Principal component analysis on atmospheric noise measured with an integrated superconducting spectrometer

Yashoda Sewnarain Sukul

Bachelor thesis
July 2019
Supervisor: Dr. A. Endo
Second examiner: Dr. A.J.L. Adam
Third examiner: S.A. Hähnle

TU Delft

Abstract

The Deep Spectroscopic High-redshift Mapper, or DESHIMA, is an integrated superconducting spectrometer which measures the redshift of photons originating from submillimeter galaxies. These photons have to travel through the Earth's atmosphere before arriving at DESHIMA, but this atmosphere adds noise to the signal. Using a principal component analysis on still-sky measurement data, the influence of the atmospheric noise on these observations is analyzed.

To achieve this, several principal component analyses are performed on still-sky observations, which are measurements of the brightness temperature of the sky, T_{sky}. These observations are assumed to be governed by three noise types: atmospheric noise, photon noise, and detector $1/f$ noise. A PCA on a still-sky observation reveals the effect of the most dominant noise sources. By creating an artificial data set, the physical origin of these noise sources can be found. This data set was produced by using an existing atmospheric model, based on the fluctuation of the precipitable water vapour, or PWV, in the atmosphere.

A principal component analysis on this artificial data set reveals the effect of this PWV fluctuation on the data. The first principal component of the artificial data is found to represent the derivatives of the T_{sky}-PWV relations for every channel. The second principal component has a non-zero explained variance and is found to represent the second-order derivatives of the T_{sky}-PWV relations for every channel, indicating that these relations are not linear. This can be explained by performing a Taylor expansion on the T_{sky}-PWV relation. Comparing the principal components of the real data to those of the artificial data shows that the first principal component generally has the same shape as the first principal component of the artificial data, which represents the first-order PWV fluctuation, confirming that PWV fluctuation is the most dominant noise source for still-sky observations.

It is also found that, when the PWV fluctuation is large, the second-order Taylor expansion term of the T_{sky}-PWV generally becomes more important in the real data. In this case, the first principal component has a very high explained variance, and the second-order term is usually represented by the second principal component. Conversely, when the PWV range is small, the first-order term explains significantly less variance, and random noise like the photon noise becomes more dominant than the higher-order terms.

The results show a few exceptions to this interpretation, so further research on these systematic errors is strongly recommended. In order to achieve better results, the experimental method can be improved by including the bandwidths of the channels and better estimation of the PWV fluctuation. This research can also be extended into a design of a random noise level and ultimately, the design of a better atmosphere calibration method for DESHIMA.

Contents

Chapter 1

Introduction

The Atacama desert is considered the driest place on earth. This makes it an excellent location for telescopes, as the sky is clear during the entire year. Places like this are very desirable for submillimeter astronomy, a branch of astronomy that observes radiation in the THz waveband. The THz band is defined as frequencies between 0.3 - 10 THz, the corresponding wavelengths are called submillimeter wavelengths. Distant dusty galaxies, which are optically faint, radiate in this frequency band because of the cold dust (40 - 100 K) that enshrouds them [1]. Observing these galaxies, also called submillimeter galaxies, in the THz band can tell us about star formation in these galaxies, but noise unavoidably influences these observations. It is known that part of this noise can be attributed to the atmosphere. However, while studied extensively in other fields, a clear physical interpretation of atmospheric noise in the field of wideband THz astronomy has not yet been achieved.

Figure 1.1: Picture of the ASTE telescope in the Atacama desert in Chile. The DESHIMA chip is placed in the cabin of ASTE, in a cryostat.

1.1 DESHIMA

DESHIMA (Deep Spectroscopic High-redshift Mapper) is an integrated superconducting spectrometer, and it is mounted in the ASTE (Atacama Submillimeter Telescope Experiment) telescope in the Atacama desert. DESHIMA is a wideband THz spectrometer that is used to measure the frequency of photons originating from submillimeter galaxies. The sources of submillimeter radiation from the dusty galaxies are line emissions and dust emission. From the dust emission, the total infrared luminosity can be derived, which is correlated with the star formation rate in the galaxy. Line emissions are photons that are excited when atoms or molecules transition from a higher energy state to a lower energy state. When the rest frequency is known, the observed frequencies of these line emissions can reveal the cosmological redshift of the photons. Cosmological redshift is an increase in photon wavelength caused by the expansion of the universe. Using that the redshift increases with the distance travelled, the distance to the source of the photons can be calculated. By observing the dust emission and line emissions in the THz band, new and optically faint galaxies can be localized.

An important emission line for DESHIMA is that of [CII] ($\lambda_{rest} = 158\mu m$). This line emission is prevalent when warm star-forming molecular gas is cooling down [2], and it is used because it is extremely luminous and because the redshifted line emissions are in the THz range of interest, which is radiation from submillimeter galaxies where star formation is at its peak [3]. In order to accurately measure the redshift of a galaxy, DESHIMA detects at least two emission lines. The ultimate goal of DESHIMA is to determine the distance to the new galaxies and make a 3D map of the universe.

What distinguishes DESHIMA from other existing spectrometers, is that instead of using optics, it uses a superconducting circuit for frequency sorting and detection. The integrated superconducting filterbank chip of DESHIMA consists of 49 channels that are, in fact, bandpass filters. These 49 filters are all connected to hybrid Microwave Kinetic Inductance Detectors. Microwave Kinetic Inductance Detectors, or MKIDs, detect the photons. MKIDs are superconducting detectors, and they work on the principle of pair breaking. At very low temperatures, electrons in a superconducting material pair up, as this lowers their energy. These pairs are called Cooper pairs. When a photon with large enough energy reaches the superconducting material, it can be absorbed by breaking up Cooper pairs, which can be seen in figure 1.2. This creates single electrons, also called quasiparticles. The MKID then detects the photon by measuring the amount of signal power that is absorbed in the channel.

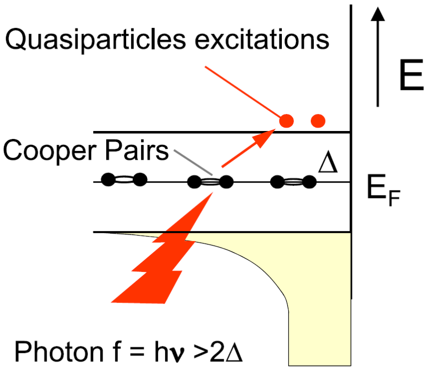

Figure 1.2: A photon with large enough energy can be absorbed by breaking Cooper pairs into quasiparticles [4]

1.2 Noise

There are three important sources of noise when it comes to detecting THz photons: photon noise, detector noise, and atmospheric noise.

1.2.1 Photon noise

The first type of noise, photon noise, is inherent to the incoming signals. Photon noise is the variation of the incident photon flux, caused by the random arrival rate of the photons at the detector. As mentioned, the detector does not detect individual photons, but it measures the absorbed signal power P. Because of the random arrival rate of the photons, this absorbed energy per second is not constant over time. The random fluctuation in P caused by this phenomenon is the photon noise.

This method of detecting photons exhibits another type of intrinsic noise. In thermal equilibrium, there is thermal generation and recombination of quasiparticles. The MKID cannot tell the difference between these thermal effects and photon absorption. This thermal generation and recombination of quasiparticles is a random process, so the noise that it adds to the actual signal is also random [4]. When the MKIDs are cold enough, and the signal power is high enough, though, the probability of thermal generation of quasiparticles is relatively low, meaning that most quasiparticle generations can be attributed to absorbed photons. The random recombination can thus be seen as the biggest source of the generation-recombination (or g-r) noise.

1.2.2 Detector $1/f$ noise

The filters and MKIDs on DESHIMA are placed on a 350 μm thick layer of c-plane sapphire, which is a dielectric material. The detector noise is inherent to the dielectrics of the detector. This noise, also called two-level system noise, originates from fluctuating two-level defect states in the dielectric material used for the chip. Two-level systems are various kinds of impurities on a dielectric material which make it easier for particles to tunnel between two states. These tunneling states act like electric dipoles, and can interact with external electric fields [5]. This interaction can change the dielectric constant of the c-plane sapphire, which in turn affects the resonance frequency of the MKIDs, and thus creates noise.

The two-level system noise spectrum is proportional to $1/f$, and it has been proven that the noise increases at low temperatures.

1.2.3 Atmospheric noise

The third type of noise is atmospheric noise. The composition of the atmosphere makes it very opaque for THz photons. There are many molecules in the atmosphere; nitrogen, water, oxygen and carbon dioxide being the most abundant. Water, especially, has hundreds of absorption lines in the THz band [6]. Even in the generally clear sky conditions of the Atacama desert, these molecules make it harder for the photons to reach the telescope unaffected.

The atmosphere contains water, mostly in the form of vapour. The depth of water vapour in a column of the atmosphere is called the precipitable water vapour, or PWV, and is usually in the order of a few mm. The amount of PWV influences the opacity of the sky, τ. This opacity can be converted into the atmospheric transmission η_{atm} with the following formula:

$$\eta_{atm} = e^{-\tau} \tag{1.1}$$

The atmospheric transmission ranges from 0 to 1, with 0 being completely opaque and 1 completely transparent. If this transmission were constant, the atmosphere would be noiseless. The transmission is not constant, however, and it is the constantly changing composition of the atmosphere above ASTE that makes the atmosphere noisy.

