

A Predictive Framework for Season-Dependent Quantum Key Distribution over Turbulent Free-Space Channels

Harshit Tiwari^a, Miguel Ángel Vázquez^a, and Satyendra Kumar Mishra^a

^aSpace and Resilient Communications and Systems (SRCOM), Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain

ABSTRACT

Free-space optical (FSO) quantum key distribution (QKD) over satellite-to-ground links is strongly influenced by altitude and temporal variations of atmospheric turbulence. Optical turbulence strength and temporal evolution along the communication path is defined by the refractive-index structure parameter, $C_n^2(z, t)$. Conventional approaches commonly represent C_n^2 using empirical or parametrized profiles, while recent data-driven studies have demonstrated the potential of machine learning to estimate or forecast turbulence from atmospheric observations. However, direct multi-step spatio-temporal forecasting of altitude-resolved C_n^2 from multivariate atmospheric conditions, and its subsequent impact on QKD performance, remains insufficiently explored.

This work proposes Cn2ConvTransformer, a convolutional neural network (CNN)–Transformer sequence-to-sequence neural network for forecasting the future $C_n^2(z, t)$ from wind speed, wind direction, temperature, and turbulence kinetic energy. The CNN component extracts local vertical structures, while the transformer captures longer-range temporal dependencies and enables multi-step prediction of the altitude-resolved turbulence profile. The proposed framework, trained on two years of CERRA reanalysis data over nine European ground stations, is evaluated across forecast lead time, altitude, and input-history length. The measured forecast error is propagated through a turbulence-aware FSO-QKD downlink model, combining a Hufnagel–Valley turbulence profile with a decoy-state BB84 link budget, to quantify its impact on quantum bit error rate (QBER) and secret key rate (SKR) across a representative range of satellite elevation angles.

Keywords: quantum key distribution, free-space optical communication, atmospheric turbulence, refractive-index structure parameter, spatio-temporal forecasting, convolutional neural network, Transformer, Hufnagel–Valley model, secret key rate

1. INTRODUCTION

Free-space optical (FSO) satellite links are strongly affected by atmospheric turbulence, which causes time-varying fluctuations in channel transmissivity – degrading classical high-capacity FSO links under severe weather¹ – and, consequently, in quantum key distribution (QKD) performance. Experimental results on a free-space QKD link showed that selecting high-transmission intervals reduced the quantum bit-error rate (QBER) from 5.51% to 4.30% and increased the final secret key rate (SKR) by 25.2% from the same raw data.² Theoretical models further relate turbulence to QKD performance under moderate and strong turbulence³ and for satellite links, where beam wander, beam broadening, and scintillation contribute significantly to key-rate degradation;⁴ hybrid FSO-fiber BB84-QKD architectures have likewise been shown to sustain viable performance across clear-air, haze, fog, and rain conditions.⁵ These results motivate predicting atmospheric turbulence before transmission rather than responding to channel fluctuations after they occur.

Turbulence is commonly characterized by the refractive-index structure parameter C_n^2 , which varies strongly with altitude.⁶ Measurements show significant turbulence throughout the troposphere and into the lower stratosphere, with reported profiles extending to 17.5 km,⁷ 30 km,⁸ and approximately 33 km.⁹ For satellite and

Correspondence: Harshit Tiwari, Space and Resilient Communications and Systems (SRCOM), Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain.

E-mail: htiwari@cttc.es

airborne QKD, where propagation follows a slant path through the atmosphere, a surface value of C_n^2 is therefore insufficient; the altitude-dependent profile along the propagation path is required.

The temporal evolution of C_n^2 has motivated forecasting at both the surface and profile levels. Surface turbulence has been forecast using recurrent neural networks,¹⁰ sequence-to-sequence LSTMs,¹¹ and mesoscale numerical weather prediction.¹² Vertical profiles have been estimated from concurrent meteorological measurements,¹³ reconstructed from reanalysis data,¹⁴ and forecast using autoregressive deep learning.¹⁵ Incorporating seasonal information enhances climatological reconstruction,¹⁶ although models trained under a single seasonal regime may lose accuracy when conditions change.¹⁷ In meteorology, multivariate spatiotemporal forecasting is typically widely used and while convolutional neural network (CNN)–Transformer architectures integrate spatial feature extraction with attention-based temporal modeling for multi-step atmospheric forecasting,¹⁸ convolutional-recurrent networks capture spatial and temporal relationships in gridded data.¹⁹ Neural-network techniques have also been applied to other stages of the satellite-to-ground optical link, such as spiral-phase-plate-integrated networks for high-fidelity orbital-angular-momentum-mode demultiplexing.²⁰

These developments establish a foundation for atmospheric forecasting to quantum link performance. Existing models map measured or simulated turbulence to QBER and SKR via log-normal and hybrid propagation models,^{2,3} while satellite studies include beam and pointing effects.⁴ Turbulence-induced crosstalk has also been used to estimate QBER and SKR in UAV-relay OAM-BB84 systems,²¹ but these approaches rely on current or past channel states rather than forecasts. Forecast-informed QKD has been explored in limited forms. A recurrent model using C_n^2 predicted turbulence on a 5.4 km link for advance link selection,²² and autoregressive channel forecasting has been used to adapt QKD source intensity, improving SKR by over 20%.²³ However, these methods do not combine altitude-resolved C_n^2 forecasting with multivariate spatio-temporal atmospheric modeling and explicit QBER/SKR prediction, motivating a framework that forecasts the vertical turbulence profile from atmospheric state variables and evaluates its impact on QKD performance.

This work proposes a data-driven framework for forecasting the altitude-resolved turbulence profile $C_n^2(z, t + \tau)$ from multivariate atmospheric state variables and assessing its implications for FSO-QKD. The study covers nine optical ground station (OGS) sites distributed across Europe, providing a geographically diverse basis for evaluating the robustness and generalization of the proposed approach. The Cn2ConvTransformer is a sequence-to-sequence CNN–Transformer architecture that combines CNN layers for local vertical structure with temporal attention for multi-horizon forecasting. The predicted C_n^2 profiles are subsequently propagated through a turbulence-aware FSO-QKD channel model to quantify their impact on QBER, SKR, and related link-performance parameters. The framework is evaluated across forecast lead time, altitude, input-history length, and seasonal conditions.

The main contributions of this work are:

1. Develop a multivariate, exogenous-input approach for forecasting altitude-resolved C_n^2 across nine European optical ground station sites, extending existing C_n^2 forecasting methods that primarily rely on autoregressive turbulence histories.
2. Introduce the Cn2ConvTransformer architecture for vertical turbulence forecasting, combining altitude-dependent convolutional features with temporal attention for multi-horizon prediction.
3. Evaluate forecasting performance across altitude, lead time, input-history length, and seasonal conditions, providing a multi-site assessment of model accuracy and generalization.
4. Quantify the impact of forecasted C_n^2 profiles on QBER, SKR, and other link-performance parameters by propagating the forecasts through an FSO-QKD channel model, extending the established turbulence-to-key-rate relationship to forecast-based channel assessment.

