

Analyzing Real-time Video Delivery over Cellular Networks for Remote Piloting Aerial Vehicles

Aygün Baltaci
ayguen.baltaci@tum.de
Technical University of Munich &
Airbus
Germany

Fabien Geyer
fabien.geyer@airbus.com
Airbus
Germany

Hendrik Cech
hendrik.cech@tum.de
Technical University of Munich
Germany

Vaibhav Bajpai
bajpai@cispa.de
CISPA Helmholtz Center for
Information Security
Germany

Dominic Schupke
dominic.schupke@airbus.com
Airbus
Germany

Nitinder Mohan
mohan@in.tum.de
Technical University of Munich
Germany

Jörg Ott
ott@in.tum.de
Technical University of Munich
Germany

ABSTRACT

Emerging Remote Piloting (RP) operations of electrified Unmanned Aerial Vehicles (UAVs) demand low-latency and high-quality video delivery to conduct safe operations in the low-altitude airspace. Although cellular networks are one of the prominent candidates to provide connectivity for such operations, their ground-centric nature limits their capabilities in achieving seamless and reliable aerial connectivity. In this paper, we study the feasibility of supporting RP operations with low latency and high-quality video delivery over commercial cellular networks. By setting up an adaptive bitrate video transmission pipeline with the Google Congestion Control (GCC) and Self-Clocked Rate Adaptation for Multimedia (SCReAM) Congestion Control (CC) algorithms, we analyze the video delivery performance for the RP application requirements and compare the performance of GCC and SCReAM against constant bitrate video delivery. Our results show that low-latency video delivery with < 300 ms playback latency between full-HD and 4K resolution can be maintained up to about 95% of the time in the air. While static bitrate video delivery outperforms adaptive streaming in urban location with abundant link capacity, the latter becomes advantageous in rural locations, where the link capacity is affected by fluctuations. Although the study's findings highlight the capabilities of cellular networks in delivering low-latency video for a safety-critical aerial service, we also discuss the potential improvements and future research challenges for enabling safe operations

and meeting the service requirements using cellular networks. We release our collected traces and the video transmission pipeline as open-source to facilitate research in this field.

CCS CONCEPTS

• **Networks** → *Network performance evaluation; Network performance analysis.*

KEYWORDS

Adaptive streaming; real-time video; cellular networks; LTE; UDP; RTP; UAV; drone; eVTOL; flying taxi; UAM; AAM

ACM Reference Format:

Aygün Baltaci, Hendrik Cech, Nitinder Mohan, Fabien Geyer, Vaibhav Bajpai, Jörg Ott, and Dominic Schupke. 2022. Analyzing Real-time Video Delivery over Cellular Networks for Remote Piloting Aerial Vehicles. In *Proceedings of the 22nd ACM Internet Measurement Conference (IMC '22)*, October 25–27, 2022, Nice, France. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3517745.3561465>

1 INTRODUCTION

The aviation industry is on the verge of a new era of environmental sustainability with emission-free operations. Novel Aerial Vehicle (AV) concepts, electric Vertical Take-off and Landing (eVTOL) and Unmanned Aerial Vehicles (UAVs), will occupy the low-altitude airspace with use cases such as passenger transportation, aerial package delivery, medical aid and remote surveillance. The majority of these AVs are planned to be operated remotely, relying on network connectivity to transmit live video streams from on-board cameras and receive control commands from a remote operator (see Figure 1). The airspace regulators, therefore, face new challenges in ensuring safe and reliable operations in the sky [12]. The AVs demand reliable and robust connectivity to safely perform Remote Piloting (RP) operations, obtain flight paths, dynamically update no-fly zones, and obtain other flight-relevant information from the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IMC '22, October 25–27, 2022, Nice, France

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9259-4/22/10...\$15.00

<https://doi.org/10.1145/3517745.3561465>

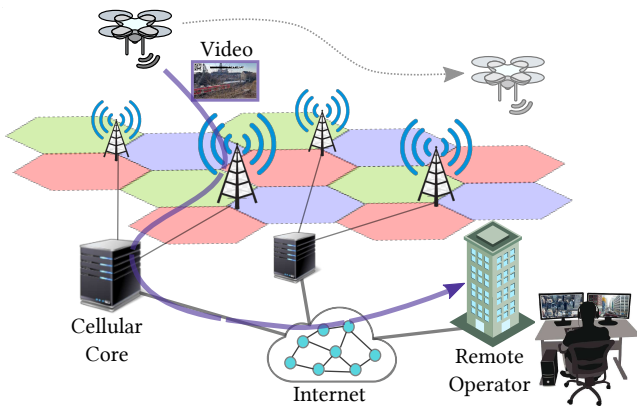


Figure 1: The Remote Piloting (RP) over cellular connectivity scenario considered in this study. The remote pilot connects to the Unmanned Aerial Vehicle (UAV) via a cellular network. Only the link between the UAV and the Base Station (BS) is wireless, while the rest is over the WAN.

unmanned traffic management systems. Simultaneously, the AVs also need to continuously broadcast their location to other AVs for avoiding conflicts in flight paths [42].

Recently, both industry and academia have started exploring the potential of utilizing cellular networks to provide wireless connectivity for aerial use cases, primarily due to already-existing network infrastructure and low-cost communications [52]. Although the current cellular standards and implementations are intrinsically designed to support connectivity on the ground, the 3rd Generation Partnership Project (3GPP) intends to incorporate non-terrestrial networks as part of the future 5G releases [3]. The evolution of the cellular networks takes the challenging connectivity requirements and low latency video delivery demands of RP applications into account to ensure reliable aerial coverage [2].

However, the 3D mobility patterns combined with the fast speeds of AVs introduce unprecedented challenges for seamless connectivity over cellular links. Although several studies have investigated the performance of cellular networks in the air in the past [25, 27, 29, 55], their focus has remained primarily on understanding the changes in the wireless medium, while their impact on the performance of AV applications has remained largely unexplored. Not only does this research lapse impair understanding of the potential limitations of current state-of-the-art cellular infrastructure in supporting aerial applications, but it also limits research in designing solutions specifically for supporting the operational requirements of such applications, e.g. using multipath transport [7].

In this paper, we plug this gap in research by evaluating the performance of low-latency, high-quality video transmission in the air (which is integral to RP operations) in cellular networks. We aim to assess the feasibility of already-existing cellular infrastructure for supporting high-quality video transmission for RP AVs (as illustrated in Figure 1) and uncover the arising networking challenges for future cellular generations. Specifically, the objective of our study is to answer the following research questions: (i) How do the network conditions in the air differ from on the ground? Can we characterize the differences between the two in terms of cellular

Quality of Service (QoS) metrics, such as Handover (HO) frequency, Handover Execution Time (HET), throughput, end-to-end latency, and Packet Error Rate (PER)? (ii) Can the currently available cellular networks support the low-latency, high-quality video transmission requirements of RP operations in aviation? What is the achievable video streaming performance, and what are the bottlenecks in supporting low-latency video transmission? We investigate these questions through the following contributions.

(1) We conduct an extensive measurement campaign in Munich, Germany, by flying an UAV in urban (Munich city center) and rural (Munich outskirts) environments. The UAV carried a payload of Raspberry Pi 4s (RPis) and Intel NUC PCs that continuously transferred high-quality Real-time Transport Protocol (RTP) video over an LTE network to a remote server hosted within the AWS cloud. In this study, we focused our attention on LTE’s aerial performance since 5G roll-out is in its nascent stages over the majority of the globe, and the current 5G (non-standalone) infrastructure predominantly relies on existing LTE base stations for last-mile connectivity [41]. In addition to the video streaming performance, we also record the low-layer LTE access information (e.g. signal strength and base station ID). This allows us to correlate the frequency of HOs in the air and on the ground to the application performance.

(2) We measure the performance of both non-adaptive and adaptive RTP video streaming workloads in our experiments to achieve low-latency, high-quality video transmission requirements that are integral to remote UAV operations. For adaptive streaming, we utilize two novel Congestion Control (CC) mechanisms with RTP, namely Google Congestion Control (GCC) and Self-Clocked Rate Adaptation for Multimedia (SCReAM) [24], primarily since these CCs are designed to support mobile applications over cellular networks [35] and real-time video traffic [14]. We record several video performance metrics such as the achievable bitrate, Frames Per Second (FPS), playback latency and received frame quality to analyze the video performance to the RP application requirements. Using static bitrate video delivery as a baseline, we also discuss the performance trade-off between GCC and SCReAM in the air.

(3) We find that the majority of the connectivity requirements desired by RP can be satisfied by the current LTE infrastructure, albeit with some limitations. Firstly, we observe that high-quality streaming between full-HD and 4K resolution was achievable in the air up to an altitude of 120 m with less than 300 ms playback latency. However, performance reliability was largely missing within LTE networks as low latency and high-quality video streaming could not always be maintained during flights. The performance largely varies depending on the geographical location and the Mobile Network Operator (MNO). The achieved video delivery performance is highly dependent on the density and availability of nearby base stations.

(4) We discuss the implications of our findings relevant to the development of future cellular standards and open research challenges that must be overcome if RP over cellular networks is to become a reality. We further discuss potential improvements in the video delivery pipeline to ensure smooth and low-latency video delivery for RP.

To foster reproducibility and motivate future research in this field, we publish our collected dataset, the video streaming pipeline

used as the workload for this study, and the parsing and visualization scripts at [11] and [10].

2 BACKGROUND AND RELATED WORK

In this section, we describe the RP use case in detail and its challenging connectivity requirements for ensuring safe aerial operations. We also provide more information about related studies and outline the novelty of our work.

2.1 Remote Piloting Scenario and its Connectivity Requirements

Remote piloting (RP) is one of the promising approaches to realize the next-generation aerial platforms in the sky, namely UAVs and eVTOLs [12]. The goal of RP is to allow pilots to remotely control a fleet of aerial vehicles from the ground operation center or elsewhere. The pilots send command packets to the UAVs and receive video and telemetry streams in return. The RP concept can be realized with different aerial vehicles types such as fixed-wing aircrafts, rotary-wing aircrafts, airship, and others [32]. We focus on rotary-wing aircrafts in this study.

Wireless links are one of the building blocks in the remote piloting architecture defined by ICAO [32]. They are used to establish bidirectional communication between the aircrafts and remote pilots. Video is a major part of the data stream from the aircraft to the ground station [54] as it enables the remote pilot to sense the environment in real-time to maneuver the aircraft and avoid conflicts in the airspace. As a result, low latency and high-quality video transmissions are essential for safe vehicle operation from the ground. Hence, wireless networks (mostly cellular due to their dense availability and reach) play a vital role in RP operations by providing reliable and robust connectivity between remote pilots and the aircraft to ensure safe aerial operations. Thus, RP operations impose challenging connectivity requirements on the end-to-end network. Specifically, the use case demands between 10–100 Mbps for acceptable video streaming, end-to-end latency within 150–300 ms, and 99.999% communication reliability [12]. Regarding the specifications of the video application and streams, aviation regulations are still at an early stage in defining the minimum requirements. Nevertheless, similar demands to those in the automotive industry can be expected, where the 5G Automotive Association (5GAA) specifies *high-quality streaming* for teleoperated driving [5].

We believe that mapping the connectivity requirements from the network to the video application layer requires extensive study. Nevertheless, our derivation of the video application requirements will help in evaluating the potential of supporting RP applications with cellular networks and uncover the outstanding research challenges. Recent works have shown that LTE can support the minimum specified latency and bandwidth requirements imposed by RP in almost all developed regions across the globe [17, 19]. In regions with less dense availability of cellular infrastructure, utilizing multiple cellular connections simultaneously through multipath transport shows promise. However, multipath transmissions are more susceptible to fluctuations caused by obstructions and handovers in last-mile [40], hence understanding the performance of RP over LTE offered by multiple operators is the first step to realizing the use case in practice.