1.3 Goal of this research

This research will focus on DESHIMA still-sky observations, which are governed by the three types of noise mentioned above. The atmosphere is an important source of noise for DESHIMA, as THz signals are affected by the atmosphere. Currently, there are already methods to calibrate the system in order to filter out the atmospheric noise signal [7]. However, as the components of the atmospheric noise have a yet unknown physical origin, these calibration methods have some room for improvement. The goal of this research is to find out the most important factors that influence the noise signal in the still-sky observations. It is expected that the PWV value of the atmosphere is one of these important factors. In order to find out whether this is true and to see how important this factor is, a principal component analysis will be performed on still-sky data measured by DESHIMA. This analysis will output principal components, variables that are responsible for the most variance in the data, revealing the underlying mechanisms of still-sky noise.

Chapter 2

Theory

This chapter describes the underlying theory needed to understand the method and the results and the discussion. First, the working principle of DESHIMA will be explained. Then, the concept of brightness temperature and the ATM model will be illustrated. Lastly, the theory behind Principal Component Analysis will be explained.

2.1 The integrated superconducting spectrometer

The DESHIMA spectrometer measures the redshifted [CII] emission line originating from submillimeter galaxies. The first version, DESHIMA 1.0, can measure the incoming photons within a frequency range of 332-377 GHz. The design of the spectrometer chip is illustrated in figure 2.1.

First, the signal enters the chip through the lens and the antenna. Then, the signal is coupled to a small transmission line, that conducts the signal towards the filterbank. This filterbank consists of 49 channels, which are all resonators. These resonators each have a different resonance frequency in the submillimeter range. The resonators all have slightly different lengths, and since resonance only occurs if the length is half the incoming photon wavelength, we can distinguish photons of different wavelengths by observing which channels resonate. These so-called half-wavelength resonators act as narrow bandpass filters and are designed to have specific Q factors in order to achieve the desired frequency resolution. This quality factor Q is defined as the resonance frequency divided by the bandwidth of the resonator $\frac{f_r}{\Delta f}$. DESHIMA 1.0 has a Q factor of $\sim$380. The channels of DESHIMA have overlapping bandwidths so that there is no risk of gaps between the channels. This causes signal power to be shared between neighbouring channels [8].

The ground plane of the chip and the transmission line are made of NbTiN, a superconducting material with a critical temperature T_c of 15 K. At this temperature, the gap frequency $F_{2\Delta}$ of NbTiN is $\sim$ 1.1 THz. This means that signals with a frequency $f < F_{2\Delta}$ can travel through NbTiN without being absorbed. The signal, usually having one strong frequency, is intercepted by one or two resonators. The resonators are connected to aluminium absorbers, which are coupled to MKIDs on the other side. Aluminium has a gap frequency $F_{2\Delta}$ of $\sim$ 80 GHz at its critical temperature of 1.3 K, much lower than that of NbTiN. As the signals have frequencies in the range of 332-377 GHz, all of them will be absorbed in the aluminium section.

MKIDs are also resonators, and they have resonance frequencies in the range of 5.6-6.4 GHz. When signal power is absorbed in the adjoining aluminium absorber, the kinetic inductance of the MKID changes, resulting in a resonance frequency shift to a lower value. This shift can be detected and measured in the readout line. MKIDs are thus used to measure the absorbed power. When it has been established which channels have absorbed signal power, the frequencies of the incoming photons can be derived [8].

In order to keep the number of thermal quasiparticle generations to a minimum, the chip must be cooled down. Because of the large gap frequency, there will be no thermal quasiparticle generations in the NbTiN components of the chip. The aluminium components, however, have a relatively small gap frequency and corresponding critical temperature T_c of 1.1 K. In order to make sure that the quasiparticle excitations can primarily be attributed to the absorption of photons, the chip must be cooled down to <0.2 K. Taking this and the saturation of the quasiparticle density at low temperatures into account, Endo et al. [3] have argued that DESHIMA should operate at a temperature between 100-200 mK. In order to reach such low temperatures, the chip must be placed in a cryostat while operating.

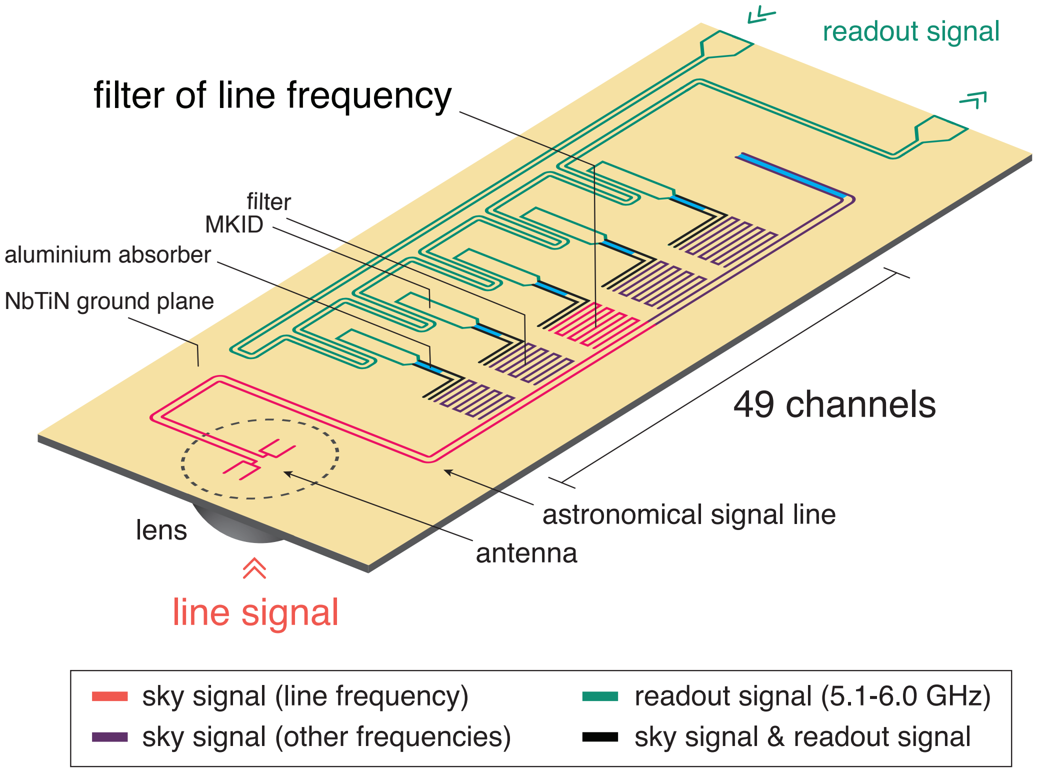

Figure 2.1: Design of the filterbank spectrometer.

2.2 Brightness temperature

By using power measurements of the detectors in the telescope, the total emitted power from a distant galaxy can be derived. This total emitted power is also called the luminosity L of a galaxy. Related to the luminosity is the surface brightness B, which is the intensity at the surface of a radiating object, equal to the power emitted at the surface per unit area (normal to direction $\mathbf{n}$) per unit solid angle $d\Omega$ [9]. This is illustrated in figure 2.2.

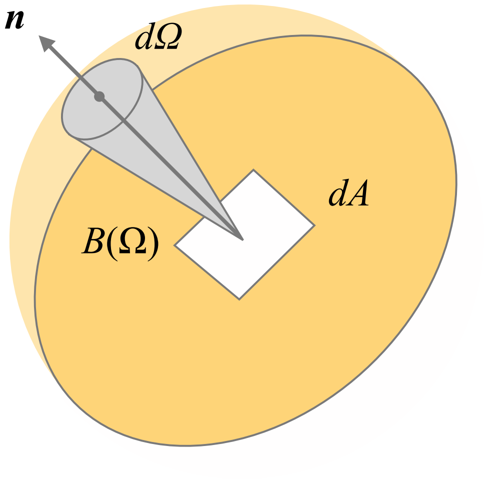

Figure 2.2: Surface brightness of a radiating object [9]

Brightness temperature is the surface brightness of the source converted into Kelvin. This is done by using Planck's law:

$$T_b = \frac{c^2}{2f^2k} B_f \tag{2.1}$$

In this equation, c is the speed of light in m/s, f is the observed frequency in Hz, k is the Boltzmann constant, and B_f is the surface brightness per unit frequency bandwidth. This equation is only valid in the Rayleigh-Jeans limit, when $hf \ll kT$, with h Planck's constant.