2. PHYSICAL AND QKD SYSTEM MODEL

2.1 Atmospheric Turbulence and Refractive-Index Structure

Atmospheric turbulence is characterized by the refractive-index structure parameter C_n^2 , which determines the strength of turbulence-induced atmospheric coherence time, beam wander, beam broadening, and scintillation

along the propagation path. For slant or vertical links, C_n^2 must therefore be represented as an altitude-dependent profile rather than by a single ground-level value. We model this dependence using the Hufnagel–Valley (HV) profile,²⁴

$$C_n^2(z) = 0.00594 \left(\frac{V_{\text{rms}}}{27} \right)^2 (10^{-5}z)^{10} \exp\left(-\frac{z}{1000}\right) + 2.7 \times 10^{-16} \exp\left(-\frac{z}{1500}\right) + A \exp\left(-\frac{z}{100}\right), \quad (1)$$

where z is the height above the ground station and V_{rms} is the RMS high-altitude wind speed. The first two terms describe the upper-atmospheric contribution, while A controls the near-ground turbulence strength. With V_{rms} fixed, A is the only adjustable parameter.

In this work, the C_n^2 field forecast by the Cn2ConvTransformer is used to construct the altitude-dependent turbulence profile required by the optical-channel model. Rather than fixing the HV ground-layer parameter A to the standard HV5/7 value, A is inferred from the Cn2ConvTransformer prediction at the corresponding ground level. Since Eq. (1) is linear in A , this inversion has a closed-form solution. Each forecasted C_n^2 value is therefore mapped to an HV profile, providing a direct and consistent conversion from the model prediction to the vertical turbulence profile used for subsequent channel and adaptive-optics analysis. The derivation of the time-varying $C_n^2(z, t)$ from the multivariate atmospheric fields is described in the following section.

2.2 Spatio-Temporal Representation of Atmospheric Turbulence

The refractive-index structure parameter $C_n^2(z, t)$ is derived for each OGS, vertical layer, and analysis time from the reanalysis fields described in Sec. 3. The calculation uses a local gradient-Richardson-number stability closure combining Tatarskii’s temperature-structure formulation²⁵ with the turbulent-Prandtl-number and hybrid-length-scale closure of Basu and Holtslag²⁶ and their revised structure-parameter relation.²⁷ In this work we used CERRA²⁸ dataset, provides four selected vertical levels; therefore, vertical derivatives are computed as layer-averaged finite differences between adjacent levels, consistent with the layer-centered target grid in Sec. 3. The turbulent kinetic energy used in the closure is taken from CERRA’s short-range forecast product and timestamp-matched to the thermodynamic and wind fields from the analysis product (Sec. 3). At each layer, potential temperature follows Poisson’s equation,

$$\theta(z, t) = T(z, t) \left(\frac{p_{\text{ref}}}{p(z, t)} \right)^\kappa, \quad \kappa = R_d/c_p, \quad (2)$$

and the layer-bulk potential-temperature gradient $\Gamma = \Delta\theta/\Delta z$ and wind shear $S = \sqrt{(\Delta u/\Delta z)^2 + (\Delta v/\Delta z)^2}$ – with (u, v) reconstructed from wind speed U and direction θ_{wind} – define the buoyancy parameter and gradient Richardson number

$$\beta = g/\theta, \quad N^2 = \beta \Gamma, \quad \text{Ri}_g = N^2/S^2, \quad (3)$$

where $\beta = g/\theta(z, t)$ uses the local potential temperature in place of a fixed reference θ_0 , a standard approximation that introduces $\lesssim 1\text{--}2\%$ error given the small fractional variation of θ across the levels used. Layers with Γ below the floor $\Gamma_{\text{min}} = 10^{-5} \text{ K m}^{-1}$ (near-neutral or convective stratification) or with S below $S_{\text{min}} = 10^{-4} \text{ s}^{-1}$ (guarding against the $\text{Ri}_g \rightarrow \infty$ singularity as $S \rightarrow 0$) fall outside the validity domain of this closure and are excluded via the validity mask of Sec. 3.

Given Ri_g , the turbulent Prandtl number Pr_t is taken as the larger root of

$$\text{Pr}_t^2 - [\text{Pr}_{t0} + \text{Ri}_g + (1 - a_p)c_P \text{Ri}_g] \text{Pr}_t + \text{Pr}_{t0} \text{Ri}_g = 0, \quad (4)$$

with neutral-limit coefficients $\text{Pr}_{t0} = 0.85$, $a_p = 0.33$, $c_P = 2.80$.²⁶ Assuming near-isotropic vertical velocity fluctuations, TKE supplies the vertical-velocity variance $\sigma_w^2 = (2/3) \text{TKE}(z, t)$; because stable stratification suppresses vertical relative to horizontal fluctuations, this isotropy assumption is a likely source of positive bias in σ_w , and hence in C_n^2 , under strong stability. σ_w combines with Ri_g and Pr_t into a hybrid turbulence length scale,

$$L_X = c_H \frac{\sigma_w}{S} \bigg/ \sqrt{\frac{1}{G(\text{Ri}_g)} - \frac{\text{Ri}_g}{\text{Pr}_t}}, \quad G(\text{Ri}_g) = \min(1, \text{Ri}_g^{-1}), \quad c_H = 0.8, \quad (5)$$

for which L_X is real-valued only while $1/G(\text{Ri}_g) > \text{Ri}_g/\text{Pr}_t$; for the coefficients above this condition, together with the discriminant of Eq. (4), is satisfied for all $\text{Ri}_g \geq 0$, so no additional runtime mask beyond the Γ and S floors above is required. This sets the potential-temperature structure parameter

$$C_\theta^2(z, t) = \frac{c}{\text{Pr}_t} L_X^{4/3} \Gamma^2, \quad c = 2.8 \text{ (range } [2.8, 3.2], \approx \pm 14\% \text{ on } C_n^2), \quad (6)$$

after Basu and Holtslag's revision²⁷ of Tatarskii's structure-function scaling, building on Wyngaard et al.²⁹ Because $\Gamma = \Delta\theta/\Delta z$ is a potential-, not actual-, temperature gradient, C_θ^2 is the potential-temperature structure parameter rather than the actual-temperature C_T^2 of the standard optical-turbulence literature.

Converting C_θ^2 to the optical refractive-index structure parameter uses $(p/p_{\text{ref}})^\kappa = T/\theta$ (Eq. (2)) to fold the potential-to-actual-temperature correction into the height-valid refractivity relation of Andrews and Phillips²⁴ and Cherubini and Businger,³⁰

$$C_n^2(z, t) = \left(\frac{A_\lambda p(z, t)}{T(z, t) \theta(z, t)} \right)^2 C_\theta^2(z, t), \quad p \text{ in hPa}, \quad (7)$$

with $A_\lambda = 77.6 \times 10^{-6} (1 + 7.52 \times 10^{-3} \lambda^{-2}) \text{ hPa}^{-1}$ (λ in μm), giving $A_\lambda \approx 7.7843 \times 10^{-5} \text{ hPa}^{-1}$ at the FSO/QKD link wavelength $\lambda = 1550 \text{ nm}$ (Sec. 2.3); this dispersion correction is small ($\sim 2\text{--}3\%$ relative to the visible-band coefficient) across the optical/near-infrared band. Equations (2)–(7) map the instantaneous atmospheric state $\mathbf{x}(z, t)$ of Eq. (11) onto the layer-resolved $C_n^2(z, t)$ field used as the forecasting target in Sec. 3 and propagated to the optical channel model of Sec. 2.3.