2.2 Related Work

Several studies have looked into video streaming from UAVs in the past. We highlight the works best fitting the explorations relevant to this study, such as bidirectional video and control data exchange, real-time video delivery, and various video-streaming algorithms.

Bi-directional video and control data transmission. Jin et al. [34] utilized commercial off-the-shelf hardware with H.265 encoding for video and the MAVlink protocol for the control traffic. The measurements reveal that the latency of video traffic is up to 1.2 s in 5G and 3 s in 4G, while the minimum control signal latency is 30 ms. In [51, 61], the authors studied the end-to-end latency of cellular networks while transmitting control and video traffic for UAV operations. The authors show that the video latency is worse than control traffic, up to 8 s in real-world measurements, due to high mobility. Furthermore, study [9] sets up a multipath communication scenario using link diversity and forward error correction mechanisms to enable real-time video transmission beyond visual line-of-sight operations of UAVs. The measurement results highlight that using uncorrelated links improves the video quality by up to 33%. Lastly, Yu et al. [60] present a channel co-allocation algorithm to enable a multi-UAV full-duplex connectivity scheme. By enabling channel reuse within multiple UAVs, they aim to reduce the co-channel interference with low transmitter complexity and higher spectrum efficiency.

UAV-based real-time streaming. In [46], the authors implement a WebRTC system on a UAV that consists of a camera, a media gateway, a backend, and a website. Streaming two different cameras over RTP, their emulation study shows that it is possible to achieve 30 FPS with ≤ 1 s latency [46]. However, the authors perform their measurements over a wired medium, which does not have the fluctuating properties of wireless channels and cellular networks. In [53], the authors describe their ARM-based 5G hardware chain for real-time video transmission. The 5G module handles the data processing, and they transmit uncompressed 480p video at 30 FPS. Without providing statistical insights, they claim a transmission rate between 19 and 112 Mbps and a playback latency between 49 and 266 ms, depending on the cellular standard used.

Improving video transmission quality of UAVs. In [6], the authors implement an automatic repeat request mechanism at the application layer to minimize the video packet losses in ad-hoc UAV scenarios. Through a measurement campaign with a flying UAV transmitting H.264 compressed full-HD video over an 802.11n link, the proposed scheme was able to reduce the packet losses by almost two orders of magnitude. The authors of [16] take a different approach and introduce Machine Learning-based scheduling to improve the video quality for omnidirectional 360° video transmissions. At each transmission interval, the scheduler prioritizes a traffic class in the time domain and a scheduling rule for that traffic in the frequency domain based on various parameters such as RF channel, traffic profile, and application requirements. Their simulations show that the novel algorithm outperforms other baseline scheduling methods in packet loss, latency, and throughput.

Xiao et al. [58] show correlations of flight altitude / speed and RF channel conditions during UAV flights in 802.11n networks and propose a new adaptive bitrate algorithm to utilize the sensor data with network observations. Due to unclear correlations, the authors

further utilize neural networks to accurately estimate the varying channel conditions from flight states. They find that the proposed scheme improves the average Quality of Experience (QoE) by 21.4% and the average bitrate by 10.8% compared to state-of-the-art adaptive bitrate algorithms. As we show below, we did not observe similar correlations between flight parameters and RF channels in cellular networks (except altitude). In [48], the authors set up a user study to assess the QoE of a cloud-based UAV simulator. With 16 participants, they measured the impact of the video encoding parameters on the experienced video quality. The study concludes that while resolution strongly affects the perceived graphics quality, participants did not observe a major difference between 30 FPS and 60 FPS. However, poor graphics quality could cause simulator sickness symptoms. The results of this study are quite relevant to us since the authors dictate the acceptable video quality metrics for operating UAVs. However, their focus was only on measuring the influence of the video encoding parameters on the video QoE in a controlled environment. On the other hand, our study complements their work by investigating the achievable video delivery performance over cellular links in the air.

Mobile HOs in the air. The authors of [25–27, 29, 50] experimentally evaluate the HO count at various altitudes. According to [25, 29], increasing altitude positively correlates with the HO count in suburban areas – increasing up to $5\times$ compared to ground users. In contrast, the authors of [27] report a negative correlation in their measurements in a rural environment, and the studies [26, 50] observe no connection between the HO count and altitude in both urban and rural areas. Thus, the studies do not converge to common findings and demand further investigation. Additionally, higher UAV speeds can increase the HO count [50] and degrade the network performance [27].

Overall, our study differs from previous research in this field in the following aspects. (1) We design and implement a realistic low-latency adaptive video transmission setup using two different CC variants specifically for wireless and mobile scenarios. (2) We assess the achievable video delivery performance over cellular networks in the air in two largely complementary environments with a safety-critical RP use case in mind. (3) Our extensive in-the-wild measurements allow us to investigate and contrast the performance of cellular connectivity experimentally (e.g., frequency of HOs, duration of HETs, etc.) along with its impact on video streaming in both air and ground.

3 MEASUREMENT METHODOLOGY

3.1 Setup Configuration

Our study aims to analyze the performance of high-quality video delivery for remotely piloting UAVs in outdoor environments over a cellular network. Hence, we design our measurement setup to mimic the requirements and operations of remote piloting applications closely – specifically focusing on the environment, altitude, data traffic and network. Furthermore, we perform our measurements in two distinct environments involving different complexities for RP applications – *urban* offering many handover opportunities thanks to the dense availability of base stations and obstacles/reflections due to tall buildings, and *rural* with limited BS

coverage but large open areas with the possibility of direct line-of-sight connectivity throughout flight duration. We base our analysis on ≈ 7 GB of collected video transmission data, including ≈ 60 M packets recorded from 130 measurement runs over a total of ≈ 90 flights in rural and urban environments.

Hardware Setup. Our measurement setup is illustrated in Figure 2. We used a DJI-M600 UAV, which can carry a payload (our measurement setup) of up to ≈ 5 kg in the air. The payload consists of two Intel NUCs, each equipped with i7-7567U CPUs and two RPis, both running Linux operating system. We duplicate our measurements across the RPi and NUC to (i) remain hardware independent and (ii) collect more data per measurement run to increase our confidence level in the resulting analysis. All four measurement nodes were connected to an LTE network via CAT4 LTE USB modems. We attached GPS antennas to the two RPis which enables them to listen to the PPS signal and maintain an accurate system clock. We further set up the RPis as NTP servers and attached them to each NUC via Ethernet – making all four nodes time-synchronized. We used two large 20 Ah batteries to power our setup.

Our remote pilot is a compute-optimized Amazon Web Services (AWS) EC2 instance equipped with 8 CPU cores (c5a.2xlarge instance) hosted in the UK region ($\approx 1,000$ km from the measurement locations around Munich, Germany). The server uses the Amazon Time Sync Service for timekeeping, which claims to be highly accurate through satellite and atomic clock sources [8]. We use the LTE connection from two major MNOs, which are known to have the best coverage in the test region. In this work, we chose not to utilize 5G for our measurements due to the following reasons. *Firstly*, like the majority of the globe, 5G roll-out in Munich is still severely restricted to a few locations (mainly around the city center) hence limiting the generality of our analysis. Furthermore, most of the current 5G deployment is still *non-standalone*, i.e., it relies on existing LTE radio access networks for last-mile connectivity instead of the mmWave links [41]. On the other hand, LTE connectivity across the globe is fully mature, and the latency and bandwidth promised by the standard should suffice for RP operations [5, 17, 19]. Note that we design our methodology and experimental configuration to be independent of the underlying cellular standard, meaning that it will be easy to undertake a similar measurement campaign over 5G once the standard is more mature and widely available.

Our LTE data plan allowed us to perform unlimited download and upload transfers, but our downlink and uplink speeds were capped at 300 Mbps and 50 Mbps, respectively. The lowest recorded Round Trip Time (RTT) between the AWS server and our measurement node over LTE is ≈ 35 ms. We also highlight the impact of different MNO infrastructure on RP performance in the same region in Section 5 and Appendix A.2.

Flight Configurations. Throughout this study, we performed a total of ≈ 90 flights in an urban and a rural location on different days and at different times. Our *urban* flying zone is in the university campus located at the center of Munich, Germany (≈ 1.5 M inhabitants), surrounded by moderately high buildings (≥ 25 m). We conduct *rural* flights in more than 20 km of open space in the outskirts of Munich’s metropolitan area ($\approx 3,700$ inhabitants). Figure 3 shows the terrain view and the locations of the BSs in close vicinity to our flight regions. The most notable differences

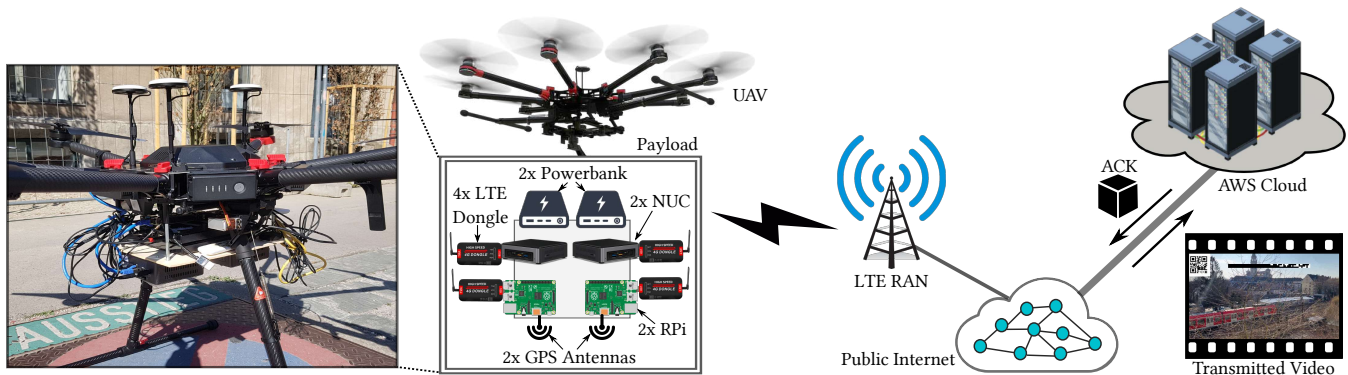


Figure 2: Our measurement setup. The UAV is equipped with two Intel NUCs, two Raspberry Pis (RPIs), two GPS antennas, and four LTE dongles, all of which are powered by two power banks. The RPIs receive accurate time readings from the GPS signal and serve as a Network Time Protocol (NTP) time server for the NUCs. Each Intel NUC sends a full-HD video over LTE to our cloud server and receives acknowledgments in return. The bottom left shows a frame of the transmitted video. Each frame contains a QR code with the current frame number and a barcode with the encoding time. These data points allow us to assess the video delivery performance.

between the two locations are the obstructions (density and size of surrounding buildings) and the number of BSs. Our two measurement locations provide a complementary-yet-complete view of different environments within which RP applications would have to operate.

We followed a pre-designed trajectory for all our flights, allowing us to perform repeatable measurements across both test locations throughout our measurement campaign. Since cellular operators do not disclose their cell coverage data, we maximized our horizontal and vertical movements within the allowed flight area to ensure HOs between different cells. We managed to observe multiple HOs during each flight. Our trajectory represents basic movements involved in typical RP flights, i.e., take-off, cruise, ascending and landing. We flew up to a height of 120 m above ground (maximum allowed height as per European regulations [22]) at a median and maximum recorded speed of 13 kmph and 60 kmph, respectively. Readers should refer to Appendix A.2 for further details on our flight trajectory.