Temperature brightness is not only used to describe the luminosity of distant galaxies, but it can also be used to describe the intensity of the atmosphere. When the sky signal is measured, this observed intensity is essentially a measure for the signal noise. The brightness temperature of the sky T_{sky}, observed by detectors in a telescope, is given by:

$$T_{sky} = T_{atm}(1 - \eta_{atm}) = T_{atm}(1 - \eta_{zenith}{}^{\frac{1}{sin(el)}}) \tag{2.2}$$

The temperature T_{atm} is the temperature of the atmosphere, which is taken to be equal to the ambient temperature T_{amb}. The variable η_{atm} is the transmission of the sky. The value of η_{atm} is related to the amount of water in the atmosphere, also called the PWV. η_{zenith} is the transmission of the atmosphere when the telescope points vertically upwards. When the telescope is directed along an angle to the horizon $\neq 90°$, η_{atm} is not equal to η_{zenith}, as the path through the atmosphere for photons is longer. This angle, el in equation (2.2), is called the elevation angle of the telescope. Using the ATM model, which is described in the next section, η_{zenith} can be calculated for a given frequency range and PWV value.

2.3 The ATM model

The ATM model is a tool used to determine the atmospheric transmission spectrum for dry and clear atmosphere conditions like in the Atacama Desert (typical PWV of ∼1 mm [10]), created especially for submillimeter astronomy. The model is not based on real data, but uses mathematical models to calculate the effects of absorption and phase delay by the H_2O, O_2, N_2 and O_3 molecules in the atmosphere [11]. Using this model, a spectrum of the atmospheric transmission can be calculated for a given frequency range and PWV value.

When considering the range of the frequencies of the photons that reach DESHIMA (332-377 GHz), the ATM model outputs the spectra for six different PWV values, shown in figure 2.3. This spectrum applies to the location and weather conditions of the ALMA telescope. As the ASTE telescope is situated next to ALMA, it is reasonably safe to assume that the same spectrum also applies for ASTE [10].

It can be seen here that the transparency rapidly drops to zero around 368 GHz. This feature seems to be independent of the PWV value. This is caused by a water absorption line [11].

ALMA measures the zenith transmission η_{zenith}, which is only valid for ASTE when it is pointing vertically upwards. When this is not the case, the real transmission η_{atm}, is lower than the zenith transmission, as the path through the atmosphere to the telescope is longer for photons.

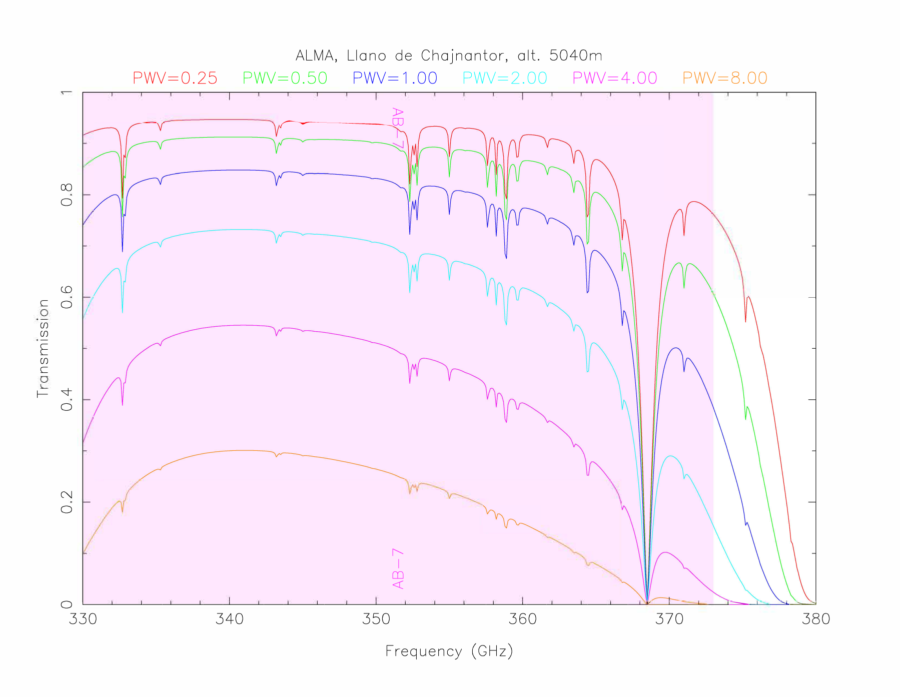

Figure 2.3: Transparency spectrum for different values of PWV in the DESHIMA frequency range [10].

2.4 PCA

Principal component analysis is a statistical technique used to reduce the dimensionality of a large data set. Large data tables with multiple variables can very often be represented by smaller data sets. The reason for this is that some of these variables are intercorrelated. PCA computes the principal components, which are orthogonal variables and thus uncorrelated. By choosing a reasonable number of principal components for a data set, the dimensionality of the data set can be reduced while still accounting for most of the variance in the original data.

2.4.1 Performing PCA

The essential goal of PCA is to find the principal components. In order to do this, the data must be represented in a table of $M \times N$ values. The columns of this table are the different features (or variables) of the data set and the rows are the different observations of these features. The principal components are actually linear combinations of the elements in the data table. For a data table $\mathbf{X}$, the first principal component can be written as [12]:

$$\mathbf{X}\bar{\alpha_1} = \alpha_{11}\mathbf{X_1} + \alpha_{12}\mathbf{X_2} + ... + \alpha_{1N}\mathbf{X_N} = \sum_{j=1}^{N} \alpha_{1j}\mathbf{X_j} \tag{2.3}$$

In this equation, the $\mathbf{X_j}$ are the columns of the data table. The first of these linear combinations, $\alpha_1'x$, has maximum variance. The second one, $\alpha_2'x$, is uncorrelated with $\alpha_1'x$ and has the second-largest variance. Theoretically, N of these linear combinations can be found. As will be seen, however, the first couple of principal components already contain most of the variance. Thus, the data set can be reduced. The following paragraphs will explain how this works.

First, the covariance matrix $\mathbf{C}$ of the data table is calculated using equation (2.4):

$$\mathbf{C} = E[\mathbf{X}^T\mathbf{X}] - {\mu_X}^T \mu_X \tag{2.4}$$

In this equation, μ_X is the mean $E[\mathbf{X}]$. The matrix $\mathbf{X}$ can be written as a vector of column vectors: $\mathbf{X} = (X_1, ..., X_N)$. Each column vector represents one variable.
In the matrix $\mathbf{C}$, the diagonal elements are the covariances of the individual variables with themselves. This means that the diagonal elements are equal to the variance of the corresponding variables:

$$\sigma_x^2 = E[|X_n - E[X_n]|^2] \tag{2.5}$$

The off-diagonal elements are equal to the covariances $c_{xx}(k, l)$ of two variables with each other:

$$c_{xx}(k, l) = E[(X_k - E[X_k])(X_l - E[X_l]] \tag{2.6}$$

Here, X_k and X_l are observations of two different variables, so $k \neq l$ and $k, l \leq N$. The covariance is a measure of the linear relationship between two variables. From equation (2.6), it becomes clear that the covariance $c_{xx}(k, l)$ is the same as $c_{xx}(l, k)$, making the covariance matrix $\mathbf{C}$ symmetric. The covariance matrix contains information about the magnitudes and directions of the spread along axes. After performing PCA, the covariance matrix of the new data table **X**', should have no off-diagonal elements, i.e. the covariances of the new variables should be equal to zero [13].

The next step is to find a vector that points in the direction of the largest variance in the data. This vector will be called $\bar{\alpha_1}$, in line with equation (2.3). The data can be projected along this vector by performing the matrix product $\mathbf{X}\bar{\alpha_1}$. An important prerequisite for $\bar{\alpha_1}$ is that it should maximize the variance of $\mathbf{X}\bar{\alpha_1}$. From equation (2.4), it can be derived that the variance can be written as $\bar{\alpha_1}^T\mathbf{C}\bar{\alpha_1}$. If $\bar{\alpha_1}$ is a normalized unit vector, the problem of finding it for maximum $\bar{\alpha_1}^T\mathbf{C}\bar{\alpha_1}$, can be formulated as a Rayleigh Quotient problem. Solving this problem gives that $\bar{\alpha_1}$ must be equal to the largest eigenvector of $\mathbf{C}$ [14]. The magnitude of an eigenvector is given by the corresponding eigenvalue. Hence, $\bar{\alpha_1}$ is equal to the eigenvector of $\mathbf{C}$ that belongs to the highest eigenvalue.

To find the second principal component, the same process must be repeated, but now the eigenvector corresponding to the second-highest eigenvalue, named $\bar{\alpha_2}$, must be found. An important prerequisite for this eigenvector is again that it maximizes the variance of $\bar{\alpha_2}^T\mathbf{X}$, but also that it is orthogonal to the eigenvector $\bar{\alpha_1}$. When real data is used, the expectation values of the data are real. This means that the covariance matrix $\mathbf{C}$ of a real matrix $\mathbf{X}$ is also real. It has already been established that $\mathbf{C}$ is symmetric and, fortunately, an essential property of real and symmetric matrices is that the eigenvectors belonging to distinct eigenvalues are mutually orthogonal [14].