2.3 From C_n^2 to Link Transmittance and Secret-Key Rate

The Cn2ConvTransformer provides an altitude-resolved forecast $C_n^2(z_i, t + \tau)$ for each OGS and forecast horizon. To obtain a continuous turbulence profile for the propagation model, each forecasted value at height z_i is used to determine the HV ground-layer coefficient A in Eq. (1). Since the HV model is linear in A , the coefficient is obtained analytically by inverting the profile at z_i . The resulting $C_n^2(h)$ profile is then available at arbitrary heights along the slant path and is used as the turbulence input to the optical-channel model.

The integrated turbulence strength is characterized by the Rytov variance,²⁴

$$\sigma_R^2 = 2.25 k^{7/6} (\sec \zeta)^{11/6} \int_{h_{\text{OGS}}}^H C_n^2(h) (h - h_{\text{OGS}})^{5/6} dh, \quad (8)$$

where H is the satellite altitude, h_{OGS} is the ground-station altitude, ζ is the zenith angle, and $k = 2\pi/\lambda$ is the optical wavenumber. The resulting σ_R^2 is converted to the scintillation index σ_I^2 using the moderate-to-strong turbulence model of.²⁴ The corresponding irradiance statistics are then used to determine the turbulence-induced transmittance loss, η_{scint} .

The total channel transmittance is modeled as

$$\eta = \eta_{\text{geo}} \eta_{\text{fspl}} \eta_{\text{atm}} \eta_{\text{scint}} \eta_{\text{point}} \eta_{\text{misc}} G_{\text{tx}} G_{\text{rx}}, \quad (9)$$

where η_{geo} and η_{fspl} account for geometric and free-space propagation losses, η_{atm} represents clear-air atmospheric extinction, η_{point} accounts for pointing errors, and η_{misc} represents fixed optical and hardware losses. The aperture gains are denoted by G_{tx} and G_{rx} . Among these terms, η_{scint} is the component that varies with the forecasted turbulence state; the remaining terms are determined by the link geometry and system parameters.

The resulting η is then used to evaluate QKD performance. Following the weak-plus-vacuum decoy-state BB84 model of,³¹ the signal and decoy gains, Q_μ and Q_ν , and their corresponding QBERs, E_μ and E_ν , are determined from the channel transmittance, signal and decoy intensities μ and ν , detector dark-count probability Y_0 , and intrinsic detector error e_{det} . Decoy-state estimation provides bounds on the single-photon gain Q_1 , yield Y_1 , and error rate e_1 . The asymptotic secret-key rate per transmitted pulse is then

$$R = q [Q_1 (1 - H_2(e_1)) - f Q_\mu H_2(E_\mu)], \quad (10)$$

where q is the sifting efficiency, $f \geq 1$ is the error-correction inefficiency, and $H_2(\cdot)$ is the binary entropy function. The reported QBER is the signal-state error rate E_μ .

Equations (1)–(10) therefore define the forward model used in this work:

$$\widehat{C}_n^2(z_i, t + \tau) \rightarrow C_n^2(h) \rightarrow \sigma_R^2 \rightarrow \eta_{\text{scint}} \rightarrow \eta \rightarrow \{E_\mu, R\}.$$

This chain provides a direct mapping from the Cn2ConvTransformer forecast to the predicted QBER and secret-key rate evaluated in Sec. 5. The resulting QBER and SKR provide the performance metrics used to assess the forecasted turbulence. The dataset and forecasting problem are formulated in the subsequent sections.

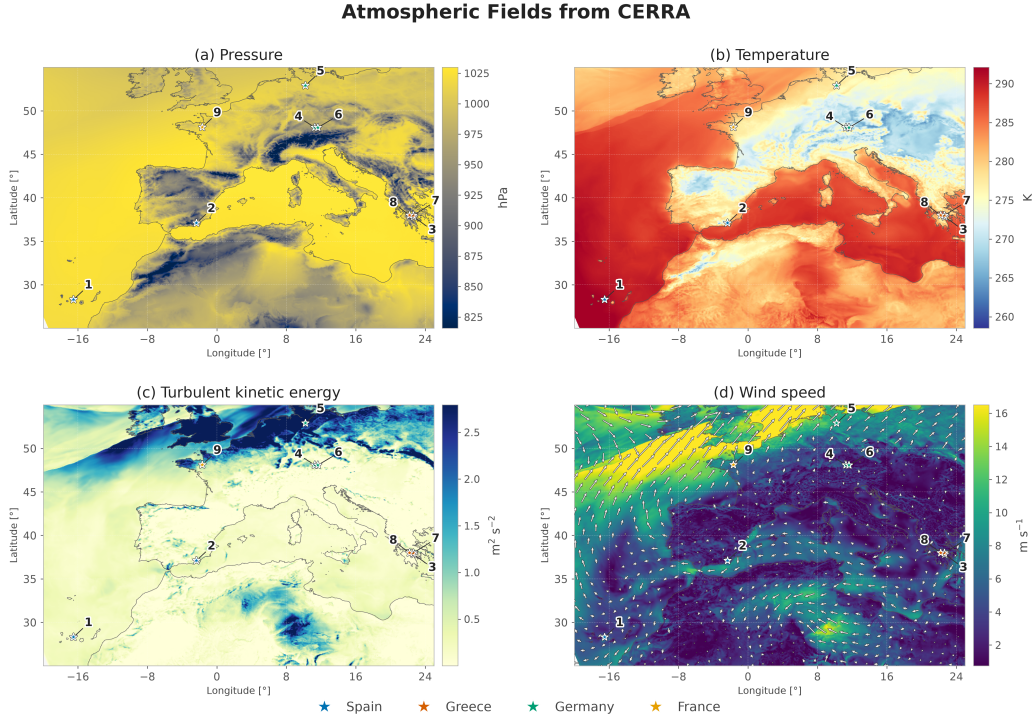


Figure 1. Representative CERRA input fields over the study domain: (a) surface pressure, (b) temperature, (c) turbulent kinetic energy, and (d) wind speed with direction vectors. Star markers mark the nine optical ground stations, colored by country and numbered to match Table 1.