3.2 Data Collection

The primary requirement of RP is to support a high-quality video stream from the onboard camera of a UAV to the remote pilot over the Internet. To closely mimic such an operation, we utilize a 30 FPS and Full-HD (1920×1080 pixels) resolution video stream encoded at bitrates between 2 and 25 Mbps using H.264 as our workload. We set our highest rate to 25 Mbps to evaluate whether a higher quality stream, e.g., 4K [30], could also be supported. Moreover, it allows us to test the maximum achievable bitrate while maintaining real-time video delivery. We transmit video in real-time using RTP over UDP and utilize two purpose-built CC algorithms for adaptively switching bitrates depending on the network conditions: GCC [14] and SCReAM [35]. GCC estimates the queuing delay gradient with a Kalman filter to detect congestion. A loss-based controller complements that algorithm. The sender sets the target bitrate based on

reports provided by the two controllers [14]. On the other hand, SCReAM sets the video target bitrate based on the packet transmission and acknowledgment rate. It applies packet pacing and limits the bytes-in-flight to a congestion window that is increased as long as the RTP queue is shorter than 300 ms. Otherwise, if packet loss is detected, the window is decreased [35]. Both algorithms rely on feedback from the receiver but work with different RTP Control Protocol (RTCP) extensions. In the implementations that we utilize, GCC relies on the RTCP extension for transport-wide congestion control [31], while SCReAM works with the extension format specified in RFC 8888 [47]. Both CCs log the timing information of the received packets – allowing us to calculate the one-way delay of the transmission. To estimate a baseline performance, we determine the maximum bitrate within the 2–25 Mbps range over repeated trial runs at which the LTE link supports a stable video delivery in each location. Based on those tests, we send the constant (or static) bitrate video at 25 Mbps in the urban area and at 8 Mbps in the rural area. To summarize, we use three different varieties of video workloads in our measurements in rural and urban environments: *adaptive* RTP transmissions using (1) GCC and (2) SCReAM CCs, and (3) *static* bitrate RTP transmissions at the maximum “support-able” bitrate.

Video Streaming Pipeline. Our video pipeline utilizes the multimedia framework GStreamer [28] and implementations of GCC and SCReAM [45]. A wrapper application [23] orchestrates the encoding and streaming of the video on the drone and the decoding and playback on the AWS server. To support repeatable and reproducible measurements, we employ a pre-recorded video (“source video”) that contains considerable detail and motion as our input instead of an actual camera. The source video is re-encoded in real-time by the VideoLAN x264 encoder [56] at a specified bitrate. For *static* video, we set the bitrate to 25 Mbps and 8 Mbps for urban and rural environments, respectively, based on findings from our trial measurements. For *adaptive* video, the bitrate is dictated by the

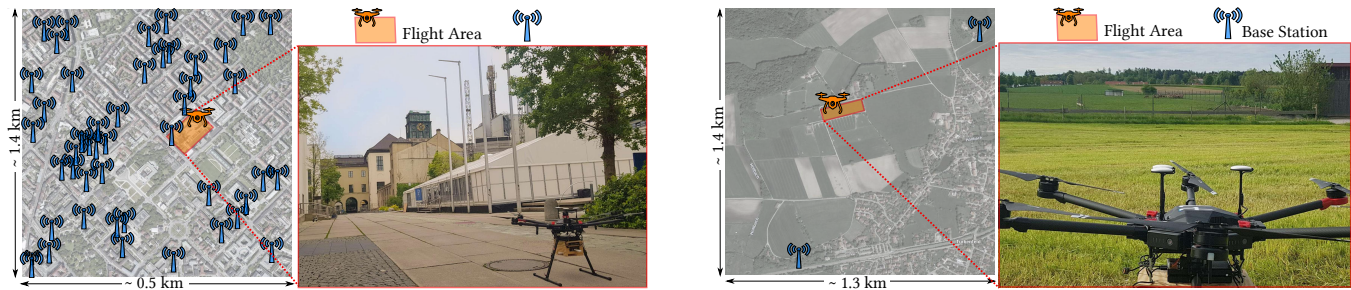


Figure 3: The measurement areas along with the BSs in their vicinity. The left-hand side depicts our *urban* flight zone that is densely surrounded by buildings and BSs. The right-hand side shows the open space and low BS density in our *rural* flight environment. We use data from [13] to plot the approximate locations of the BSs. During our measurements, we were connected to 32 and 18 different cells in the urban and rural locations, respectively.

employed CC algorithm. To allow calculation of the video delivery performance (detailed below), we embed additional information into each frame (see the example in the bottom-right corner of Figure 2). The pre-encoded video contains a QR code corresponding to a uniquely increasing number. Additionally, we built a GStreamer element that dynamically embeds the current time into each frame during encoding. The encoded video is written to a file and in parallel split into UDP-enclosed RTP packets that are sent to our server and placed in an instance of GStreamer’s RTP jitter buffer. We buffer packets for 150 ms to cushion the variable packet arrival rate and handle out-of-order packets. The player at the AWS server also writes the received video to a file, allowing us to calculate the per-frame performance of the stream. Our video delivery pipeline is available as open source at [10].

Data Logging. We collect the packet-level data with tcpdump, the sent and received video streams, and data logs generated by GCC and SReAM. We also use QCSuper [37] to record the LTE Radio Resource Control (RRC) messages which allows us to accurately detect the start and end of HO events. We utilize these to calculate cellular performance metrics such as *HO frequency*, i.e., the number of HO events encountered in a time interval, and *Handover Execution Time (HET)*, i.e., the time between the reception of the *RRCConnectionReconfiguration* packet from the source BS to the transmission of the *RRCConnectionReconfigurationComplete* packet at the target BS [1]. We draw our resulting analysis from our collected dataset that spans upwards of 7 GB and includes $\approx 60M$ data packets from over 130 measurement runs, which we publish at [11].

Video Delivery Performance Metrics. In addition to the network-level performance metrics, such as RTT, packet loss, HO, etc. we also estimate the quality of video transmission (and later correlate it to the requirements imposed by the RP use case) by calculating the following metrics.

- (1) *Frame Rate (FPS)*. The frame rate of the video is directly logged by a GStreamer element at the receiver, which can be compared to the original video frame rate of 30 FPS. We additionally compute the FPS in post-processing from the received video.
- (2) *Playback Latency*. The playback latency is the duration between the start of the encoding of a frame at the sender and the playback of the decoded frame by the receiver. We calculate the playback latency by correlating the encoding timestamp that is stored as a barcode

in each frame and the decoding timestamp that is logged by the receiver. Note that FPS and playback latency are directly correlated, i.e., a lower frame rate results in an increase in playback latency and vice-versa. The playback latency can, however, suddenly drop without an FPS increase if frames are skipped, and it can stay at an elevated level even though the playback frame rate matches the source video frame rate.

(3) *Structural SIMilarity (SSIM)*. The SSIM index assesses the received frame quality by measuring the degradation of the luminance, contrast, and structure information [57]. We calculate SSIM by comparing the source and received frames.

(4) *Video Stall*. A video stall event occurs when the inter-frame time exceeds the latency requirement of the RP use case, which we estimate to be ≈ 300 ms [12].

4 MEASUREMENT RESULTS

In this section, we first elaborate on the networking performance and compare how the network conditions in the air differ from those on the ground. Next, we analyze the video delivery performance in Subsection 4.2 by analyzing the achieved video throughput. We also evaluate the playback performance of the adaptive video delivery methods w.r.t. the achieved frame rate, playback latency, and video quality calculated with the Structural SIMilarity (SSIM) method. These metrics allow us to pinpoint video stalls as well as frame drops and to evaluate whether the playback latency necessary for RP is achievable. We conclude with a discussion of our results.

4.1 Networking in the Air

To highlight the aerial networking challenges that affect the video delivery application, we compare the networking performance in the air to the ground in terms of HO frequency, HET, end-to-end latency, goodput, and PER. We collected the aerial data with the help of a UAV that flew both horizontally and vertically (see Section 3). We used a motorbike, at speeds similar to the average flight speed, to mimic the horizontal movements of the UAV on the ground.

Figure 4 compares the HO occurrences on the ground and in the air in terms of HO frequency and HET duration. In Figure 4(a), we observe that the frequency of HOs during flight is about an order of magnitude higher than on the ground. Comparing the two test locations, the average HO frequency is higher in the urban

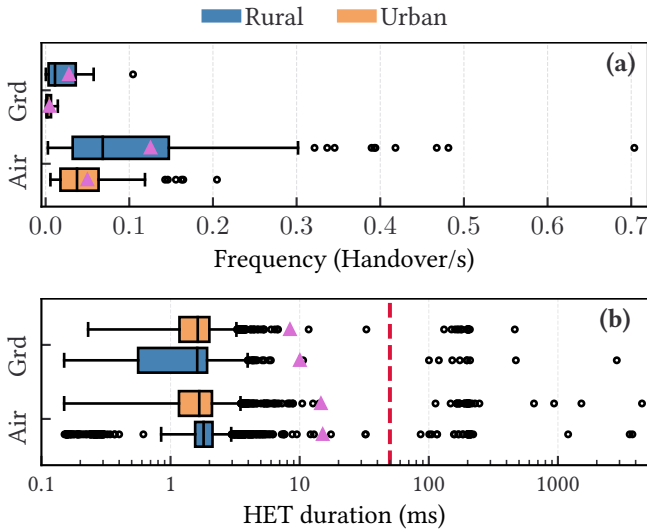


Figure 4: Comparison of HO performance in the air and on the ground. (a) HO frequency, (b) HET in the air and on the ground. In (a), the HO frequency is an order of magnitude higher in the air than on the ground. In (b), we clearly observe HET outliers above 100 ms, which cause long service interruptions in the air. The purple triangle marks the average of each boxplot.

area both on the ground and during flight. We see two factors that can cause the increased HO frequency in the air. *Firstly*, the number of line-of-sight channels to different BSs increases in the air, causing increased inter-cell interference levels along with link fluctuations. *Secondly*, as the BS antennas are down-tilted to provide optimal coverage for ground subscribers, the UAV can enter the side-lobe coverage area of the antennas, which can contribute to the link fluctuations. We note that the collected ground dataset likely includes longer durations without horizontal movements than the air dataset. The results may, therefore, be skewed towards a lower HO frequency on the ground.

Figure 4(b) shows that the majority of HOs last less than 49.5 ms, which 3GPP specified as the HET threshold for successful HO events [1]. An important observation is the number of outliers that last longer than 50 ms. The majority of excessive outliers above 500 ms occur in the air and range up to 4 s. Such large interrupts can severely impact the video delivery and the experience of the remote pilot. Possible explanations for the increased density of outliers in the air could be fluctuating Radio Signal Strength Information (RSSI) levels as well as a higher noise floor, as noted by previous works [29, 55]. These factors could have prolonged the link interruption until the User Equipment (UE) has established a stable link with the next serving BS.

We compare the one-way latency of the RTP packets from sender to receiver between ground and air measurements in Figure 5. We find that for $\approx 99\%$ packets, the latencies lie below 100 ms of the ground transfers and $\approx 96\%$ of the packets with some outliers that last more than 1 s in the air. Our analysis shows that these outliers

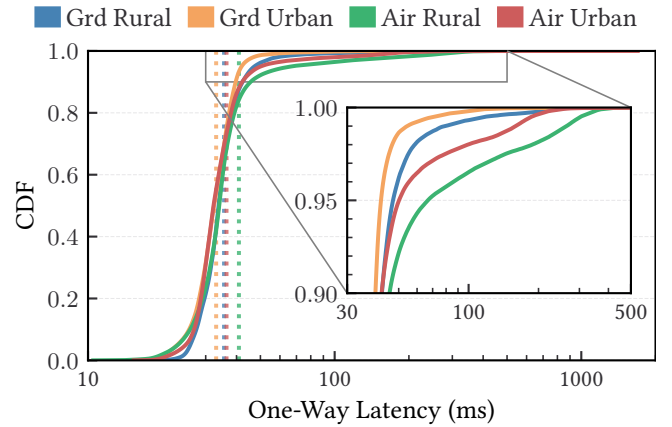


Figure 5: Comparison of end-to-end latency on the ground and in the air in the urban and rural environments. The dashed lines represent the average latency. The plot shows the increased end-to-end latency in the air in both environments and the higher latency in the rural environment compared to the urban test location.

most often happen at altitudes > 100 m (see Figure 13). Multiple factors are involved in the degraded latency conditions in the air. The increased HET and the HO frequency in the air can cause briefly elevated one-way latencies. We discuss this behavior in Section 4.2.2, where we show that latency spikes tend to precede HO events. Neither flight velocity nor altitude significantly affects the PER and goodput performance. The PER stays within 0.06–0.07% and the average goodput is similar in the air and on the ground. Due to deep network buffers of cellular operators [33], worsened HO conditions and link fluctuations in the air do not impact PER. Most of the observed packet drops occurred consecutively.