This process can be continued up until the N-th principal component. As mentioned before, however, most of the variance is already contained in the first couple of principal components. So, the next step is to choose the number of components, p. The more components are included, the more accurate the new data set. The next chapter will explain a method to visualize the significance of each principal component, which makes it easier to choose an appropriate number of components, p. After having done this, the original data can be projected onto the new axes. This is done by making a matrix $\mathbf{P}$ which contains the p largest eigenvectors as columns. Then the new data set $\mathbf{Y}$ (size $N \times p$) is calculated with the following matrix product:

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{P} \tag{2.7}$$

2.4.2 Visualizing the principal components

An important part of PCA is the visualization of the principal components. In order to show this, an example superconducting spectrometer chip is considered. This chip has only two channels. After some scanning time, the collected data from these two channels can be represented in a data table. The columns are the variables, which are the two different channel responses, making this a two-dimensional data set. The rows are the observations of these variables at different time stamps. The next step is to calculate the covariance matrix of this table. In most cases, this 2×2 matrix will have non-zero values for all elements, indicating that the variables share some relationship. This relationship can be visualized by plotting the channel responses against each other. An example of what this could look like is shown in figure 2.4.

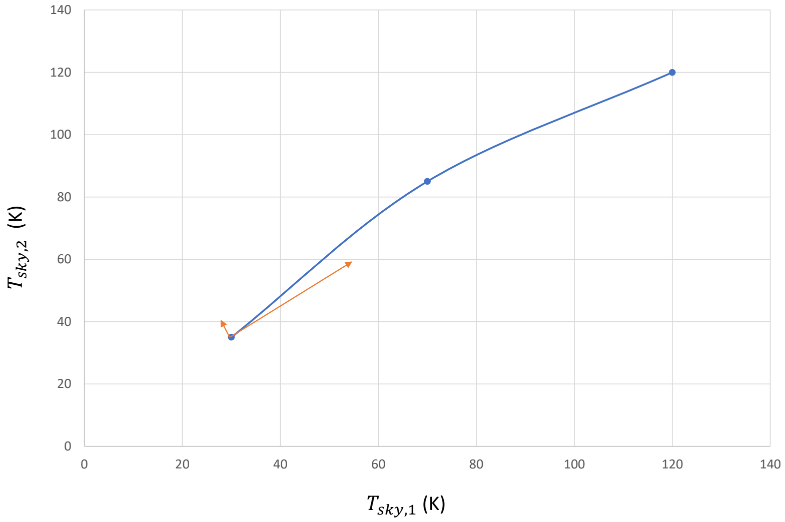

Figure 2.4: Channel responses of a two channel system at three different time stamps plotted against each other. The arrows point in the direction of the largest spread and are scaled to the magnitude of the spread in the corresponding directions.

The orange arrows point in the directions of the most spread, and their magnitude is scaled to the magnitude of the spread in the corresponding directions. These arrows represent the eigenvectors of the covariance matrix of this data set, and their magnitudes are scaled to the eigenvalues of these eigenvectors. If the relation between the channel responses were linear, only one principal component would be needed to fully describe the data set, as the spread would only be in one direction. As the relation between the channel responses is curved, however, two principal components are needed to fully describe the data set. When both eigenvectors are combined into a matrix **P**, the data set can be projected onto the principal component axes using equation (2.7). This projection can be seen as rotating the data such that the orange arrows align with the horizontal and vertical axes. Since the spread of this new data set is directed along the axes only, the new covariance matrix will only have diagonal elements.

DESHIMA has 49 channels, making the real data set 49-dimensional. This is a lot harder to visualize, but the exact same principle applies.

2.4.3 Applications of PCA

There are numerous reasons to perform PCA on a data set. The first one is to extract relevant information from a data table. For the case of this research, for example, the data table solely contains noise data. Some of this noise originates from a non-random physical mechanism, while the rest of the noise is random, like the photon noise. PCA can help separate random and non-random data.

Another application of PCA is to reduce the noise in a data set. When a data table, for example, contains data coming from a strong signal and some background noise, most of the variance can be explained by the signal. When the signal and the noise are uncorrelated, they will be represented by separate principal components. By only retaining the components that represent the signal, the noise can be reduced.

Chapter 3

Method

This chapter describes the experimental method and explains how the number of principal components for the PCA will be chosen.

3.1 Setup of the experiment

During real observations, the telescope is moving in order to scan a specific area. For this research, however, only still-sky data will be used. The main goal is to see how the atmospheric noise changes as a function of time and why it changes, so in order to simplify the problem, the rotation of the telescope is not yet taken into account. Figure 3.1 shows the general setup of the spectrometer system.

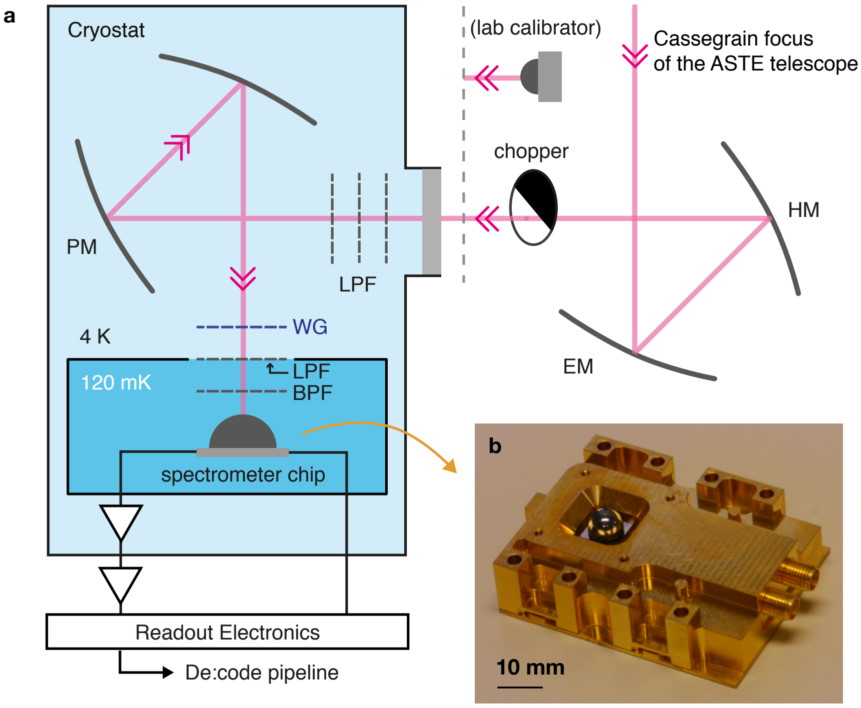

Figure 3.1: General setup for the DESHIMA spectrometer system [7]

In this figure, the telescope is pointed vertically upwards with an angle of 90° to the horizon.

This angle, called the elevation angle, can be adjusted. It must be noted that when the elevation angle is <90°, the path for photons through the atmosphere is longer. Hence, these signals will generally be noisier than signals that come in from a 90° elevation angle, as the longer path length means more absorption and thus a lower value of η_{atm} for a particular value of PWV.

The telescope can also vary its azimuth angle, which is necessary when it needs to be pointed at a specific galaxy. Unlike the elevation angle, the azimuth angle has less influence on the absolute value of η_{atm}, as the path length towards the telescope stays the same for incoming photons. When there are small local variations of PWV, however, the absolute value of η_{atm} also depends on the azimuth angle. This research will use measurement data where the elevation angle is between 60° and 88°, while the azimuth angle is not taken into account.

When they arrive at the telescope, incoming photons will be guided towards the cryostat by two mirrors. These mirrors are curved in order to focus the signal. Before entering the cryostat, the signal path moves through the chopper wheel. A chopper is used for system calibration. The chopper is a disk of which one half is opaque (black body with temperature T_amb), and the other half is transparent. If the chopper is rotating while the spectrometer is pointed at the sky and measures T_{sky}, the readout frequency changes with the frequency of the chopper rotation. By comparing the resulting outputs, the system can be calibrated, and the changes in the atmospheric absorption can be corrected. This calibration will not be used for this research, as the main point of interest is the atmospheric noise, which includes the atmospheric absorption [15].

The signal then enters the cryostat at 4K and goes through a stack of low pass filters (LPF in the figure). The signals do not only contain submillimeter photons, but also higher frequency components. In order to retain only the THz photons, the signal must go through these low pass filters. The signal then guided towards the chip with two parabolic mirrors and passes through a wire grid polarizer, which only lets waves with one particular polarization through. The now linearly polarized signal goes through a low pass filter, and then a bandpass filter in order to filter out the irrelevant frequencies. The signal then enters the chip, which is cooled down to 120 mK.

3.2 The data

In order to perform a principal component analysis, the data must be represented in a data table. In this table, the columns are the variables and the rows are the different measurements of these variables. The variables, in this case, are the frequency responses of the different channels of the filterbank spectrometer, converted into T_{sky}. The rows are the different observations of T_{sky} for the channels.

In order to interpret the principal component analysis correctly, a principal component analysis on some artificial data is also performed. This data will contain highly correlated data, making it easier to interpret what the first couple of principal components really represent. The following sections will explain how the artificial data and the real data are obtained.

3.2.1 Real data

The real data is data from DESHIMA 1.0. This is data from when the telescope is pointed at a single point in the sky, meaning that the elevation and azimuth angles are constant. It contains measured values for T_{sky} for each channel at several timestamps during the scanning time. These still-sky observations are assumed to contain noise data governed by three noise types: the atmospheric noise, photon noise, and detector $1/f$ noise.