3. DATASET AND FORECASTING PROBLEM FORMULATION

3.1 Atmospheric Dataset

The forecasting problem is built on the CERRA regional reanalysis,²⁸ obtained from the Copernicus Climate Data Store on CERRA’s native 5.5 km Lambert-conformal grid. Height-level fields are extracted at four levels above ground – 15, 100, 200, and 500 m – at the reanalysis’ native 3-h cadence (00, 03, $\dots$, 21 UTC), covering the two-year period 2024–2025, for a network of nine European optical ground stations (Table 1), each represented by a station-centered spatial patch on the nearest CERRA grid point. At every station and analysis time, the reanalysis supplies the vertically resolved thermodynamic and dynamical state of the atmosphere – temperature, pressure, and horizontal wind – together with a short-range-forecast estimate of turbulent kinetic energy (TKE). The refractive-index structure parameter C_n^2 is not a reanalysis output; it is instead *derived* from these fields through the local, gradient-Richardson-number-based stability closure of Sec. 2.2, following Tatarskii²⁵ and the TKE-based length-scale and structure-parameter revisions of Basu and Holtslag.^{26,27} The resulting dataset comprises of 52,380 data points (6000/station) atmospheric profiles.

Table 1. Region, country, and coordinates of the nine European optical ground stations used to construct the dataset.

No.	Region	Country	Latitude (°N)	Longitude (°)
1	Canary Islands (Tenerife)	Spain	28.2980	−16.5118
2	Andalusia	Spain	37.0951	−2.3561
3	Peloponnese (Corinthia)	Greece	37.8200	22.6600
4	Bavaria	Germany	48.0848	11.2780
5	Lower Saxony	Germany	52.9000	10.2000
6	Bavaria	Germany	48.0700	11.6300
7	Peloponnese (Corinthia)	Greece	37.9719	22.6186
8	Peloponnese (Achaia)	Greece	37.9844	22.1961
9	Brittany	France	48.1100	−1.6800

At each input height level z and time t , the model receives horizontal wind speed $U(z, t)$, temperature $T(z, t)$, wind direction $\theta(z, t)$, and turbulent kinetic energy $\text{TKE}(z, t)$. U and TKE characterize the mechanical generation of turbulence that mixes the temperature inhomogeneities underlying C_n^2 ; T fixes the thermodynamic background entering the potential-temperature gradient that scales turbulence intensity. Wind direction adds information on the advective regime and shear between levels; being circular, θ is encoded as $(\sin \theta, \cos \theta)$ to avoid the $0^\circ/360^\circ$ discontinuity, giving the per-level input vector

$$\mathbf{x}(z, t) = [U(z, t), T(z, t), \sin \theta(z, t), \cos \theta(z, t), \text{TKE}(z, t)]. \quad (11)$$

Figure 1 illustrates representative CERRA fields across the study domain, including surface pressure, temperature, turbulent kinetic energy, and wind speed with direction vectors at 15m above ground. The nine optical ground stations are marked by stars, colored by country, and numbered according to Table 1. These variables are not assumed to contribute equally, and their relative contribution to forecast skill is expected to vary with lead time and altitude.

3.2 Forecasting Problem Formulation and Data Preprocessing

The forecasting problem is formulated as supervised multi-step prediction of the altitude-resolved turbulence profile from a history of multivariate atmospheric states. At each analysis time, C_n^2 is derived over N_z vertical layers defined as bulk, cell-centered intervals between adjacent CERRA levels. With four input levels, this gives $N_z = 3$ layers centered at approximately 57.5, 150, and 350 m over the 15–500 m altitude range. The target vector is

$$\mathbf{y}(t) = [C_n^2(z_1, t), C_n^2(z_2, t), \dots, C_n^2(z_{N_z}, t)]. \quad (12)$$

Since C_n^2 spans several orders of magnitude across altitude and atmospheric regimes,²⁴ the model predicts $\log_{10} C_n^2$ rather than the raw values,

$$\mathbf{y}_{\log}(t) = \log_{10}(\mathbf{y}(t)). \quad (13)$$

This transformation reduces the dynamic range and prevents strongly turbulent samples from dominating the training loss.

Before sequence construction, the atmospheric inputs are temporally aligned to the exact CERRA analysis timestamp corresponding to each valid TKE forecast time. No temporal interpolation is applied. Each input feature is standardized using statistics computed from the training set only,

$$\tilde{x} = \frac{x - \mu_x}{\sigma_x}, \quad (14)$$

while TKE is first transformed as $\log(1 + \text{TKE})$ to reduce its positive skew before standardization.

At each time t , the input is an L -step history,

$$\mathbf{X}_{t-L+1:t} = \{\mathbf{X}_{t-L+1}, \dots, \mathbf{X}_t\}, \quad (15)$$

where $\mathbf{X}_t \in \mathbb{R}^{N_z^{\text{in}} \times N_f}$ contains the N_f atmospheric variables at the $N_z^{\text{in}} = 4$ input levels. The forecasting task is to predict the sequence of future turbulence profiles,

$$\hat{\mathbf{Y}}_{t+1:t+H} = F_{\Theta}(\mathbf{X}_{t-L+1:t}), \quad (16)$$

where Θ denotes the trainable parameters, $H = 6$ is the forecast horizon, corresponding to 18 h at the native 3-h temporal resolution, and each target frame satisfies $\mathbf{Y}_t \in \mathbb{R}^{N_z}$.

Three look-back windows, $L \in \{3, 6, 9\}$ steps, corresponding to 9, 18, and 27 h of atmospheric history, are considered. A separate model is trained for each look-back setting, allowing the effect of input-history length on forecast skill to be evaluated across the fixed 18-h forecast horizon. The complete forecasting problem can therefore be written as

$$\{U, T, \sin \theta, \cos \theta, \text{TKE}\}_{t-L+1:t} \longrightarrow \{C_n^2(z, t+1), \dots, C_n^2(z, t+H)\}. \quad (17)$$

The resulting formulation requires the model to learn both the vertical structure of turbulence and its temporal evolution from the preceding atmospheric state. Training, validation, and test samples are separated chronologically to preserve temporal causality and avoid information leakage. The proposed Cn2ConvTransformer sequence-to-sequence architecture is described in the next section.

4. PROPOSED SPATIO-TEMPORAL FORECASTING FRAMEWORK

4.1 Cn2ConvTransformer

The forecasting function F_{Θ} in Eq. (16) is implemented as a CNN–Transformer hybrid, termed Cn2ConvTransformer. The architecture explicitly separates spatial and temporal modeling, the CNN layers extract local spatial features from each atmospheric field, while a Transformer models temporal dependencies across the input history. A cross-attention decoder then maps the temporal representation to the forecast horizon, followed by a CNN decoder that reconstructs the altitude-resolved output fields. The overall architecture is

$$\hat{\mathbf{Y}}_{t+1:t+H} = \mathcal{D}(\mathcal{T}(\mathcal{E}(\mathbf{X}_{t-L+1:t}))), \quad (18)$$

where $\mathcal{E}$, $\mathcal{T}$, and $\mathcal{D}$ denote the CNN encoder, temporal Transformer, and decoder, respectively. The complete architecture is shown in Fig. 2. The encoder and CNN decoder use shared weights across input frames and forecast steps, respectively, while temporal interactions are handled by the Transformer and decoder cross-attention.