Overall, our analysis highlights the increased HO frequency in the air and a higher probability of long HET outliers. Consequently, the end-to-end latency performance degrades while the PER and goodput performances are comparable across air and ground. For RP, these results imply a higher number of link disruptions in the air, up to 0.7 HOs/second, with longer outage durations. Such events can trigger undesired service interruptions for the RP use case with unbearable outage durations. Next, we analyze the achievable video delivery performance in the air, bearing in mind this section’s findings and the derived RP video requirements from Section 2.1.

Takeaway – More frequent HOs and excessive HET outliers in the air can degrade the end-to-end latency while throughput and PER are not impacted.

4.2 Video Streaming Performance

In this section, we analyze the video delivery performance in terms of bitrate, FPS, playback latency and SSIM. We first evaluate the achievable video delivery performance with the requirements of the RP use case in mind and then compare the performance of GCC and SCReAM with the static bitrate video delivery baseline. Afterward, we dig deeper into the collected dataset to further understand the

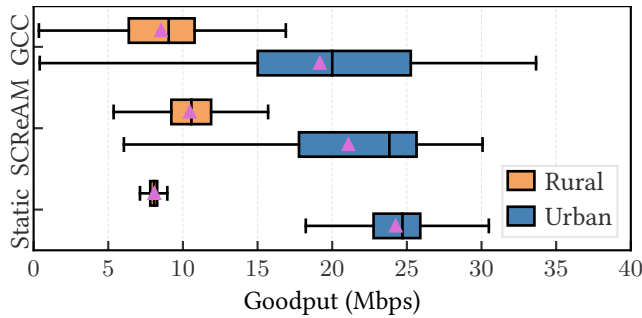


Figure 6: The baseline performance of video delivery was set with a constant bitrate stream of 8 Mbps in the rural and 25 Mbps in the urban area. The selected bitrates were adapted to the link capacities in each area. With abundant link capacity in the urban area, the constant bitrate (static) stream managed to maximize the bitrate. With limited link capacity in the rural area, SCReAM is best at maximizing the bitrate, especially during link fluctuations.

influence of the network layer on the video performance and look closely at the impact of HO events. In the rest of the paper, *high bitrate* and *low bitrate* refer to the video transmission at 25 Mbps and 8 Mbps (see Subsection 3.2). Also, the term *low-latency* refers to latencies below the threshold of 300 ms and *high-quality video* is a video with a minimum of HD resolution.

4.2.1 Bitrate and FPS Analysis. The goodput metric is an indicator of the achievable video quality, and the FPS describes the smoothness and stability of the video transmission. Figure 6 shows the achieved bitrate performance in the urban and rural test environments, respectively. In the urban location, the achieved goodput varies between 20–25 Mbps depending on the bitrate adaptation method while the goodput in the rural area is limited to 8–10 Mbps.

The goodput analysis of the different bitrate adaption methods in the urban environment shows that SCReAM comes closest to the hand-picked static bitrate of 25 Mbps with 21 Mbps on average while GCC achieved 19 Mbps. In the urban environment, the MNO provides an uplink capacity of up to 40 Mbps which allows the static transmission to maintain the high bitrate (see P1 in Figure 10). The CCs increase the target bitrate conservatively and lower it if signs of congestion are detected to optimize for stable and low-latency video delivery. The bitrate ramp-up phase of GCC and SCReAM at the beginning of video transmission takes ≈ 12 and 25 s until they can reach the target of 25 Mbps, respectively. This contributes to the lower tail of CC goodput. One reason that causes a bitrate decrease is packet loss that occurs at altitudes above 80 m in the urban environment. In the rural environment under limited bandwidth, SCReAM is better at utilizing the available link capacity than GCC. SCReAM achieves an average bitrate of ≈ 10.5 Mbps compared to GCC’s ≈ 8.5 Mbps. However, both numbers are higher than the static transmission at 8 Mbps – highlighting the high variability of the available link capacity in the rural area.

In our analysis, we found cases in which SCReAM wrongly detected packet losses which caused it to lower the target bitrate needlessly. This behavior stems from the implementation of the

RFC8888 acknowledgment feedback generation in the SCReAM library from Ericsson Research that we used [45]. Every 10 ms, an RTCP packet is generated that carries information about the status of the RTP packet with the currently highest received sequence number and, by default, the 63 preceding packets. At rates higher than ≈ 7 Mbps, more than 64 RTP packets arrive between two consecutive RTCP packets, which causes some RTP packets to remain unacknowledged. We increased the number of acknowledged RTP packets per RTCP packet from 64 to 256 to lower the probability of these events. SCReAM does, however, discard its RTP send queue whenever it exceeds the length of 100 ms, which causes instantaneous and unpredictably large jumps of the highest received RTP sequence number. A more intelligent acknowledgment generation scheme is necessary to handle those events correctly and not leave received packets unacknowledged. Our measurements, therefore, include cases where SCReAM receives misleading feedback that causes it to lower its bitrate needlessly. The performance issues in the current implementation of SCReAM are already known to its developers. The design and implementation of an adequate acknowledgment scheme remains outstanding for future work.

We find that the FPS performance is primarily influenced by congestion events and packet losses. Figure 7(a) shows the FPS distributions. Both adaptive bitrate transmission methods can maintain 30 FPS $\approx 90\%$ of the time in the urban area, but the video is played with less than 10 FPS $\approx 1.5\%$ and 3% of the time with SCReAM and GCC, respectively. Both performance levels are worse than that of the static transmission, which manages to maintain a minimum of 8 FPS. On one hand, the static transmission performance highlights the reliability of the cellular network that supports a high bitrate stream in the face of link capacity fluctuations. On the other hand, this result is surprising since the CCs are designed to maintain a reliable video stream. Counterintuitively, we find that the low FPS outliers most often occur when the CC significantly decreases the target bitrate. In this case, already queued frames need to be sent with the decreased target bitrate before frames that are encoded to match the updated target bitrate are at the front of the send queue. The temporary mismatch of video and send bitrate starves the player’s buffer, which proactively lowers the playback rate to avoid running out of video data. The lower overall bitrate in the rural area alleviates the just discussed problem and, e.g., allows GCC to completely avoid 0 FPS situations.

Lastly, we use the FPS data to compute how frequently the video stalled (defined as an inter-frame time > 300 ms), which is very relevant for the RP use case. SCReAM and GCC produced 0.89 stalls/min and 1.37 stalls/min, respectively, while the video that was transmitted at a static bitrate experienced 0.11 stalls/min. These results are aligned with the analysis of FPS deviations and together highlight the negative effect of the adaptive bitrate adjustment.

Overall, our analysis highlights that video can be transmitted at a bitrate of up to 25 Mbps if the MNO provides sufficient resources. Nevertheless, even 8 Mbps fulfills the bitrate requirements of a full-HD video, while 25 Mbps can even enable 4K streaming [30]. The occurrence of FPS deviations that can be classified as video stalls indicates problems that must be solved to ensure a smooth and reliable video stream for the safe operation of remote vehicles. Next, we elaborate on the achievable playback latency while delivering video in the air.

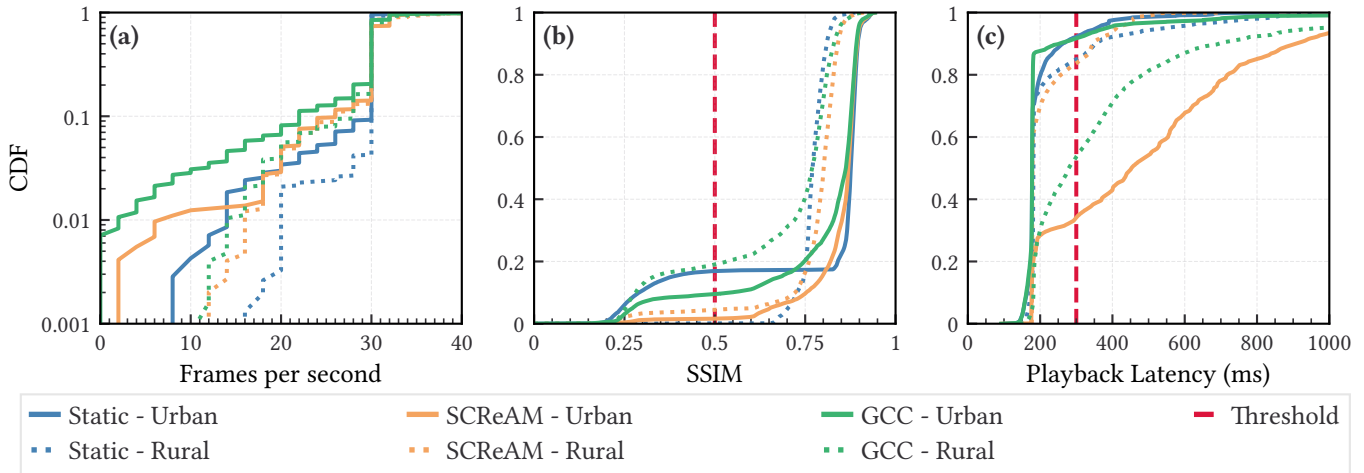


Figure 7: Adaptive video delivery performance in the urban and rural environments. GCC and SCRaM deviate from 30 FPS more often than static delivery due to sudden target bitrate decreases. While SCRaM manages to minimize SSIM drops below the quality threshold in both environments, its playback latency performance is unacceptable in the well-provisioned urban environment. GCC on the other hand fails to achieve a low playback latency in the rural area.

Takeaway – The aerial bitrates enable the delivery of high-quality video for RP operations and the static transmission method demonstrates stable playback rates. The performance of the CCs, however, shows room for improvement with more severe FPS deviations and more frequent video stalls.

4.2.2 Video Playback Latency Analysis. Figure 7(c) shows the playback latency during measurements, which remains less than 200 ms under stable conditions. The main latency contributors are the one-way network latency (≈ 50 ms) and the RTP jitter buffer (≈ 150 ms). Depending on the bitrate control method, the playback latency threshold of 300 ms could be achieved 30–90% and 55–85% of the time in the urban and rural locations, respectively.

In the urban location, GCC is able to maintain the playback latency threshold of 300 ms for $\approx 90\%$ of the time and performs similar to the static bitrate transmission. The link capacity in the urban area provides sufficient headroom to maintain a static 25 Mbps video transmission with low playback latency. SCRaM, however, performs significantly worse when compared to the other methods in the urban location; the playback latency threshold is only met $\approx 38\%$ of the time. We find that the playback latency during SCRaM transmissions exhibits high fluctuations and occasionally remains at a plateau of around 1 s for a longer period. We suspect the culprit of those plateaus to lie within the player pipeline as the CC and network metrics do not show unusual values: the sending rate and network latency remain relatively constant while no losses occur. The RTP jitter buffer in the player pipeline that controls the video buffer timing is most likely at fault. We cannot pinpoint the issue more precisely and do not have a convincing explanation of why the other transmission methods do not trigger this behavior. On the other hand, we observe that SCRaM is effective at quickly reducing the playback latency by discarding the queued RTP packets.