3.2.2 Artificial data

The artificial data is created using the ATM model. Using the ALMA Atmosphere Model, the zenith transparency of the sky, η_{zenith}, in a frequency range (100-1000 GHz) will be calculated for different values of PWV between 0.1-2.6 mm. The model makes a data table that gives the zenith atmospheric transparency for numerous frequency values in the chosen range and the different PWV values. By interpolating the data of this table, equation (2.2) can be used to calculate ηatm for any frequency in the chosen frequency range and any PWV value in the PWV range.

The artificial data table has a slightly different composition than that of the real data table. The columns are again the frequency responses of the different channels, now calculated with the ATM model using the resonance frequencies of the resonators as the channel frequencies. The different observations, however, are not calculated for different times but for different values of the PWV. The PWV values for each of the observations are initially obtained from DESHIMA 1.0 data. This data contains one PWV value for every data set, which is measured by ALMA at the start of each measurement. However, it is unlikely that the PWV value stays constant during the scanning times, which range from 300 to 3000 s. Since the artificial data should match with the real data in order to make a good comparison between the real and artificial data PCAs, the original ALMA PWV values are corrected by making the maximum and minimum T_{sky} responses of one channel in the data set match, setting a new PWV range for every measurement.

3.2.3 Comparison

Performing a principal component analysis on the artificial data will most likely lead to one principal component accounting for almost 100% of the variance because T_{sky} for the different channels is calculated with a function where only one variable is changed, the PWV value. It is thus expected that the first principal component represents the variation of the PWV value.

Performing a principal component analysis on the real data will reveal the influence of the most dominant noise source on the data. Comparing it to the artificial data will reveal if the most dominant noise source is the fluctuating atmosphere.

3.3 Number of principal components

An essential part of principal component analysis is choosing the number of components. When projecting the data on the new principal component axes, the number of principal components determines the accuracy of the new data set. There are multiple ways to determine a reasonable number for the principal components.

As mentioned in section 2.4 about PCA, the orthonormal eigenvectors point in the directions of the maximum variance and the corresponding eigenvalues are a measure for the amount of variance in the directions of these eigenvectors.

In this research, the concept of explained variance is used. As one eigenvalue of the covariance matrix represents the magnitude of variance along one principal component axis, the total sum of all the eigenvalues of the covariance represents the total variance in the data set. A straightforward way to visualize how important a principal component is, is to divide the eigenvalues by the total sum of the eigenvalues. This gives a percentage of explained variance for each principal component. These percentages will be sorted and displayed in a graph, along with their cumulative sum. This cumulative sum can often be very helpful because it tells how much variance is accounted for when choosing a certain number of components. Plotting these percentages and their cumulative sum for the example data set from section 2.4.2, results in figure 3.2.

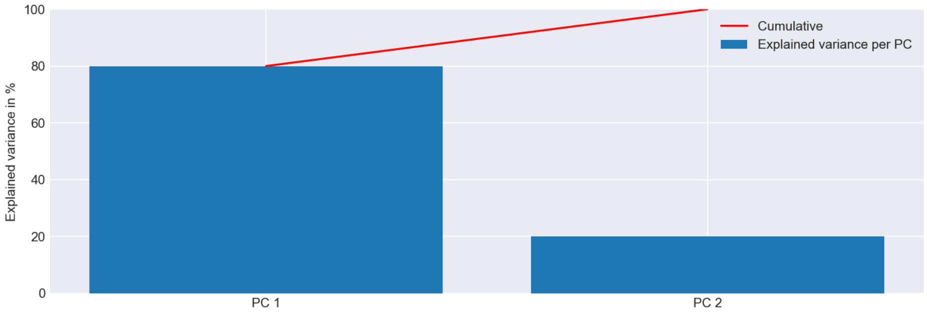

Figure 3.2: Explained variance plot of the example data set from section 2.4.2. Together, the two principal components explain 100% of the variance.

There is no rule as to how much explained variance the number of principal components should account for. This depends on the goal of the PCA and is thus rather subjective. In this research, the number of principal components is chosen to be the same for every data set. This way, the principal components can be easily compared. The number of components will be chosen such that the last component also varies significantly for different data sets.

There are some other methods that help when choosing the number of components, but all of them involve a large amount of subjectivity. Since the data will not be projected in this paper, the principal components are only calculated to visualize what they represent and to compare the components of different data sets. For this reason, the explained variance plot is a suitable method and also less subjective than other existing methods.

Chapter 4

Results and Discussion

This chapter presents the results and interpretation of the principal component analysis. For the real data, only the frequency responses of 46 channels are used, as three of the KIDs were not working correctly. In line with the real data, the artificial data is also only calculated for these 46 channels. The peak frequencies of the 46 channels are shown in figure 4.1.

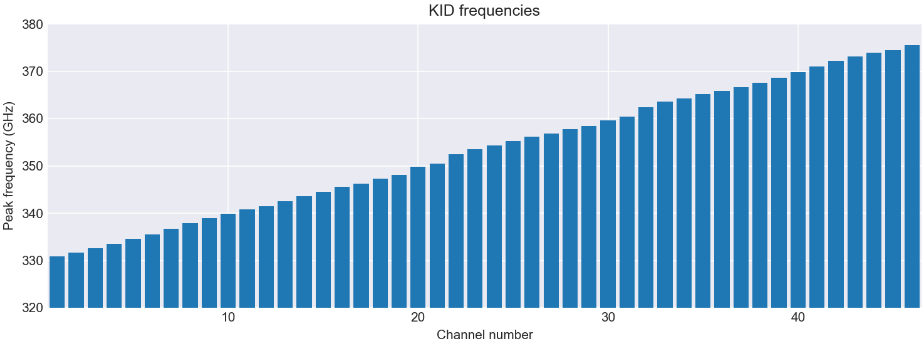

Figure 4.1: Peak frequencies of the 46 filterbank channels that are used. The frequencies range from 330.8 to 375.5 GHz

4.1 Artificial data

As mentioned in the previous chapter, a principal component analysis on artificial data is performed first. For this data, the peak frequencies of the 46 channels and a range for the PWV values are used. This first principal component analysis uses a trial range for the PWV ($0.98 \leq$ PWV$\leq$ 1.02) that is not related to any real data. It is merely used to portray the explained variance of the first principal component and to show what the first principal components represent.

4.1.1 Eigenvalue plot

Figure 4.2 shows the explained variance of the different principal components. As expected, the first principal component has an explained variance of nearly 100%. This means that the sky temperature measured by a channel, $T_{sky,i}$, is dependent on only one variable instead of 46. In order to determine what this principal component represents, the following step is to look at the eigenvectors belonging to the highest eigenvalues.

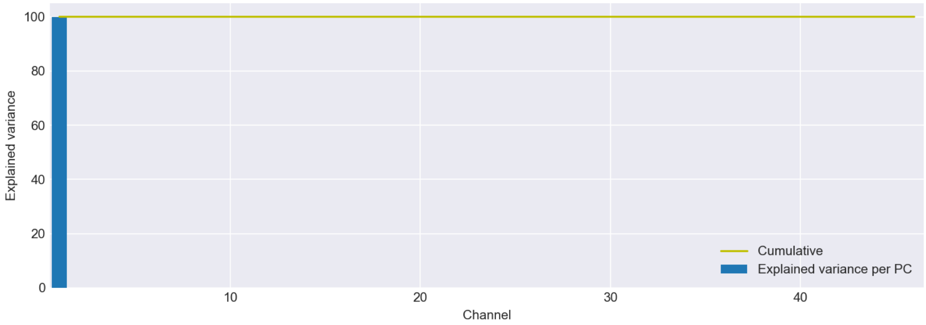

Figure 4.2: Explained variance plot for the artificial data. The first principal component explains ∼ 100% of the variance in the data set.

4.1.2 Interpreting the principal components

Another principal component analysis is performed on the artificial data, for different values of the mean PWV: 0.5, 1.0, 1.5, and 2.0 mm. The range is the same for every data set: 0.04 mm. The number of principal components is chosen to be four, because the fourth PC still shows some clear difference for the different data sets. Figure 4.3 shows the four eigenvectors belonging to the first four principal components. Even though the first principal component explains almost 100% of the variance, the other three principal components still seem to show that there are some other factors that explain a tiny portion of the total variance, as they are not zero and seem to represent some type of relation.