Each input frame $\mathbf{X}_t \in \mathbb{R}^{N_z^{\text{in}} \times P \times P \times N_f}$ is processed by the CNN encoder, with the vertical levels incorporated into the feature representation. The encoder comprises $K = 3$ downsampling blocks with channel widths $\{32, 64, 128\}$. Each block consists of a stride-1 convolution followed by a stride-2 convolution, batch normalization, and ReLU activation. The spatial resolution is reduced from $P \times P$ to $s = \frac{P}{2^K}$, followed by a 1×1 convolution that projects the feature maps to $E = 256$ channels. This produces $S = s^2$ spatial tokens per input frame. The encoder is applied independently to all L history frames using shared weights. Intermediate feature maps from the encoder are retained from the most recent input frame for U-Net skip connections in the decoder.

The encoded spatial tokens are augmented with learned spatial and temporal positional embeddings and processed by the temporal Transformer. The Transformer comprises six pre-normalization self-attention blocks with eight attention heads and an MLP dimension of $4E$. Self-attention is applied along the temporal dimension for each spatial token, allowing the model to capture the evolution of atmospheric conditions at each spatial location without mixing spatial information at this stage. The resulting representation contains $L \times S$ context tokens.

Forecast generation uses a non-autoregressive cross-attention decoder. For each forecast step and spatial location, a learned query token is generated, giving $H \times S$ queries in total. These queries attend to the encoded

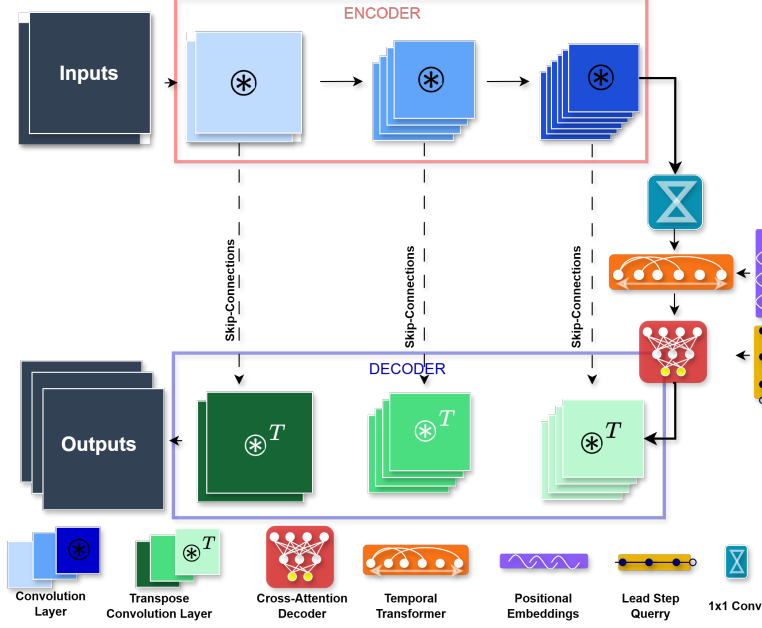


Figure 2. Cn2ConvTransformer architecture. Each input frame is independently encoded by a CNN, followed by temporal self-attention across the input history. A cross-attention decoder generates all forecast steps in parallel, and a mirrored CNN decoder reconstructs the altitude-resolved output using U-Net skip connections.

temporal memory through four pre-normalization cross-attention blocks, producing all H forecast steps simultaneously. This avoids recursive prediction and prevents forecast errors from being propagated between lead steps.

The decoded tokens are reshaped into H feature maps of size $s \times s \times E$ and passed through a CNN decoder that mirrors the encoder. The decoder contains $K = 3$ upsampling blocks with channel widths $\{128, 64, 32\}$, using transposed convolutions, batch normalization, and ReLU activation. At each scale, the corresponding encoder features are fused through U-Net skip connections. A final 3×3 convolution projects the reconstructed feature maps to the N_z turbulence layers, producing $\hat{\mathbf{Y}}_{t+1:t+H} \in \mathbb{R}^{H \times N_z \times P \times P}$.

4.2 Training Objective and Optimization

As the adopted C_n^2 closure is not defined under near-neutral or convective conditions (Sec. 2.2), the target contains invalid entries at the corresponding pixels. These entries are masked (refer Sec. 3.2) during the training, rather than being imputed. The primary loss is a masked element-wise Huber loss with $\delta = 1$ between the predicted and target $\log_{10} C_n^2$, computed only over valid pixels. To preserve spatial structure, a second Huber term is applied to the spatial gradients of the prediction and target along both horizontal axes. Gradient pairs are included only when both endpoints are valid. The total loss is $\mathcal{L} = \mathcal{L}_{\text{Huber}} + \lambda_{\text{grad}} \mathcal{L}_{\text{grad}}$ with $\lambda_{\text{grad}} = 0.5$.

The gradient term penalizes spatially smooth predictions that may achieve a low pointwise error while failing to reproduce local turbulence structure. Model parameters are optimized using AdamW with a weight decay of 10^{-4} and gradient-norm clipping at 1.0. The learning rate follows a linear warm-up for three epochs to a maximum of 3×10^{-4} , followed by cosine decay. Training uses a batch size of 128 pooled across all stations and is distributed across available TPU (v5e-8).

5. RESULTS

5.1 C_n^2 Forecasting Performance

We performed the ablation study on the held-out test set, the three input-history lengths yield similar overall forecasting accuracy in both $\log_{10} C_n^2$ and linear C_n^2 space (Table 2). Among them, $L = 6$ gives the lowest

$\log_{10}$ -RMSE (0.7237), followed by $L = 9$ (0.7273, 0.5% higher) and $L = 3$ (0.7328, 1.3% higher). The same ranking is obtained in linear space. Given the small differences between configurations, the results suggest that forecast accuracy is only weakly affected by input-history length over the tested range, $L \in \{3, 6, 9\}$.

Table 2. Overall test-set performance for the three history-length ablation configurations, in $\log_{10} C_n^2$ space and linear (physical) C_n^2 space.

Model ($L \times H$)	MAE ($\log_{10}$)	RMSE ($\log_{10}$)	MAE (linear)	RMSE (linear)
3×6	0.552	0.733	1.116×10^{-15}	3.168×10^{-15}
6×6	0.539	0.724	1.100×10^{-15}	3.114×10^{-15}
9×6	0.544	0.727	1.101×10^{-15}	3.315×10^{-15}

In contrast, forecast lead time has a clearer effect on accuracy. For the fixed $H = 6$ -step (18 h) horizon, $\log_{10}$ -RMSE increases steadily with lead time for all three history lengths (Fig. 3(a)). From the first lead step ($w = 1$, 3 h ahead) to the final step ($w = 6$, 18 h ahead), the error increases by 21.7–24.2%, depending on configuration. The gradual increase indicates a progressive loss of forecast skill rather than an abrupt degradation at longer lead times. Moreover, the differences between history lengths remain small compared with the lead-time effect. The analysis is restricted to the $L \in \{3, 6, 9\}$, $H = 6$ configurations evaluated on the held-out test set.