We observe that the two CCs perform differently, depending on the available link capacity. While SCRaM remains below the

latency threshold only $\approx 30\%$ of the time in the urban area where 25 Mbps transmissions are supported, this rate increases to $\approx 85\%$ in the bandwidth-constrained rural area. The playback latency of GCC, however, has the opposite relationship with the transmission bitrate. In its current state, GCC is therefore more suitable for RP scenarios that require high bitrates, e.g., 4K video at 20–25 Mbps, whereas SCRaM could rather support lower video qualities.

We find that the playback latency is mainly determined by the network latency and present a section of a video transmission that demonstrates this influence in Figure 8(b). An increase of the network latency above the RTP jitter buffer size of 150 ms causes the playback rate to decrease and the playback latency to increase. The decrease of the network latency is accompanied by a normalization of the playback latency. Another cause of playback latency spikes are packet losses which only rarely occur in cellular networks due to mechanisms like Hybrid Automatic Repeat Request (HARQ) and deep buffers [33] but multiple consecutive packets drop if they fail.

We observe that network latency spikes often occur just before an HO. This correlation can be observed in Figure 8(a), which shows one of the HO events from Figure 8(b) on a narrow x-scale. In our measurements, some HO events cause the playback latency to increase to 900 ms. The spikes usually occur ≈ 0.5 s before HOs and last ≈ 1 s until the network latency is again stable around 50 ms, as shown in Figure 8(a). Other studies [43, 44] find similar correlations between HO events and spikes in the network latency.

To determine the impact of the HO events on the network latency, we analyze the ratio of the maximum-to-minimum latency in the 1-second time window before and after HOs, as visualized in Figure 8(a). The resulting maximum-to-minimum latency ratio in these time windows from all HO occurrences during flights is shown in Figure 9. Before an HO occurs, the maximum latency on average becomes 8 times higher than the minimum latency in the same time window. After an HO, the ratio is ≈ 5 on average. Some spikes can increase the before-HO ratio to more than 15 and even

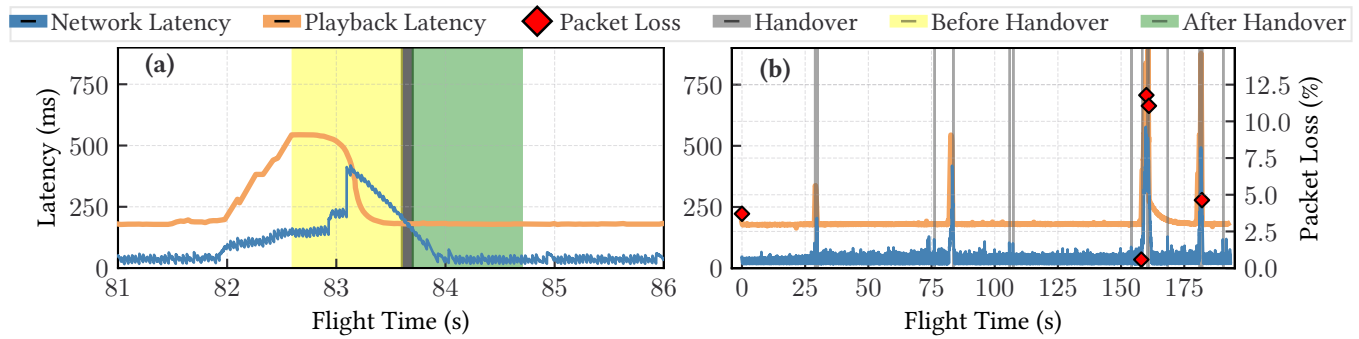


Figure 8: (b) Correlation between the playback latency, network latency, packet losses and handovers during one of the flights with GCC. Network latency experiences spikes during HO events and consequently, playback latency also increases when network latency goes above 200 ms. (a) Before and After Handover time windows around a handover instance from (b). By computing the maximum-to-minimum network latency ratio in these time windows during every HO events from all flights, the magnitude of the spikes on the network latency is quantified in Figure 9.

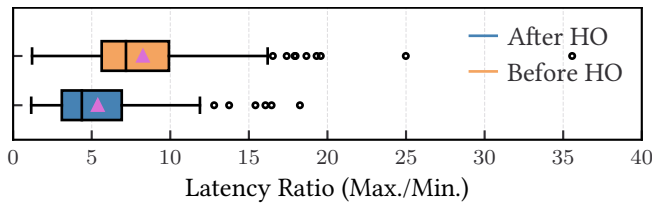


Figure 9: Maximum-to-minimum latency ratio before and after aerial HOs. The latency ratios are computed in 1-second time windows (see Figure 8(a)). The plot shows that the latency spikes that occur before HOs increase the latency on average 8 times, whereas this ratio is ≈ 5 on average after HOs. Some outliers increase the before-HO ratio up to 37 times.

up to 37. The increased aerial HO frequency exacerbates the issue of temporarily degraded playback latencies due to network latency increases prior to HOs.

In summary, the playback latency is kept below 300 ms for as much as 90% and 85% of the time during our measurements in the urban and rural areas, respectively. These video playback latencies are among the lowest when compared to other papers that worked with cellular links, even though we transmit at up to 25 Mbps.

Takeaway – The playback latency can be kept < 300 ms under stable conditions which is promising for the RP use case. This threshold is however not always met and CCs and video pipeline improvements are required. The network latency spikes that HOs cause are challenging for stable low-latency video delivery.

4.2.3 SSIM Analysis. We calculate the SSIM index to evaluate the quality of the received frames at the player by comparing them frame-by-frame in post-processing with the source video frames. The SSIM score ranges from 0 to 1, where 1 is reached if the source and received frames are identical and 0 if the frame was not played. The SSIM is closely correlated with the bitrate at which the encoder operates; a higher bitrate target allows the encoder to keep more detail in each frame. However, the SSIM is also sensitive to packet

losses, which cause visual artifacts in the output of the video decoder. We consider frames with an SSIM score of 0.5 or higher to meet the video quality requirements of RP. This threshold only applies to the video that we use in our tests and is based on our subjective measure of the quality of the received video.

Figure 7(b) shows the overall SSIM performance from our measurements. In our urban tests, SSIM stays above ≈ 0.9 for 90% of the time while this measure reduces to ≈ 0.8 during our rural tests. In both environments, the SSIM threshold of 0.5 cannot be met for ≈ 0.37 –19.09% of the time depending on the video delivery method. In practice, this result indicates the percentage of the time during which the remote pilot does not receive video with sufficient quality to maneuver the UAV safely. Hence, the selection of the video transmission method plays a significant role in minimizing the video interruptions caused by insufficient video quality.

In our urban measurements, we observe a benefit of the adaptive bitrate transmission methods in minimizing video quality interruptions (SSIM score < 0.5). SCReAM is able to minimize the video interruptions to 1.63% while those occur 16.93% of the time with static bitrate transmissions. GCC has more video quality interruptions but less playback stalls in the bandwidth-limited rural test area compared to the urban environment. The video quality is impaired by artifacts that are caused by packet losses.

Overall, the findings highlight that sufficient quality to properly perform RP could be achieved up to 98.27% and 99.63% in urban and rural locations, respectively. In our comparison, SCReAM is overall most successful in reducing the SSIM outliers in both environments. The low aerial video interruption rates are already promising but need to be avoided entirely to support safe RP operations. Next, we discuss our findings and elaborate on their practical implications for the real-life implementation of high-quality video delivery for RP operations.

Takeaway – High-quality video delivery can be maintained up to 99% of the time. The selection of the video delivery method has a significant impact on minimizing the disruptions due to insufficient video frame qualities.

Analysis Overview. Our measurement analysis suggests that up to 25 Mbps stable video delivery at altitudes up to 120 m is achievable if the MNO provides sufficient resources. Such achievable bitrates can support full-HD streaming and provide high-quality video for RP operations. However, the observed FPS stalls and fluctuations during measurements indicate disruptions in video transmission, which can severely degrade the pilot’s capabilities to control a UAV. Furthermore, our SSIM analysis reveals that there is still an open room for improvement to maintain sufficient quality on the received frames to ensure high-quality video > 99.999% of the time. The RTP jitter buffer size can be adjusted to reduce playback latency further.

5 DISCUSSION

Analyzing cellular performance without visibility into low-level RAN operation. Measurement studies utilizing cellular networks as an access technology in the wild are tasked with accurately correlating changes in the underlying cellular last-mile link to application performance impact [36]. The problem, however, is not straightforward since network subscribers only have limited insight into the cellular network’s configuration, usage, and behavior. While we record and utilize the LTE RRC messages in our analysis, we lack information about the network’s internal configuration parameters, cross-traffic, and communication channel resources. The impairment in visibility limits our analysis, specifically to accurately correlate between network characteristics and the drone location, height, speed, and, more importantly, the performance of the video streaming pipeline. Future research can consider cross-traffic and how the variance in LTE performance arising from peak hours might influence video delivery performance. Furthermore, due to the weight and power restrictions, we could not include a channel sounder in our payload to accurately measure the signal and interference levels. Moreover, the Long Term Evolution (LTE) dongles that we use only report the signal strength in one-second intervals, which we find too coarse to correlate with application performance and draw meaningful conclusions.

Generality of measurements and analysis. Our measurement methodology included an urban metropolitan and a rural location to be as generic and geographically independent as possible. Our urban and rural measurements demonstrate the LTE performance in locations with contrasting properties in RF environment and the density of base station deployment. In this sense, our urban location represents an ideal cellular deployment for RP operation with dense base station deployment and abundant coverage. On the other hand, our rural measurements act as an extreme case for RP with limited base station availability. Furthermore, we also used LTE connections from two different MNOs in our measurements to avoid operator-dependency (see Appendix A.3 on the performance comparison of different MNOs). Hence, we aimed to generalize our findings by presenting the performance under *best* and *worst*-case scenarios in terms of locations, and provide an operator-independent analysis. Nevertheless, cellular coverage and configuration may vary in other geographical locations, resulting in different operational experience [19]. Further research can evaluate scenarios with different operators in different locations and countries to obtain more generalized findings. Furthermore, replicating the RP application on a controlled testbed that allows

granular simulation/emulation of cellular networks would enable researchers to (i) investigate the impact of different configurations on application performance and (ii) surpass the many challenges of conducting aerial measurements in real-world. In the future, we would also like to evaluate the RP performance with increased traffic in the airspace sharing the same cellular link.

SWaP requirements of UAVs. The small-scale UAV platforms have challenging Size, Weight and Power (SWaP) requirements that affect the video delivery setup. An early version of our measurement setup used small-scale and power-efficient RPIs to stream video and relied on the hardware video encoding capabilities. However, the hardware encoder could not maintain a constant frame rate, which forced us to migrate to considerably larger Intel NUCs with laptop-grade CPUs. The Intel hardware encoder—accessible through the Video Acceleration API (VA-API)—outputs frames reliably but at a relatively high and unstable encoding latency. After multiple trials, we used an H.264 software encoder with the NUCs, which could consistently output video at low latency. However, the encoding performance comes at the cost of significantly higher energy consumption than a hardware-accelerated solution, which limited the number of consecutive flights we could perform. We hope our findings serve as helpful inputs for researchers planning to conduct measurements using UAVs in the future.