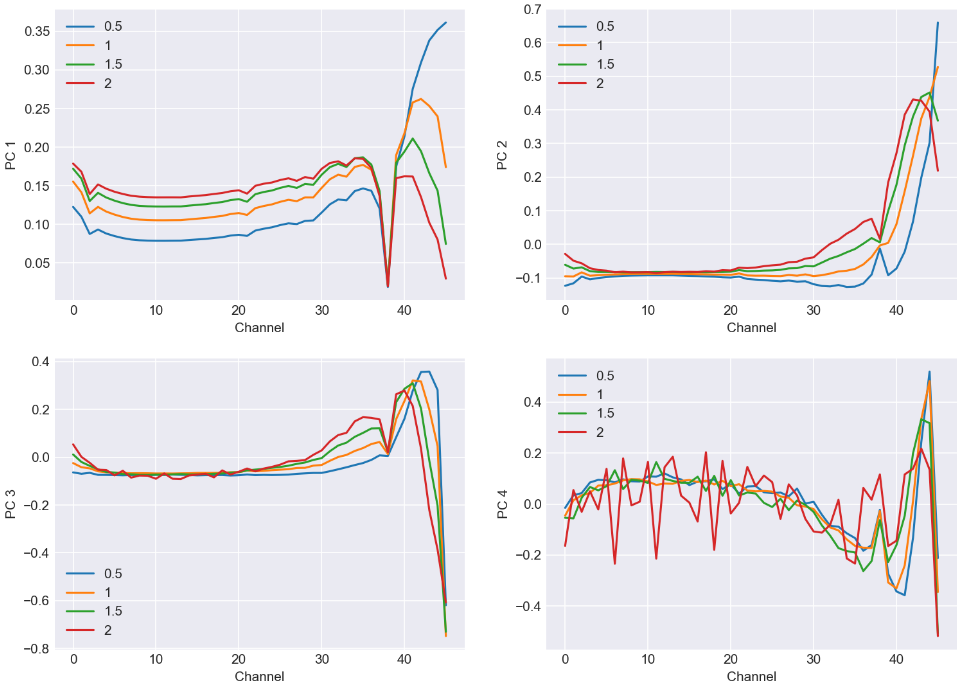

Figure 4.3: Plots of the eigenvectors belonging to the first four principal components of an artificial data set. The plots are made for four different mean values of the PWV. The different PWV plots show that the principal components are dependent on the PWV value.

This figure shows that the eigenvectors of the first four principal components have slightly different shapes for different absolute values of the PWV, but the different principal components are distinguishable. Interesting to see is how PC 4 becomes spikier and more random-looking as the absolute PWV value increases.

As the PWV value is the only changing variable that is used for the artificial data, the principal components will all represent some relation between $T_{sky,i}$ and PWV. These relations are further examined in the following sections.

The first principal component

The first principal component is expected to represent the fluctuation of the PWV value, as that is the only variable that changes for the artificial data. To show that this is true, figure 4.4 shows the brightness temperature response of one channel for all the PWV values in the range 1.25 mm$\leq$ PWV$\leq$ 1.75 mm. The brightness temperature is calculated for 100 PWV values in this range.

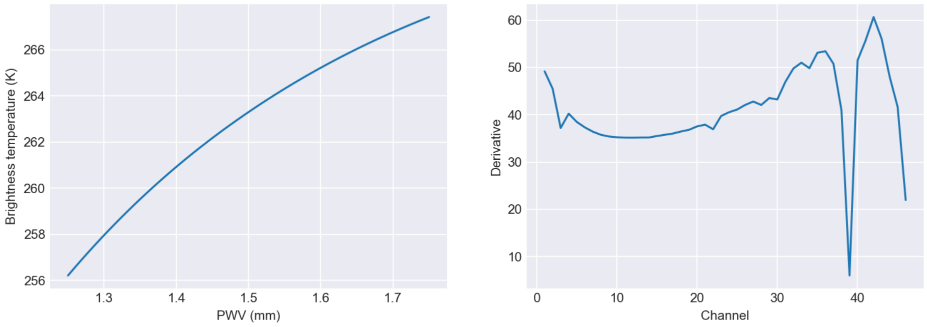

Figure 4.4: The first plot shows the brightness temperature of a channel (channel 34) as a function of the PWV value. The second plot shows the slopes of this line for every channel.

This plot shows that the frequency response and the PWV value share an almost linear relationship. The other channels show the same type of relationship, some more linear and some more curved. When plotting the slope of these lines for every channel, the second plot is the result. This line looks very similar to the yellow line PC1 plot, which is valid for the same PWV range. The magnitudes are not the same, but this is because the eigenvectors are normalized, while the artificial data is not. The slope plot tells us how the brightness temperature for each channel changes as a function of the PWV value. The similarity in shape of the slope plot and the PC1 plot suggests that the first principal component indeed represents the fluctuation of the PWV value. Also, when looking back at figure 2.3, it can be seen that the transmission of the atmosphere is ~0 for all PWV values around 368 GHz, which corresponds to channels 38 and 39. The PC1 plot also drops to 0 for these channels, which serves as another confirmation that PC1 is the derivative of the relation between T_{sky} and PWV.

The second principal component

As is visible in figure 4.4, the $T_{sky,i}$-PWV relations are slightly curved. The first principal component represents the slopes of these relations, which is effectively only the linear part. The other three principal components in figure 4.3 show that the relation between the frequency response and the PWV value is not entirely linear. This suggests that the non-linear part of the relation is represented by another principal components. In order to see if this is true, the derivative of one of the $T_{sky,i}$-PWV relations is plotted.

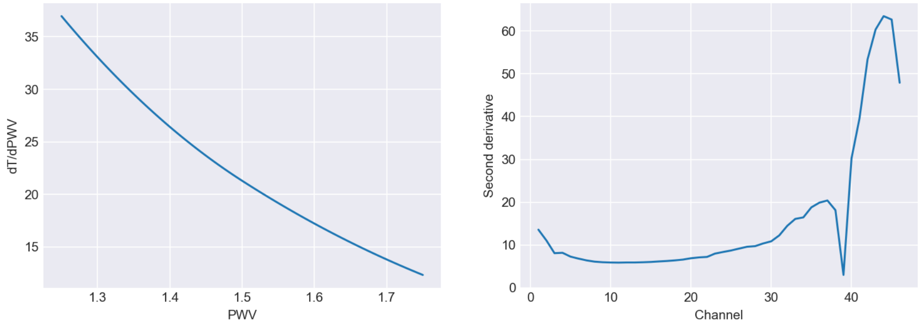

Figure 4.5: The first plot shows the derivative of the channel response with respect to PWV (channel 34), as a function of the PWV value. The second plot shows the slopes of this line for every channel.

The derivative plot shows that the derivative of the relation is not a constant value, but changes with respect to PWV due to the curvature. The slope of the derivative will represent some of this curvature. These slopes of the derivatives of the $T_{sky,i}$-PWV relations for all the channels are also plotted in figure 4.5. This line looks very similar to the PC2 plot in figure 4.3, suggesting that the second principal component of the artificial data set represents the second-order derivative of the $T_{sky,i}$-PWV relation.

As can be seen, though, the derivative plot is also curved. This suggests that the same process could be continued, perhaps also revealing the third and fourth principal components. The explained variance for these other principal components is also so low, however, that only the first two principal components are examined.

Explanation

The plots show clear similarities between the first two principal components and the first and second-order derivatives of the channel responses. This can be explained by taking a closer look at the relation between T_{sky} and PWV. First, let us take a look at equation (2.2):

$$T_{sky} = T_{atm}(1 - e^{-\tau}) \tag{2.2 revisited}$$

The opacity τ can be rewritten in terms of PWV. In order to see how these two variables are related, the next plot shows scatter plots of τ versus PWV for a few different channels. Here, τ is calculated with equation (1.1), and η_{atm} with the ALMA ATM model [10].

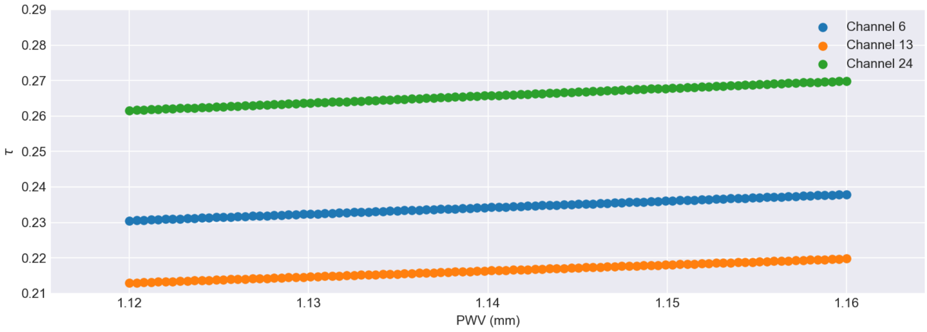

Figure 4.6: Caption

This figure suggests that τ and PWV are linearly related. We can then write a general expression for this relation:

$$\tau_i \approx a_i \cdot PWV + b_i \tag{4.1}$$

In this equation, a_i and b_i are channel-specific. Combining this equation and equation (2.2) gives a general expression for T_{sky} as a function of PWV. The principal components of the artificial data set are variables that cause maximum variance in the data set. The next step is then to write an expression for ∂T_{sky}:

$$\partial T_{sky,i} = T_{atm}(1 - e^{-\partial \tau_i}) \approx T_{atm}(\partial \tau - \frac{1}{2}\partial \tau^2) \tag{4.2}$$

The expression on the right hand side of equation (4.2) is the Taylor series expansion around $\partial \tau = 0$ up to the second order. By using that $\partial \tau_i = a_i \cdot \partial PWV$, equation (4.2) can be written as:

$$\partial T_{sky,i} = w_1 \cdot \partial PWV + w_2 \cdot \partial PWV^2 \tag{4.3}$$

This equation shows where the first two principal components originate from. The first principal component is a scaled first-order derivative of the channel responses, and the second principal component is the scaled second-order derivative. The other principal components could originate from the higher-order terms of the Taylor series expansion.