5.2 Spatial and Altitude-Resolved Forecasting Analysis

Furthermore, the vertical error profile is similarly consistent across the three configurations (Fig. 3(b)). Test-set $\log_{10}$ -RMSE increases slightly and monotonically with altitude, rising by approximately 1.6–3.3% from the lowest target layer (57.5 m) to the highest (350 m), depending on the history length. The relatively small variation indicates comparable forecasting skill across the three target layers, with only a mild loss of accuracy at higher altitudes. This analysis resolves errors vertically across the target layers; a horizontal, cross-station spatial analysis is not included here.

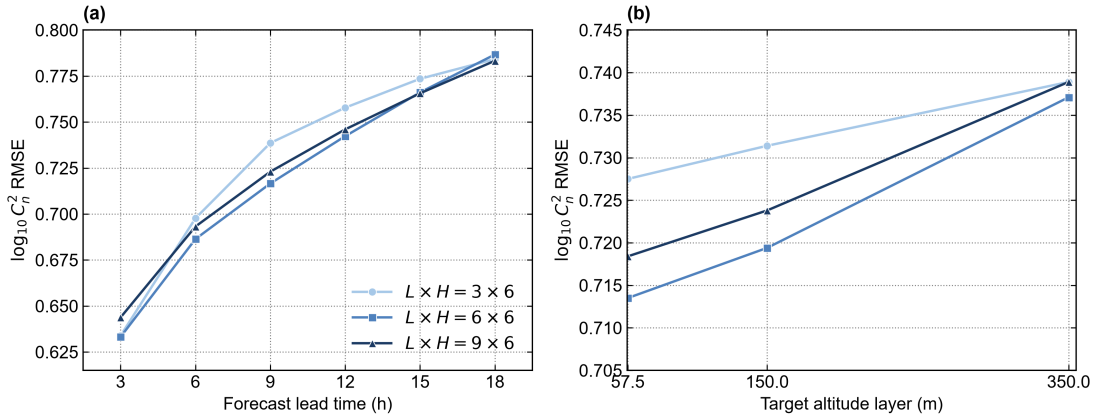


Figure 3. $\log_{10} C_n^2$ RMSE for the three history-length configurations, resolved (a) by forecast lead time (3 h per step, up to 18 h), and (b) by target altitude layer (57.5, 150, and 350 m). Colors and markers as in (a).

Fig. 4 illustrates this qualitatively for the 6×6 configuration on two representative test-set forecasts, the predicted $\log_{10} C_n^2$ maps closely reproduce the spatial pattern of the true turbulent structures at all three target levels, with the pointwise error concentrated along the edges and finer-scale features of these structures rather than spread uniformly across the domain. The fraction of the domain with a physically-defined (unmasked) C_n^2 also grows visibly from the lowest to the highest target level, consistent with the near-neutral/convective-stratification masking of Sec. 2.2.

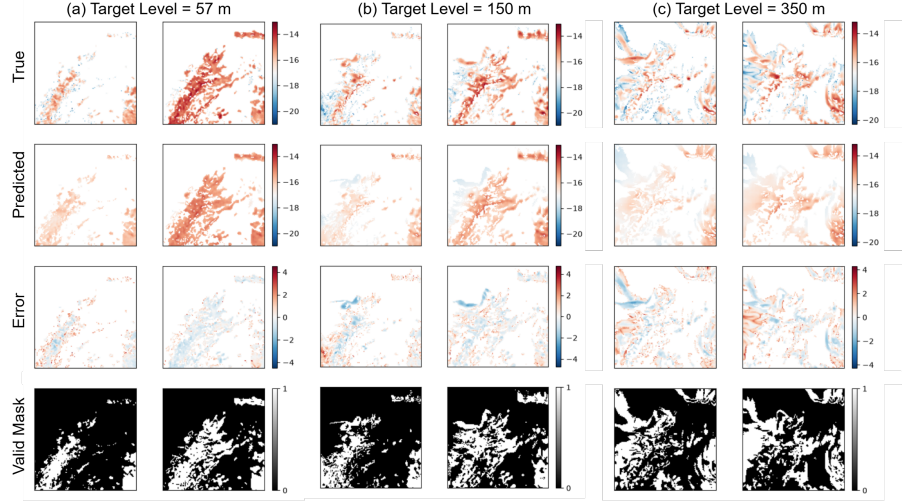


Figure 4. True and 6×6 -model-predicted $\log_{10} C_n^2$ maps on two representative test-set forecast windows (the two columns within each group), for the three target altitude levels: (a) 57.5 m, (b) 150 m, (c) 350 m. Within each group, rows show, top to bottom: the true $\log_{10} C_n^2$; the predicted $\log_{10} C_n^2$; the pointwise error (true – predicted, in $\log_{10}$ units); and the per-pixel validity mask (white: valid, physically-defined pixels included in training and evaluation; black: pixels masked out under near-neutral or convective stratification, Sec. 2.2).

6. DISCUSSION

The results in Sec. 5 indicate that the proposed Cn2ConvTransformer provides consistent, although not yet high-precision, forecasting performance across the tested configurations. The three history lengths differ by only 1.3% in overall $\log_{10}$ -RMSE (Table 2), with $L = 6$ performing marginally better than $L = 9$ and $L = 3$. This weak dependence on history length suggests that most of the predictive information in the 3-h CERRA sequence is captured within the preceding 9 h. In contrast, forecast lead time has a clearer effect, the $\log_{10}$ -RMSE increases by 21.7–24.2% between the 3- and 18-h leads, while the vertical error changes by only 1.6–3.3% between 57.5 and 350 m (Fig. 3). Thus, within the evaluated range, forecast skill is more sensitive to lead time than to either input-history length or target altitude. This stability across configurations is encouraging for practical use, since extending the input history provides little additional complexity benefit while maintaining comparable performance across the target column.

The current turbulence-to-channel mapping also introduces an important modeling assumption. The HV profile in Eq. (1) is anchored to a single forecasted C_n^2 layer through the closed-form inversion of A , while the remaining vertical structure follows the fixed HV terms. Although this provides a direct and computationally simple mapping from the forecast to a continuous turbulence profile, it does not independently constrain the profile at all three predicted layers. Jointly fitting the HV profile to the complete forecasted C_n^2 column would therefore provide a more direct use of the altitude-resolved prediction and is a natural extension of the present framework.

The QKD results further show that the effect of the forecast should be interpreted together with deterministic link geometry. Within the tested pass geometry, elevation is the dominant factor controlling SKR and QBER (Fig. 5(a)). SKR increases from 6.66×10^{-4} to 2.24×10^{-3} bits/pulse, while QBER decreases from 1.094% to 1.029% as elevation increases from 32.3° to 90° . At a fixed elevation, the station-to-station variation is much smaller, with relative deviations of at most approximately 0.1 percentage points. This distinction is operationally important, the elevation can be predicted from orbital geometry, whereas station-dependent turbulence must be estimated or forecast. Nevertheless, the elevation-averaged distributions in Fig. 5(b)–(c) show that station-dependent turbulence is not negligible in all cases. For example, station 4 exhibits a broader and lower-tailed SKR distribution, indicating that station-specific turbulence statistics may provide useful secondary information for pass selection and scheduling.