Mitigating influence of HOs on RP. We observe that network latency frequently spikes before HOs – affecting the end-to-end playback latency. Several methods could lessen the impact of HOs on video transmission performance. Firstly, the hysteresis margin for the change in link quality and the time-to-trigger parameters for time intervals are basic configuration options that control the trigger time of HO events. These parameters can be optimized for aerial scenarios to (1) minimize the frequency of HOs in the air and (2) avoid unnecessary ping-pong HOs (as shown by [59]) that we also observed in our rural measurements. Secondly, the novel HO mechanism *Dual Active Protocol Stack Solution* [4] introduces a *make-before-break* link establishment architecture. In contrast to *break-before-make* methods, they avoid link disruptions in the air and could hence remove the observed latency spikes. Finally, other studies that also observed latency spikes around HOs in LTE (e.g. [43, 44, 49]) report that such spikes are largely missing in 5G stand-alone deployments. These findings suggest that relying on 5G instead of LTE might improve video delivery. We aim to explore these findings in future iterations of this work.

Impact of infrastructure operator. To understand the impact of different cellular infrastructure operators on performance, we performed several experiments using an LTE connection from a competing operator (P2) in the rural region and the default operator used throughout our study (P1). While the BS density of both operators is quite similar in our urban test area, P1 has significantly less density in the rural region. Consequently, P2 experiences more frequent HOs and offers us a higher bandwidth capacity in the rural region compared to P1, as shown in Figure 10. Therefore, the available infrastructure of the operators in the region of interest can significantly impact the achievable video delivery performance. Readers should refer to Appendix A.3 which elaborates further on the performance comparison between the video delivery over both network operators.

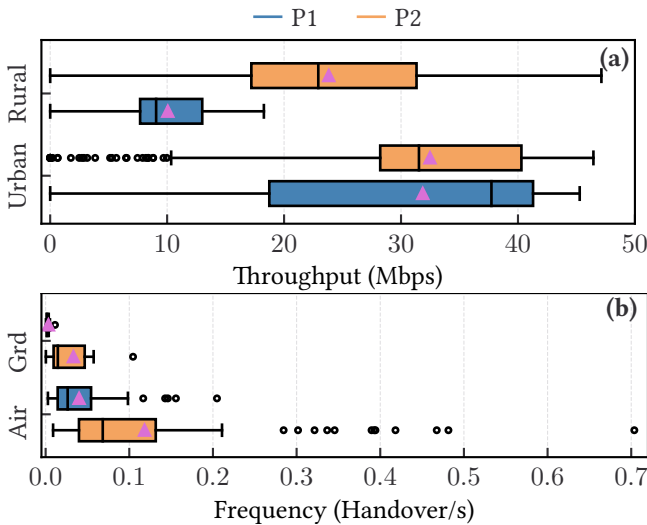


Figure 10: (a) Comparison of the achievable throughput and (b) HO frequency from two competing operators in the rural region.

Supporting low-latency video delivery for RP. Our results show that many requirements of the RP use case, especially concerning low-latency video delivery (target playback latency), can already be supported by LTE (and shows promise in 5G). However, several challenges remain to be tackled. Firstly, safety-critical RP applications can demand dedicated network resources and very high network availability (99.999% [12]), where network slicing can be the enabler. Additionally, optimizing deep network queues for video traffic transmission can mitigate the late arrivals that cause large latencies over short periods and result in significant performance gains [20]. Significant research in improving bufferbloating within cellular networks through smart queue management schemes already exists [15], which might prove beneficial for RP. Utilizing in-network computation to perform pre-processing of video and control stream within cellular infrastructure also shows promise [18, 21, 38, 39]. Moreover, utilizing multiple access links towards the ground station, e.g. multiple cellular operators or satellites, through multipath transport (such as MPTCP or MP-QUIC) can help improve the reliability of transmissions when one of the underlying networks is experiencing deteriorations [7, 40, 49]. Furthermore, the adaptive video CC algorithms can be further improved to reduce their sensitivity towards fluctuations over the last-mile wireless, and the ramp-up phase could be shortened to ensure seamless video playback. We explore extension possibilities in Appendix A.4.

6 CONCLUSION

Due to the technological advancements in automation and emission-free operations within the aviation industry, several use cases incorporating Unmanned Aerial Vehicles (UAVs) have gained popularity in recent years. The most prominent amongst them is Remote Piloting (RP) UAVs over a network, which is a cornerstone of applications within the medical, military, and logistics sectors. In this work, we conducted an extensive measurement study investigating the state

of current cellular standards for supporting the stringent operational requirements imposed by RP. Specifically, we performed ≈ 90 UAV flights, at altitudes up to 120 m, in an urban and rural area in and around a large metropolitan European city (Munich, Germany) and measured the performance of a video stream from the drone to a remote server over LTE. We used two different Congestion Control (CC) algorithms for adaptive streaming, namely Self-Clocked Rate Adaptation for Multimedia (SCReAM) and Google Congestion Control (GCC), and correlated the results of their performance to a constant bitrate stream and changes in the underlying network conditions as the UAV performed common RP maneuvers. Our results show that the cellular network becomes increasingly unpredictable in the air, experiencing more frequent Handovers (HOs) and excessive Handover Execution Times (HETs). However, the performance is significantly dependent on the Base Station (BS) coverage, as we found negligible impact on video stream performance in urban areas compared to rural. We also found that low-latency video delivery between full-HD and 4K resolution can be maintained up to about 95% of the time in the air over LTE, and sufficient video quality for RP can be maintained up to 99.63%. In this case, it might be beneficial to utilize adaptive streaming in rural environments where link capacity is usually a bottleneck due to lack of available BSs. Our findings showcase that already-existing cellular infrastructure (read LTE) can be utilized for RP operations without needing to deploy dedicated infrastructure to provide connectivity (up to 120 m altitude), provided abundant cellular coverage exists. However, HO parameters and the core networks must be optimized for RP traffic to meet the stringent Quality of Service (QoS) requirements with sufficient network resources. In this case, utilizing techniques such as network slicing show promise and might have a crucial role in widespread support of the use case in the future. Our results also indicate a potential for multipath networking (utilizing parallel links to multiple BSs simultaneously), motivating further research in this direction.

ACKNOWLEDGMENTS

This work was partly supported by the EU CELTIC-NEXT project PICCOLO (Contract No. C2019/2-2) by the German Federal Ministry of Economics and Climate (BMWK) and managed by the project agency of the German Aerospace Center (DLR) (Contract No. 01MT20005A), Federal Ministry of Education and Research of Germany (BMBF) project 6G-Life (16KISK002), Bavarian Ministry of Economic Affairs project 6G Future Lab Bavaria and the Volkswagenstiftung Niedersächsisches Vorab (Funding No. ZN3695).

REFERENCES

- [1] 3GPP. 2016. *Evolved Universal Terrestrial Radio Access (E-UTRA); Study on latency reduction techniques for LTE (Release 14)*. Technical Report TR 36.881. [Online]. Available at: https://www.3gpp.org/ftp/Specs/archive/36_series/36.881/ [Accessed on 27/10/2021].
- [2] 3GPP. 2017. *Annex H: Field Trials Results on Mobility*. Technical Report TR 36.777. [Online]. Available at: https://www.3gpp.org/ftp/Specs/archive/36_series/36.777/ [Accessed on 27/10/2021].
- [3] 3GPP. 2018. *Study on New Radio (NR) to Support Non Terrestrial Networks (Release 15)*. Technical Report 38.811. [Online]. Available at: https://www.3gpp.org/ftp/specs/archive/38_series/38.811/ [Accessed on 16/01/2021].
- [4] 3GPP. 2022. *Technical Specification Group Radio Access Network; NR; NR and NG-RAN Overall Description; (Release 17)*. Technical Report TR 38.300. [Online]. Available at: https://www.3gpp.org/ftp/Specs/archive/38_series/38.300/ [Accessed on 11/05/2022].

