of prior (see x_a in Section 4.1.3)	mixed
Sa	Covariance of prior (see S_a in Section 4.2.2)	mixed
lat	Latitude where station is located	°North
lon	Longitude where station is located	°East
alt	Height above mean sea level where station is located	m