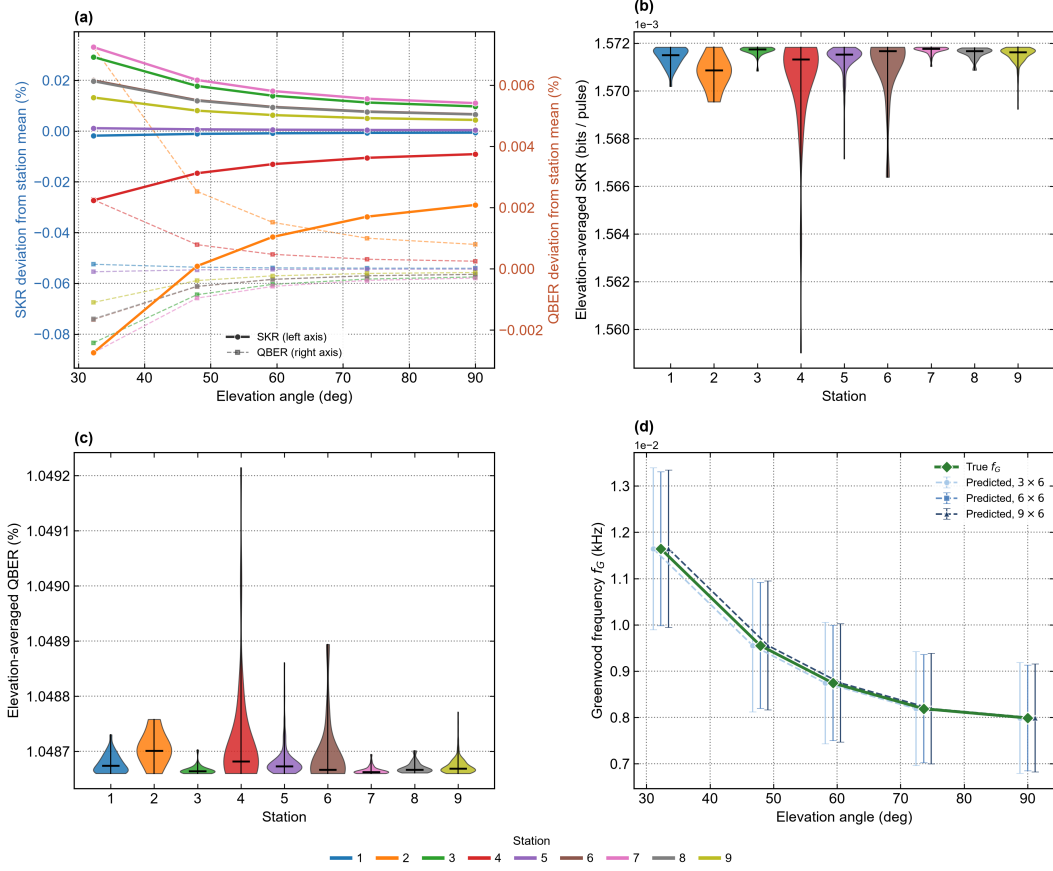


Figure 5. QKD link performance and adaptive-optics (AO) implications of the CERRA-derived test-set C_n^2 for the nine OGSs, numbered as in Table 1, over elevation angles of 32.3°–90°. (a) True SKR and QBER versus elevation, expressed as percentage deviation from the cross-station mean at each elevation (solid: SKR, left axis; dashed: QBER, right axis). (b)–(c) Per-station distributions of true SKR and QBER, respectively, averaged over the five tested elevations (violins: kernel density over test-set windows and forecast lead steps; horizontal bars: medians; whiskers: minima and maxima). (d) True and forecast-derived Greenwood frequency f_G versus elevation for each history-length configuration. Forecast values are obtained by perturbing the true test-set C_n^2 using the measured $\log_{10} C_n^2$ RMSE and the same Monte Carlo procedure used for the SKR/QBER analysis; markers show the median across lead steps and Monte Carlo draws, with error bars denoting the median per-draw RMSE. The three configurations are offset slightly in elevation for visibility.

The same turbulence forecasts can also be used to assess adaptive-optics requirements. Figure 5(d) compares the true Greenwood frequency f_G with the value obtained from the forecasted C_n^2 profiles. The predicted median follows the elevation-dependent trend of the true f_G for all three history lengths, while the forecast-induced relative error remains approximately 14–15% across elevation. This relatively stable error is consistent with the power-law dependence of f_G on the integrated turbulence strength and indicates that the uncertainty in the forecast propagates in a predictable manner to the AO bandwidth requirement. The forecasts show a small positive median bias of approximately 2.5%, meaning that they tend to slightly overestimate rather than underestimate the required correction bandwidth. Such an estimate could therefore support pre-pass selection of AO loop parameters, although it should not be interpreted as a replacement for real-time wavefront sensing and control.

Finally, the station-level comparison should be interpreted with the unequal number of test windows contributed by the nine OGS sites in mind. The present results therefore characterize the aggregate behavior of the available test samples rather than a spatially balanced comparison of stations. A balanced station-wise evaluation would provide a stronger assessment of geographical generalization and is an important direction for

extending the present analysis.

7. CONCLUSION AND FUTURE WORK

This work presented Cn2ConvTransformer, a CNN–Transformer sequence-to-sequence framework for forecasting altitude-resolved C_n^2 from multivariate CERRA reanalysis fields. The model showed consistent performance across the tested history lengths, with gradual degradation over the 18-h forecast horizon and only weak variation with altitude. The corresponding test-set C_n^2 fields were then evaluated through an HV-based FSO-QKD channel model, showing that, over the tested downlink geometry, elevation has a substantially larger effect on SKR and QBER than station-to-station turbulence variability, although station-specific turbulence remains relevant for some sites. Future work will extend the analysis to stronger-turbulence and season-dependent regimes and evaluate the complete forecast-to-QKD pipeline using live Cn2ConvTransformer predictions rather than the measured test-set C_n^2 .

ACKNOWLEDGMENTS

This research is part of SOFIA PID2023-147305OB-C32, funded by MICIU/AEI/10.13039/501100011033 and FEDER/UE.