- [5] 5GAA. 2021. *C-V2X Use Cases and Service Level Requirements Volume II*. Technical Report T-200xxx. 5GAA. [Online]. Available at: https://5gaa.org/wp-content/uploads/2021/01/5GAA_T-200116_TR-C-V2X_Use_Cases_and_Service_Level_Requirements_Vol_II_V2.1.pdf [Accessed on 24/10/2020].
- [6] Albert Abilov, Andrei Chunaev, Mohammed Amin Lamri, and Irina Kaisina. 2021. Real-Time Video Streaming with Application Layer ARQ in UAV Networks: Field Tests. In *2021 44th International Conference on Telecommunications and Signal Processing (TSP)*. 324–328. <https://doi.org/10.1109/TSP52935.2021.9522615>
- [7] Florian Aschenbrenner, Tanya Shreedhar, Oliver Gasser, Nitinder Mohan, and Jörg Ott. 2021. From Single Lane to Highways: Analyzing the Adoption of Multipath TCP in the Internet. In *2021 IFIP Networking Conference (IFIP Networking)*. 1–9. <https://doi.org/10.23919/IFIPNetworking52078.2021.9472785>
- [8] AWS. 2017. Keeping Time With Amazon Time Sync Service. [Online]. Available at: <https://aws.amazon.com/blogs/aws/keeping-time-with-amazon-time-sync-service/> [Accessed on 03/05/2022].
- [9] Manlio Bacco, Pietro Cassara, Alberto Gotta, and Vincenzo Pellegrini. 2019. Real-Time Multipath Multimedia Traffic in Cellular Networks for Command and Control Applications. In *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. 1–5. <https://doi.org/10.1109/VTCFall.2019.8891090>.
- [10] Aygün Baltacı, Hendrik Cech, Nitinder Mohan, Fabien Geyer, Vaibhav Bajpai, Jörg Ott, and Dominic Schupke. 2022. Analyzing Real-time Video Delivery over Cellular Networks for Remote Piloting Aerial Vehicles: Code. <https://github.com/hendrikcech/imc22-remote-piloting>
- [11] Aygün Baltacı, Hendrik Cech, Nitinder Mohan, Fabien Geyer, Vaibhav Bajpai, Jörg Ott, and Dominic Schupke. 2022. Analyzing Real-time Video Delivery over Cellular Networks for Remote Piloting Aerial Vehicles: Dataset. <https://doi.org/10.14459/2022mp1687221>
- [12] Aygün Baltacı, Ergin Dinc, Mustafa Ozger, Abdulrahman Alabbasi, Cicek Cavdar, and Dominic Schupke. 2021. A Survey of Wireless Networks for Future Aerial Communications (FACOM). *IEEE Communications Surveys and Tutorials* (2021), 1–1. <https://doi.org/10.1109/COMST.2021.3103044>.
- [13] Bundesnetzagentur. 2022. Elektromagnetische Felder (EMF). Retrieved March 12, 2022 from <https://www.bundesnetzagentur.de/DE/Vportal/TK/Funktechnik/EMF/start.html>
- [14] Gaetano Carlucci, Luca De Cicco, Stefan Holmer, and Saverio Mascolo. 2016. Analysis and Design of the Google Congestion Control for Web Real-Time Communication (WebRTC). In *Proceedings of the 7th International Conference on Multimedia Systems (Klagenfurt, Austria) (MMSys '16)*. Association for Computing Machinery, New York, NY, USA, Article 13, 12 pages. <https://doi.org/10.1145/2910017.2910605>
- [15] Bufferbloat community. 2022. Bufferbloat. [Online]. Available at: <https://www.bufferbloat.net/projects/> [Accessed on 15/09/2022].
- [16] Ioan-Sorin Comşa, Gabriel-Miro Muntean, and Ramona Trestian. 2021. An Innovative Machine-Learning-Based Scheduling Solution for Improving Live UHD Video Streaming Quality in Highly Dynamic Network Environments. *IEEE Transactions on Broadcasting* 67, 1 (2021), 212–224. <https://doi.org/10.1109/TBC.2020.2983298>
- [17] Lorenzo Corneo, Maximilian Eder, Nitinder Mohan, Aleksandr Zavodovski, Suzan Bayhan, Walter Wong, Per Gunningberg, Jussi Kangasharju, and Jörg Ott. 2021. Surrounded by the Clouds: A Comprehensive Cloud Reachability Study. In *Proceedings of the Web Conference 2021 (Ljubljana, Slovenia) (WWW '21)*. Association for Computing Machinery, New York, NY, USA, 295–304. <https://doi.org/10.1145/3442381.3449854>
- [18] Lorenzo Corneo, Nitinder Mohan, Aleksandr Zavodovski, Walter Wong, Christian Rohner, Per Gunningberg, and Jussi Kangasharju. 2021. (How Much) Can Edge Computing Change Network Latency?. In *2021 IFIP Networking Conference (IFIP Networking)*. 1–9. <https://doi.org/10.23919/IFIPNetworking52078.2021.9472847>
- [19] The Khang Dang, Nitinder Mohan, Lorenzo Corneo, Aleksandr Zavodovski, Jörg Ott, and Jussi Kangasharju. 2021. Cloudy with a Chance of Short RTTs: Analyzing Cloud Connectivity in the Internet. In *Proceedings of the 21st ACM Internet Measurement Conference (Virtual Event) (IMC '21)*. Association for Computing Machinery, New York, NY, USA, 62–79. <https://doi.org/10.1145/3487552.3487854>
- [20] Wladimir Gonçalves de Moraes, Carlos Eduardo Maffini Santos, and Carlos Marcelo Pedrosa. 2022. Application of active queue management for real-time adaptive video streaming. *Telecommunication Systems* 79, 2 (2022), 261–270.
- [21] Aaron Yi Ding, Ella Peltonen, Tobias Meuser, Atakan Aral, Christian Becker, Schahram Dustdar, Thomas Hiessl, Dieter Kranzlmüller, Madhusanka Liyanage, Setareh Maghsudi, Nitinder Mohan, Jörg Ott, Jan S. Rellermeier, Stefan Schulte, Henning Schulzrinne, Gürkan Solmaz, Sasu Tarkoma, Blesson Varghese, and Lars Wolf. 2022. Roadmap for Edge AI: A Dagstuhl Perspective. *SIGCOMM Comput. Commun. Rev.* 52, 1 (mar 2022), 28–33. <https://doi.org/10.1145/3523230.3523235>
- [22] European Union Aviation Safety Agency (EASA). 2022. Provisions applicable to both 'open' and 'specific' category. [Online]. Available at: <https://www.easa.europa.eu/the-agency/faqs/drones-uas>.
- [23] Mathis Engelbart. 2021. RTP over QUIC. [Online]. Available at: <https://github.com/mengelbart/rtp-over-quick-draft> [Accessed on 03/05/2022].
- [24] Mathis Engelbart and Jörg Ott. 2021. Congestion Control for Real-Time Media over QUIC (EPIQ '21). Association for Computing Machinery, New York, NY, USA, 1–7. <https://doi.org/10.1145/3488660.3493801>
- [25] Aymen Fakhreddine, Christian Bettstetter, Samira Hayat, Raheeb Muzaffar, and Driton Emini. 2019. Handover Challenges for Cellular-Connected Drones (DroNet'19). Association for Computing Machinery, New York, NY, USA, 9–14. <https://doi.org/10.1145/3325421.3329770>.
- [26] Boris Galkin, Erika Fonseca, Gavin Lee, Conor Duff, Marvin Kelly, Edward Emmanuel, and Ivana Dusparic. 2021. Experimental Evaluation of a UAV User QoS from a Two-Tier 3.6GHz Spectrum Network. arXiv:2011.03236 [cs.NI] <https://arxiv.org/pdf/2011.03236.pdf>.
- [27] Mohammed Gharib, Shashidhar Nandadapu, and Fatemeh Afghah. 2021. An Exhaustive Study of Using Commercial LTE Network for UAV Communication in Rural Areas. In *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*. 1–6. <https://doi.org/10.1109/ICCWorkshops50388.2021.9473547>.
- [28] gstreamer. 2022. gstreamer open source multimedia framework. Retrieved March 12, 2022 from <https://gstreamer.freedesktop.org/>
- [29] Samira Hayat, Christian Bettstetter, Aymen Fakhreddine, Raheeb Muzaffar, and Driton Emini. 2019. An Experimental Evaluation of LTE-A Throughput for Drones (DroNet'19). Association for Computing Machinery, New York, NY, USA, 3–8. <https://doi.org/10.1145/3325421.3329765>.
- [30] Youtube Help. 2022. System requirements. [Online]. Available at: <https://support.google.com/youtube/answer/78358?hl=en> [Accessed on 26/04/2022].
- [31] Stefan Holmer, Magnus Flodman, and Erik Sprang. 2015. *RTP Extensions for Transport-wide Congestion Control*. Internet-Draft draft-holmer-rmcat-transport-wide-cc-extensions-01. Internet Engineering Task Force. <https://datatracker.ietf.org/doc/html/draft-holmer-rmcat-transport-wide-cc-extensions-01> Work in Progress.
- [32] ICAO ICAO. Unknown. *REMOTELY PILOTED AIRCRAFT SYSTEM (RPAS) CONCEPT OF OPERATIONS (CONOPS) FOR INTERNATIONAL IFR OPERATIONS*. Technical Report xx. [Online]. Available at: <https://www.icao.int/safety/UA/Documents/ICAO%20RPAS%20CONOPS.pdf> [Accessed on 29/08/2022].
- [33] Haiqing Jiang, Zeyu Liu, Yaogong Wang, Kyunghan Lee, and Injong Rhee. 2012. Understanding Bufferbloat in Cellular Networks. In *Proceedings of the 2012 ACM SIGCOMM Workshop on Cellular Networks: Operations, Challenges, and Future Design (Helsinki, Finland) (CellNet '12)*. Association for Computing Machinery, New York, NY, USA, 1–6. <https://doi.org/10.1145/2342468.2342470>
- [34] Jiaqi Jin, Junjie Ma, Limin Liu, Linhai Lu, Guotong Wu, Deqing Huang, and Na Qin. 2021. Design of UAV video and control signal real-time transmission system based on 5G network. In *2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA)*. 533–537. <https://doi.org/10.1109/ICIEA51954.2021.9516048>.
- [35] Ingemar Johansson. 2014. Self-Clocked Rate Adaptation for Conversational Video in LTE (CSWS '14). Association for Computing Machinery, New York, NY, USA, 51–56. <https://doi.org/10.1145/2630088.2631976>
- [36] Severin Kacianka and Hermann Hellwagner. 2015. Adaptive Video Streaming for UAV Networks (MoViD '15). Association for Computing Machinery, New York, NY, USA, 25–30. <https://doi.org/10.1145/2727040.2727043>.
- [37] Marin, Benoit Michau, and C. Shawn. 2022. QCSuper. Retrieved March 12, 2022 from <https://github.com/P1sec/QCSuper>
- [38] Patrick McEnroe, Shen Wang, and Madhusanka Liyanage. 2022. A Survey on the Convergence of Edge Computing and AI for UAVs: Opportunities and Challenges. *IEEE Internet of Things Journal* 9, 17 (2022), 15435–15459. <https://doi.org/10.1109/IJOT.2022.3176400>
- [39] Nitinder Mohan, Lorenzo Corneo, Aleksandr Zavodovski, Suzan Bayhan, Walter Wong, and Jussi Kangasharju. 2020. Pruning Edge Research with Latency Shears. In *Proceedings of the 19th ACM Workshop on Hot Topics in Networks (Virtual Event, USA) (HotNets '20)*. Association for Computing Machinery, New York, NY, USA, 182–189. <https://doi.org/10.1145/3422604.3425943>
- [40] Nitinder Mohan, Tanya Shreedhar, Aleksandr Zavodovski, Jussi Kangasharju, and Sanjit K. Kaul. 2019. Is two greater than one?: Analyzing Multipath TCP over Dual-LTE in the Wild. *CoRR* abs/1909.02601 (2019). arXiv:1909.02601 <http://arxiv.org/abs/1909.02601>
- [41] Arvind Narayanan, Xumiao Zhang, Ruiyang Zhu, Ahmad Hassan, Shuwei Jin, Xiao Zhu, Xiaoxuan Zhang, Denis Rybkin, Zhengxuan Yang, Zhuoqing Morley Mao, Feng Qian, and Zhi-Li Zhang. 2021. A Variegated Look at 5G in the Wild: Performance, Power, and QoE Implications. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference (Virtual Event, USA) (SIGCOMM '21)*. Association for Computing Machinery, New York, NY, USA, 610–625. <https://doi.org/10.1145/3452296.3472923>
- [42] NextGEN. 2020. *Concept of Operations V2.0 - Unmanned Aircraft System (UAS) Traffic Management (UTM)*. Technical Report. FAA and NASA. [Online]. Available at: https://www.faa.gov/uas/research_development/traffic_management/media/UTM_ConOps_v2.pdf [Accessed on: 28/10/2021].
- [43] Binh Nguyen, Arijit Banerjee, Vijay Gopalakrishnan, Sneha Kaseria, Seungjoon Lee, Aman Shaikh, and Jacobus Van der Merwe. 2014. Towards Understanding TCP Performance on LTE/EPC Mobile Networks. In *Proceedings of the 4th Workshop on All Things Cellular: Operations, Applications, and Challenges (Chicago, Illinois, USA) (AllThingsCellular '14)*. Association for Computing Machinery, New York, NY, USA, 41–46. <https://doi.org/10.1145/2627585.2627594>