A principal component analysis converts the original variables into new variables that are linear combinations of the original variables. This explains why even the artificial data set has more than just one principal component, as the relations between T_{sky} and the PWV values are not linear. Performing a Taylor series expansion decomposes the non-linear $T_{sky,i}$ into linear orthogonal terms, precisely the conditions that apply for the principal components.

4.2 Real data

Table 4.1 shows a summary of fourteen still-sky measurements by DESHIMA 1.0. The first column contains the RunIDs of the measurements, which serve as labels for the specific data sets. The second column contains the corresponding PWV values, measured by ALMA. These values are corrected in accordance with the real data and the ATM model. The new PWV values and ranges are shown in the fifth column. The fourth column displays the scanning times of the measurements. These can serve as an indicator for the magnitude of the PWV range, as it is more likely that the PWV value changes significantly when the scanning time is longer. The data files also contain timestream data of the cabin temperature, the pressure, wind speed and direction. These measurements are not used in this research, but could improve accuracy when they are included.

Table 4.1: Summary of the DESHIMA 1.0 data used for the PCA.

RunID	ALMA PWV (mm)	Elevation (°)	Time (s)	Estimated PWV range (mm)
20171104235158	N.A.	60	314	1.12 - 1.16
20171111023644	0.68	60	318	0.80 - 0.85
20171111062033	0.67	60	315	0.74 - 0.82
20171112042102	0.65	60	314	0.63 - 0.66
20171112053522	0.58	60	315	0.60 - 0.66
20171112100903	0.81	60	315	0.70 - 0.73
20171112203617	1.13	60	317	1.09 - 1.20
20171115073724	1.98	85	314	2.10 - 2.18
20171115080822	1.78	88	316	1.91 - 2.00
20171115081732	1.72	88	3016	1.49 - 1.95
20171115091754	1.72	88	315	1.45 - 1.50
20171115133249	1.79	85	317	1.76 - 1.88
20171115135427	1.72	85	788	1.56 - 1.93
20171115141509	1.70	60	3018	1.67 - 1.83

4.2.1 New PWV ranges

The ALMA PWV values were measured at one time instant at the start of the DESHIMA observation, making them too inaccurate to use for the artificial data PCA. New PWV values are estimated along with their range, based on the real data. This is done by looking at the T_{sky} responses of one channel for each measurement and adjusting the mean and the range of the PWV such that the maximum an minimum values for T_{sky} match.

This method, though very simple, is not as accurate. This can be seen by looking at the frequency responses of the channels of the real data at one time instant and the responses of the artificial data for the corresponding estimated PWV value. This is plotted in figure 4.7.

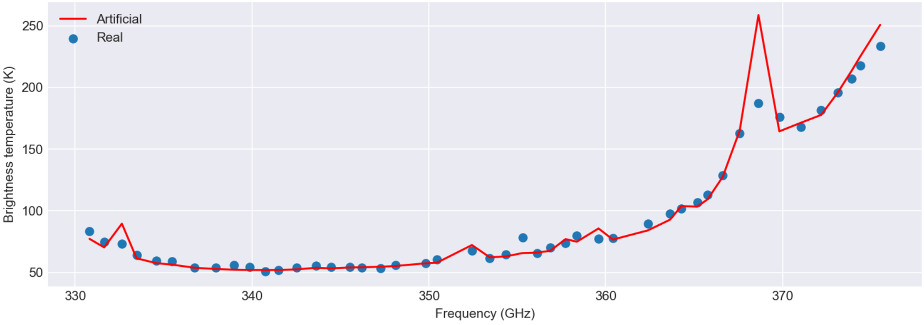

Figure 4.7: Comparison of observations of the artificial data and the real data.

From this figure, it can be seen that the real frequency responses of the different channels are not exactly related to each other like the artificial frequency responses. This means that the same PWV value has a different effect on the different channels. By just looking at one channel, this is not accounted for.

The discrepancy between the estimated PWV ranges and the actual PWV ranges can further be visualized by plotting the ranges of the T_{sky} responses of the different channels for both the artificial data and the real data. This is shown in figure 4.8:

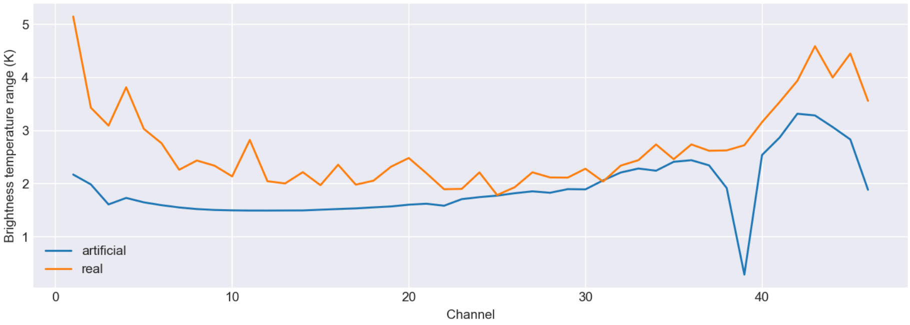

Figure 4.8: Magnitudes of the brightness temperature ranges for every channel. The blue line is the magnitude of the T_{sky} range for the estimated PWV range, and the orange line is the T_{sky} range of the real data during the scanning time.

This plot uses data from data set 20171104235158 for the orange line and the corresponding estimated PWV range for the blue line. This PWV range was estimated by matching the artificial and real T_{sky} ranges of channel 36. The plot indeed shows that the ranges are nearly equal for this channel and also some other channels. For the most part, however, the ranges of the real data are larger, indicating that the estimated PWV range is not in complete agreement with the actual PWV range.

This method of estimating the PWV range has much room for improvement, but figures 4.7 and 4.8 also show that making a perfect estimation is very difficult. For the purpose of this research, a perfect estimation is not necessary, but it is important to keep in mind that the slightly inaccurate estimated PWV ranges could yield to some discrepancies in the comparison of the principal component analyses of the real and artificial data.

4.3 Principal Component Analysis

The results of the PCA on the real and artificial data sets are presented in appendix A. The upper right figure shows the timestream response of channel 36 in blue. This figure also shows a red vertical line, which is the range of the T_{sky} response of the artificial data. The upper-middle plot shows the explained variance plot of the real data in blue and the cumulative explained variance in green. The explained variance of the artificial data is represented by the red dots in this figure, and it can be seen that the first principal component always has an explained variance of $\sim 100\%$. The remaining four figures show the first four principal components of both the artificial data (red) and the real data (blue).

What catches the eye immediately when looking at all of the PCA plots, is how well the first principal components of the real and the artificial data sets match. This means that the most dominant noise source for the still-sky measurements is the atmosphere, or more specifically, the PWV fluctuation. This is an interesting result, as it seems to be independent of the magnitude of the PWV range. The real data sets also contain detector $1/f$ noise and photon noise, but these seem to be less dominant, even when the PWV fluctuation is small. The other PC plots are less unambiguous and need a closer look.

Important to note when looking at the results is that the eigenvectors are always normalized. When the influence of the PWV fluctuation is small because of a small fluctuation, the other noise sources seem to become larger. What this actually implies, however, is that the relative noise contribution caused by the PWV fluctuation becomes smaller, bringing it closer to the photon noise. The photon noise is always random noise, and the absolute influence that it has on the noise signal is mostly quite constant. It is thus important to keep in mind that the explained variance plot shows the relative significance of the principal components, not the absolute significance.

The results can be sorted into one of three categories: data sets with large PWV ranges, data sets with small PWV ranges, and data sets that show a systematic error. The following paragraphs will highlight these three cases by looking at the PCA plots of three different data sets.

Case 1: Large PWV range

Figure 4.9 shows the resulting plots for data set 20171115135427, which has a large estimated PWV range (1.56-1.93 mm). This can be seen in the upper right plot. The explained variance of the first principal component is very high, nearly equal to that of the artificial data set. This makes sense, as the relative noise contribution of the atmosphere increases when the atmosphere's fluctuation is larger. The second principal component also looks very similar to that of the artificial data. This indicates that the non-linear part of the T_{sky}-PWV relation is also a more dominant noise source than the detector $1/f$ noise and the photon noise. PC 3 and PC 4, although less convincing, also show some similarities to PC 3 and PC 4 of the artificial data set. They look more random, however, meaning that they could also contain information about the detector and photon noise.

Case 2: Small PWV range

Figure 4.10 shows the results of the principal component analysis on the first of the data sets, RUNid 20171104235158. The estimated PWV range of this data set is 1.12-1.16 mm, which is a very small fluctuation. The relative noise contribution of the atmosphere is then expected to be lower than it would be for a larger fluctuation, which is also seen in the explained variance plot. The first principal component again represents the first order T_{sky}-PWV relation, but now explains only $\sim$ 58% of the total variance. The relative noise contribution of the higher-order parts of the T_{sky}-PWV relation are not clearly recognizable in the other PC plots, indicating that they become less significant when the PWV fluctuation is small. This makes sense when taking a look at figure 4.4. If the PWV range is very small, the T_{sky}-PWV can almost entirely be described by a linear function, making the higher-order terms in the Taylor series expansion minimal. This is also in agreement with the previous, large PWV range data set.