REFERENCES

- [1] Kumari, M., Malhotra, J., and Mishra, S. K., “High-capacity 16 x 10 gbps quad lp modal mdm system using an integrated mmf-fso link under severe climate scenarios,” *Photonics* **13**(4), 392 (2026).
- [2] Erven, C., Heim, B., Meyer-Scott, E., Bourgoin, J. P., Laflamme, R., Weihs, G., and Jennewein, T., “Studying free-space transmission statistics and improving free-space quantum key distribution in the turbulent atmosphere,” *New Journal of Physics* **14**, 123018 (2012).
- [3] Ghalaii, M. and Pirandola, S., “Quantum communications in a moderate-to-strong turbulent space,” *Communications Physics* **5** (2022).
- [4] Liorni, C., Kampermann, H., and Bruß, D., “Satellite-based links for quantum key distribution: beam effects and weather dependence,” *New Journal of Physics* **21**, 093055 (2019). arXiv:1906.08115.
- [5] Kumari, M. and Mishra, S. K., “Simulation investigation of quantum FSO–fiber system using the BB84 QKD protocol under severe weather conditions,” *Photonics* **12**(7), 712 (2025).
- [6] Tatarski, V. I., [*Wave Propagation in a Turbulent Medium*], McGraw-Hill, New York (1961). Translated by R. A. Silverman.
- [7] Cuevas, O., Marín, J. C., Blázquez, J., and Meyer, C., “Combining cn2 models to forecast the optical turbulence at paranal,” *Monthly Notices of the Royal Astronomical Society* **529**(3), 2208–2219 (2024).
- [8] Zhou, Q., Liu, Y., Song, T., Sha, F., Liu, Q., Li, H., Zhang, X., Huang, W., Lu, L., Wang, X., Luo, Q., Chen, Y., and Zhang, H., “Atmospheric turbulence parameters characterization and seeing forecasting at the daocheng site using era5 reanalysis data,” *Monthly Notices of the Royal Astronomical Society* **545**(2), staf2103 (2025).
- [9] Ko, H.-C., Chun, H.-Y., Wilson, R., and Geller, M. A., “Characteristics of atmospheric turbulence retrieved from high vertical-resolution radiosonde data in the united states,” *Journal of Geophysical Research: Atmospheres* **124**(14), 7553–7579 (2019).
- [10] Grose, M. G. and Watson, E. A., “Forecasting atmospheric turbulence conditions from prior environmental parameters using artificial neural networks,” *Applied Optics* **62**(13), 3370–3379 (2023).
- [11] Li, Y., Zhang, H., Li, L., Shi, L., Huang, Y., and Fu, S., “Multistep ahead atmospheric optical turbulence forecasting for free-space optical communication using empirical mode decomposition and lstm-based sequence-to-sequence learning,” *Frontiers in Physics* **11**, 1070762 (2023).
- [12] Yang, Q., Wu, X., Luo, T., Qing, C., Yuan, R., Su, C., Xu, C., Wu, Y., Ma, X., and Wang, Z., “Forecasting surface-layer optical turbulence above the tibetan plateau using the wrf model,” *Optics & Laser Technology* **153**, 108217 (2022).

- [13] Su, C., Wu, X., Wu, S., Yang, Q., Han, Y., Qing, C., Luo, T., and Liu, Y., “In situ measurements and neural network analysis of the profiles of optical turbulence over the tibetan plateau,” *Monthly Notices of the Royal Astronomical Society* **506**(3), 3430–3441 (2021).
- [14] Pierzyna, M., Basu, S., and Saathof, R., “Otprof: Estimating high-resolution profiles of optical turbulence (cn2) from reanalysis using deep learning,” (2026). arXiv:2604.09346.
- [15] Hou, M., Gong, S., Li, X., Xiao, D., Zuo, Y., and Liu, Y., “Forecasting the atmospheric refractive index structure constant profile with an altitude-time correlations-inspired deep learning model,” *Optics Express* **31**(2), 2426–2444 (2023).
- [16] Pierzyna, M., Basu, S., and Saathof, R., “Otcim: Generating a near-surface climatology of optical turbulence strength (cn2) using gradient boosting,” *Artificial Intelligence for the Earth Systems* **4**(2) (2025).
- [17] Celik, U., Yasar, H. A., Yigitoglu Keskin, M., Bayar, C., Aslantas, I., and Midilli, Y., “Estimation of ground-based atmospheric turbulence strength (cn2) by neural network architecture,” *Applied Optics* **63**(28), 7402–7409 (2024).
- [18] Wu, R., Liang, Y., Lin, L., and Zhang, Z., “Spatiotemporal multivariate weather prediction network based on cnn-transformer,” *Sensors* **24**(23), 7837 (2024).
- [19] Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-k., and Woo, W.-c., “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in [*Advances in Neural Information Processing Systems 28 (NeurIPS)*], (2015). arXiv:1506.04214.
- [20] Tiwari, H., Vázquez, M. Á., and Mishra, S. K., “SPPINN: Spiral phase plate integrated neural network for high-fidelity OAM demultiplexing in satellite-to-ground optical links,” in [*2025 IEEE International Conference on Space Optical Systems and Applications (ICSOS)*], 1–8 (2025).
- [21] Deng, L., “Uav-deployed oam-bb84 qkd: Turbulence- and misalignment-resilient decoy-state finite-key security with ai-assisted calibration,” (2026). arXiv:2601.11117.
- [22] Jaouni, T., Scarfe, L., Bouchard, F., Krenn, M., Heshami, K., Di Colandrea, F., and Karimi, E., “Predicting atmospheric turbulence for secure quantum communications in free space,” *Optics Express* **33**(5) (2025). arXiv:2406.14768 (submitted 2024).
- [23] Sun, X., Djordjevic, I. B., and Neifeld, M. A., “Secret key rates and optimization of bb84 and decoy state protocols over time-varying free-space optical channels,” *IEEE Photonics Journal* **8**(3) (2016).
- [24] Andrews, L. C. and Phillips, R. L., “Laser beam propagation through random media,” SPIE (2005).
- [25] Tatarskii, V. I., “The effects of the turbulent atmosphere on wave propagation,” *Jerusalem: Israel Program for Scientific Translations, 1971* (1971).
- [26] Basu, S. and Holtslag, A. A. M., “Turbulent prandtl number and characteristic length scales in stably stratified flows: steady-state analytical solutions.” arXiv:2104.14378 (2021).
- [27] Basu, S. and Holtslag, A. A., “Revisiting and revising tatarskii’s formulation for the temperature structure parameter (ct 2) in atmospheric flows,” *Environmental Fluid Mechanics* **22**(5), 1107–1119 (2022).
- [28] Schimanke, S., Ridal, M., Le Moigne, P., Berggren, L., Undén, P., Randriamampianina, R., Andrea, U., Bazile, E., Bertelsen, A., Brousseau, P., Dahlgren, P., Edvinsson, L., El Said, A., Glington, M., Hopsch, S., Isaksson, L., Mladek, R., Olsson, E., Verrelle, A., and Wang, Z. Q., “Cerra sub-daily regional reanalysis data for europe on height levels from 1984 to present.” Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (2021). Accessed: 10-18-2026.
- [29] Wyngaard, J. C., Izumi, Y., and Collins, S. A., “Behavior of the refractive-index-structure parameter near the ground,” *Journal of the Optical Society of America* **61**(12), 1646–1650 (1971).
- [30] Cherubini, T. and Businger, S., “Another look at the refractive index structure function,” *Journal of Applied Meteorology and Climatology* **52**(2), 498–506 (2013).
- [31] Ntanos, A., Lyras, N. K., Zavitsanos, D., Giannoulis, G., Panagopoulos, A. D., and Avramopoulos, H., “Leo satellites constellation-to-ground qkd links: Greek quantum communication infrastructure paradigm,” *Photonics* **8**(12), 544 (2021).