- [44] Yueyang Pan, Ruihan Li, and Chenren Xu. 2022. The First 5G-LTE Comparative Study in Extreme Mobility. *Proc. ACM Meas. Anal. Comput. Syst.* 6, 1, Article 20 (Feb 2022), 22 pages. <https://doi.org/10.1145/3508040>
- [45] Ericsson Research. 2015. SCReAM. [Online]. Available at: <https://github.com/EricssonResearch/scream> [Accessed on 03/05/2022.].
- [46] Ricardo Sacoto-Martins, João Madeira, J. P. Matos-Carvalho, Fábio Azevedo, and Luis M. Campos. 2020. Multi-purpose Low Latency Streaming Using Unmanned Aerial Vehicles. In *2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*. 1–6. <https://doi.org/10.1109/CSNDSP49049.2020.9249562>
- [47] Zaheduzzaman Sarker, Colin Perkins, Varun Singh, and Michael A. Ramalho. 2021. RTP Control Protocol (RTP) Feedback for Congestion Control. RFC 8888. <https://doi.org/10.17487/RFC8888>
- [48] Matko Silic, Mirko Suznjevic, and Lea Skorin-Kapov. 2021. QoE Assessment of FPV Drone Control in a Cloud Gaming Based Simulation. In *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*. 175–180. <https://doi.org/10.1109/QoMEX51781.2021.9465385>
- [49] Hassan Sinky, Bechir Hamdaoui, and Mohsen Guizani. 2016. Proactive Multipath TCP for Seamless Handoff in Heterogeneous Wireless Access Networks. *IEEE Transactions on Wireless Communications* 15, 7 (2016), 4754–4764. <https://doi.org/10.1109/TWC.2016.2545656>
- [50] Jędrzej Stanczak, Istvan Z. Kovacs, Dawid Koziol, Jeroen Wigard, Raphael Amorim, and Huan Nguyen. 2018. Mobility Challenges for Unmanned Aerial Vehicles Connected to Cellular LTE Networks. In *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. 1–5. <https://doi.org/10.1109/VTCSpring.2018.8417736>
- [51] Antonia Stornig, Aymen Fakhreddine, Hermann Hellwagner, Petar Popovski, and Christian Bettstetter. 2021. Video Quality and Latency for UAV Teleoperation over LTE: A Study with ns3. In *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. 1–7. <https://doi.org/10.1109/VTC2021-Spring51267.2021.9448676>
- [52] A. Takacs, X. Lin, S. Hayes, and E. Tajedor. 2018. *Drones and networks: Ensuring safe and secure operations*. Technical Report GFMC-18:000526. Ericsson. [Online]. Available at: https://www.ericsson.com/4adfc0/assets/local/reports-papers/white-papers/10201110_wp_dronesandmobilenetworks_nov18.pdf [Accessed on 19/05/2022.].
- [53] Guangmin Tang, Yibo Hu, Hong Xiao, Liangqian Zheng, Xingbin She, and Na Qin. 2021. Design of Real-time video transmission system based on 5G network. In *2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA)*. 522–526. <https://doi.org/10.1109/ICIEA51954.2021.9516191>
- [54] International Telecommunication Union. 2009. *Characteristics of unmanned aircraft systems and spectrum requirements to support their safe operation in non-segregated airspace*. Technical Report M.2171. ITU. [Online]. Available at: <https://www.itu.int/en/ITU-R/space/snl/Documents/R-REP-M.2171-2009-PDF-E.pdf> [Accessed on 03/09/2022.].
- [55] Bertold Van Der Bergh, Alessandro Chiumiento, and Sofie Pollin. 2016. LTE in the sky: trading off propagation benefits with interference costs for aerial nodes. *IEEE Communications Magazine* 54, 5 (2016), 44–50. <https://doi.org/10.1109/MCOM.2016.7470934>
- [56] VideoLAN. 2018. x264. [Online]. Available at: <https://www.videolan.org/developers/x264.html> [Accessed on 17/05/2022.].
- [57] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612. <https://doi.org/10.1109/TIP.2003.819861>
- [58] Xuedou Xiao, Wei Wang, Taobin Chen, Yang Cao, Tao Jiang, and Qian Zhang. 2020. Sensor-Augmented Neural Adaptive Bitrate Video Streaming on UAVs. *IEEE Transactions on Multimedia* 22, 6 (2020), 1567–1576. <https://doi.org/10.1109/TMM.2019.2945167>
- [59] Fang Yang, Honggui Deng, Fangqing Jiang, and Xu Deng. 2015. Handover Optimization Algorithm in LTE High-Speed Railway Environment. *Wireless Personal Communications* 84 (2015), 1577–1589.
- [60] Tao Yu, Shunya Imada, Kiyomichi Araki, and Kei Sakaguchi. 2021. Multi-UAV Full-Duplex Communication Systems for Joint Video Transmission and Flight Control. In *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*. 1423–1428. <https://doi.org/10.1109/CCWC51732.2021.9375952>
- [61] Hui Zhou, Fenghe Hu, Michal Juras, Ashish B. Mehta, and Yansha Deng. 2021. Real-Time Video Streaming and Control of Cellular-Connected UAV System: Prototype and Performance Evaluation. *IEEE Wireless Communications Letters* 10, 8 (2021), 1657–1661. <https://doi.org/10.1109/LWC.2021.3076415>

A APPENDICES

A.1 Ethics

We conducted the UAV flights following the relevant laws and regulations. We also ensured a safe distance between the flying UAV and bystanders. We conducted our experiments with unaltered

and commercial modems over the cellular network of two public mobile network operators. Furthermore, we contacted the network providers and got permission to access their network in-flight. This work does not raise any ethical issues to the best of our knowledge.

A.2 Flight Trajectory during Measurements

Figure 11 shows our flight trajectory. We designed this trajectory to represent the basic movements of UAV during takeoff, landing, and cruise. In each flight, we lift off and fly vertically up to 40 m, complete a horizontal leap of ≈ 200 m and repeat this procedure at 80 m and 120 m. After the last leap at 120 m, we descent straight down to the takeoff location. Overall, our air time per flight is ≈ 6 minutes.

A.3 MNO Performance

Figure 12 shows the diversity in available link capacity from different MNOs and its influence on the achievable video delivery performance. Note that operator P1 is the default MNO used throughout our study while operator P2 is a competing provider in the region. However, P2 connection offered us a larger bandwidth cap (500 Mbps downlink + 50 Mbps uplink) compared to P1 (300 Mbps downlink + 50 Mbps uplink). Also note that while P1 and P2 had similar deployment density in urban region, the density of P1 is significantly less compared to P1 in rural region.

As the operator P2 provided more link capacity than that of P1, achieved goodput rates over both operators significantly differ in Figure 12(a). Consequently, the quality of the received frames could improve as can be seen in the Structural SIMilarity (SSIM) analysis in Figure 12(d). However, larger capacity did not necessarily improve the playback latency and Frames Per Second (FPS) performance since SCReAM performed significantly poorer with the operator P2 at higher bitrates. This is mainly due to the limitation on the acknowledgement mechanism of SCReAM at high bitrates as we discuss in Section 4.2.1.

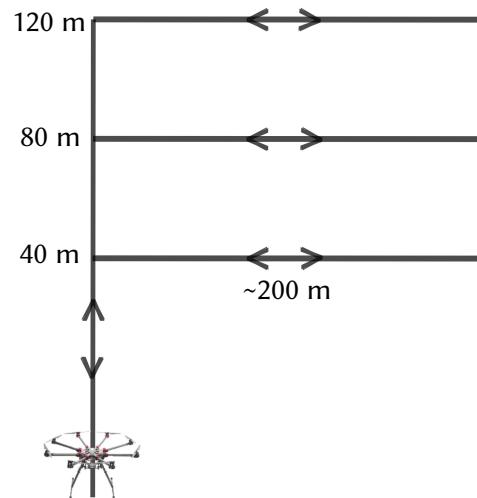


Figure 11: The flight trajectory of our experiments.

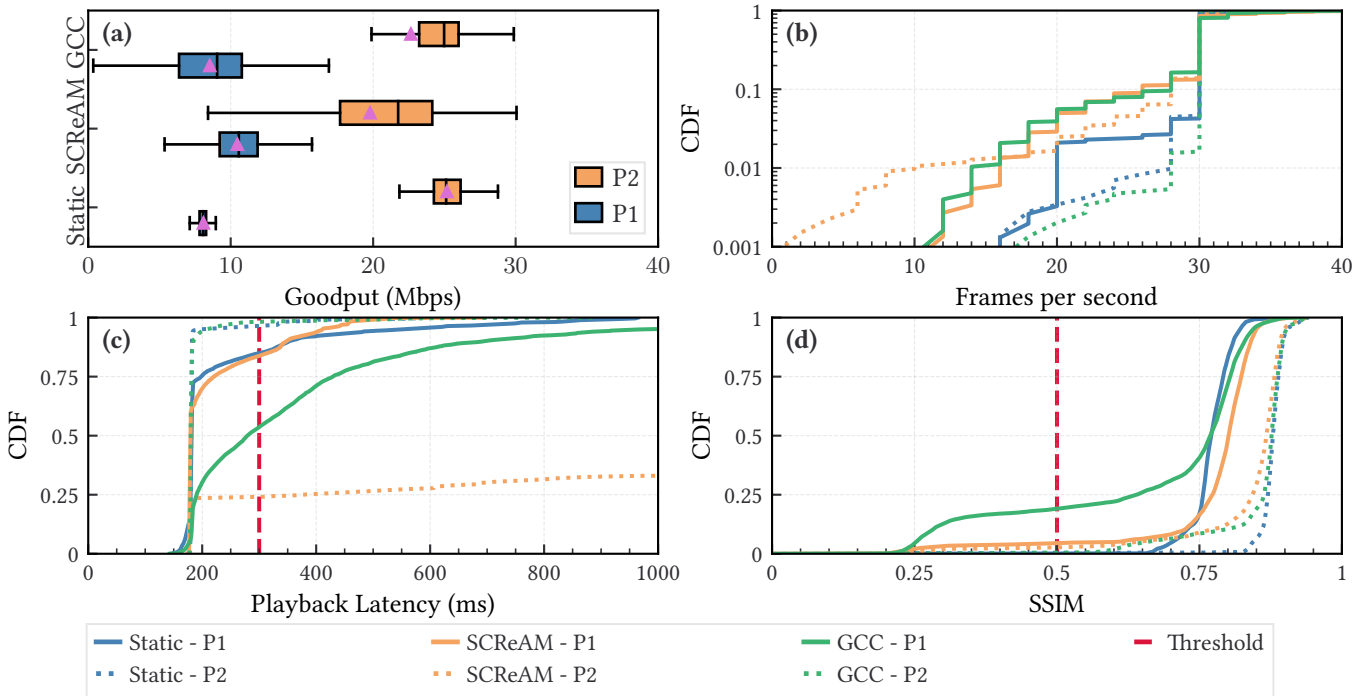


Figure 12: The video delivery performance with respect to Mobile Network Operators (MNOs) in the rural environment. This analysis highlights the influence of available link capacity of the MNOs on the achievable video delivery performance.

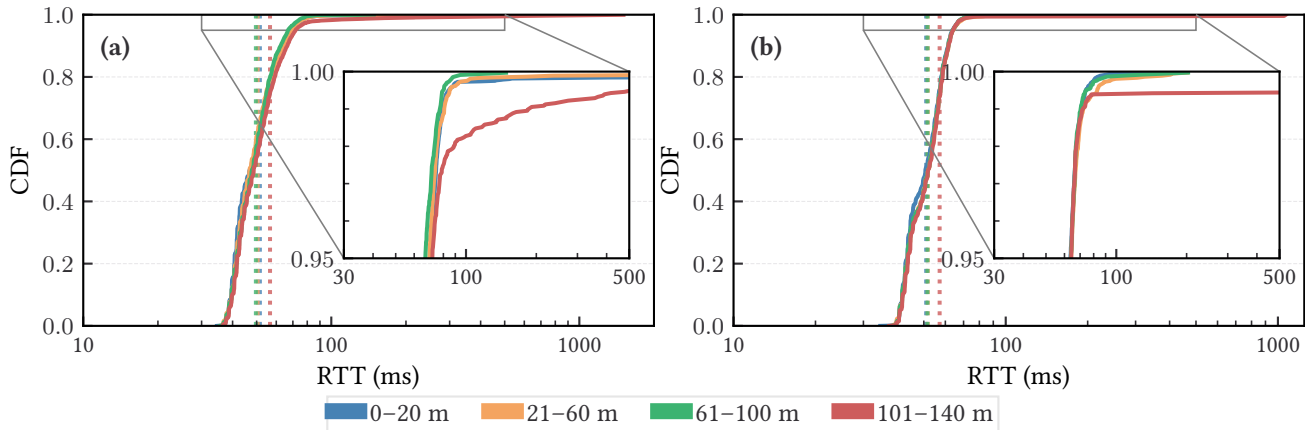


Figure 13: The RTT measured by ICMP pings at different altitudes in the (a) urban and (b) rural environments without cross traffic. No clear trend is discernible below 100 m. Above that, the proportion of high RTT outliers increases.

A.4 Potential Video Pipeline Improvement for Future Studies

The GStreamer components that constitute our video pipeline optimize for a pleasant viewing experience under link congestion. The playback speed reduces proactively when the video buffer runs low to avoid freezes and sudden playback speed changes. Once the delayed packets arrive, the playback speed increases to cut down on the elevated playback latency until the baseline is reached again.

For RP operations, a different strategy might be better suited. Instead of manipulating the playback speed, the pilot could always see the most recently recorded frame as it contains the most relevant information. Such an operation could be possible by setting the “drop-on-latency” property on the `rtptimebuffer` element in our GStreamer pipeline that drops all frames older than a threshold from the buffer. This property might decrease the time it takes the playback latency to come down to the baseline after a playback latency spike. The decision over how to buffer and play frames should be based on evaluating the pilots’ needs.