Case 3: Systematic error

Figure 4.11 shows the plots for RUNid 20171112042102. This data has a small PWV range (0.63-0.66 mm), and we would thus expect the plots to look similar to those of data set 20171104235158. This is not the case, though, and this result shows several interesting features. The explained variance plot shows that the first principal component explains around 72% of the total variance. This first principal component again describes the first-order term of the T_{sky}-PWV relation, but it does not entirely match with the artificial PC 1. This is presumably caused by an inaccurate estimation of the PWV range. Looking at figure 4.3, the PWV estimation is likely a bit higher than the actual PWV, which is caused by the method of PWV estimation mentioned in section 4.2.1. The explained variance of the second principal component is quite high - around 10% - and now it seems to match with PC 2 of the artificial data. This is strange, as in the previous paragraph, it was suggested that the higher-order terms of the T_{sky}-PWV relation become very small when the PWV fluctuation is small. We would thus expect the second principal component to represent random noise rather than the second-order term. Looking at the explained variance of PC 2 of other data sets with small PWV fluctuation, the explained variance of 10% is also unusual. This all suggests that there might be some systematic error in the PCA algorithm that was used.

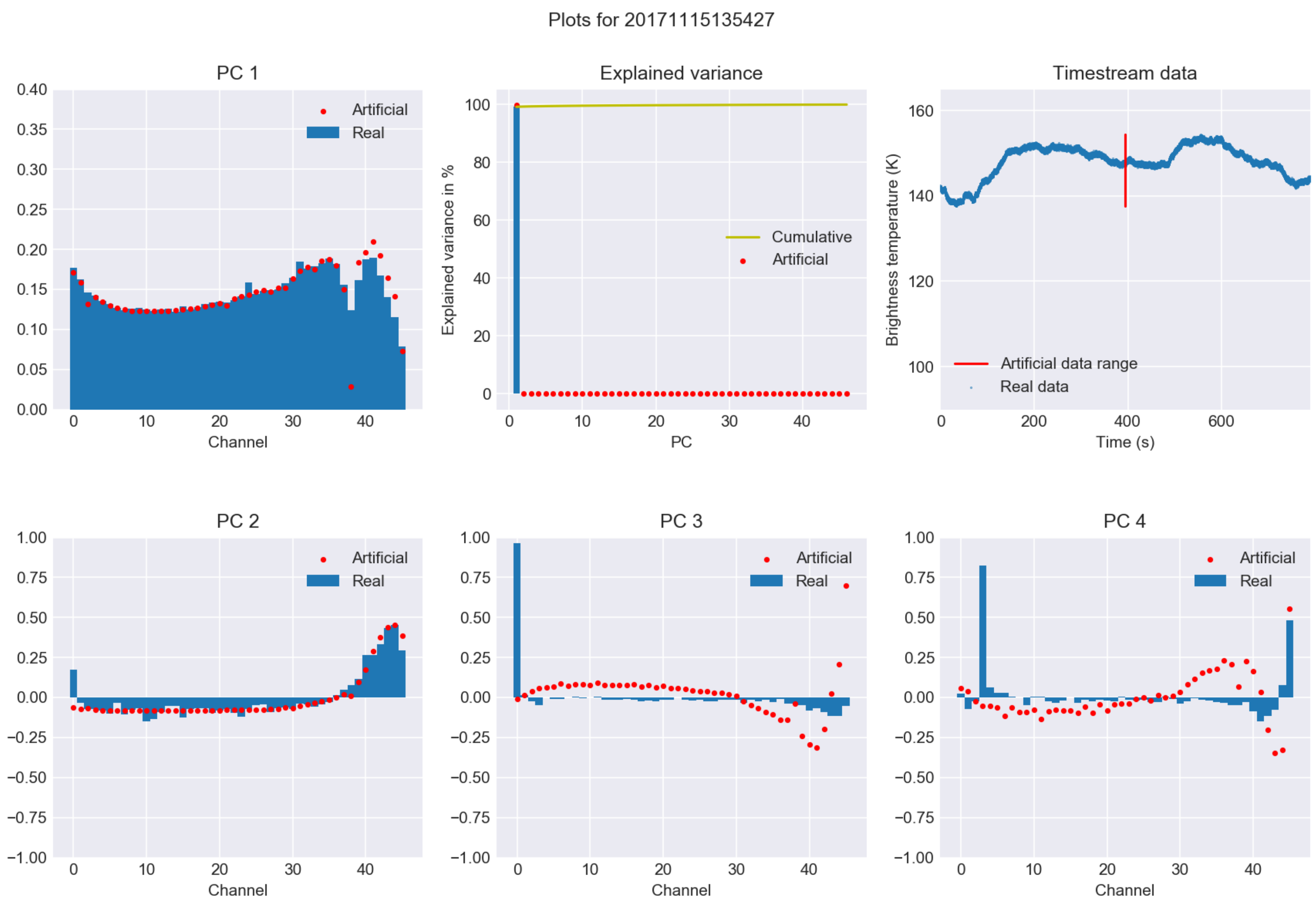

Figure 4.9: Plots of the principal component analysis of data set 20171115135427. This is an example of a measurement with large PWV fluctuation.

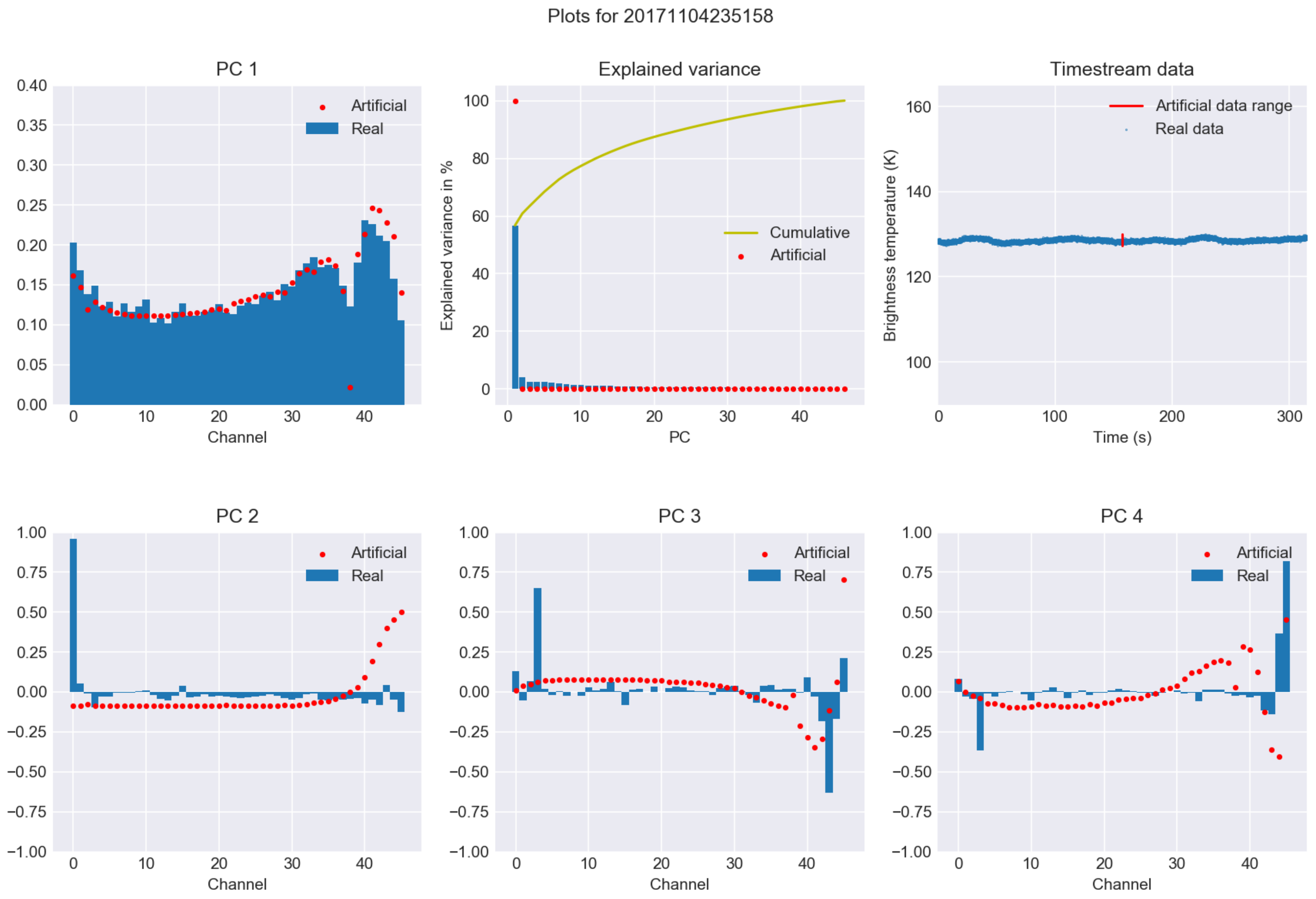

Figure 4.10: Plots of the principal component analysis of data set 20171104235158. This is an example of a measurement with small PWV fluctuation.

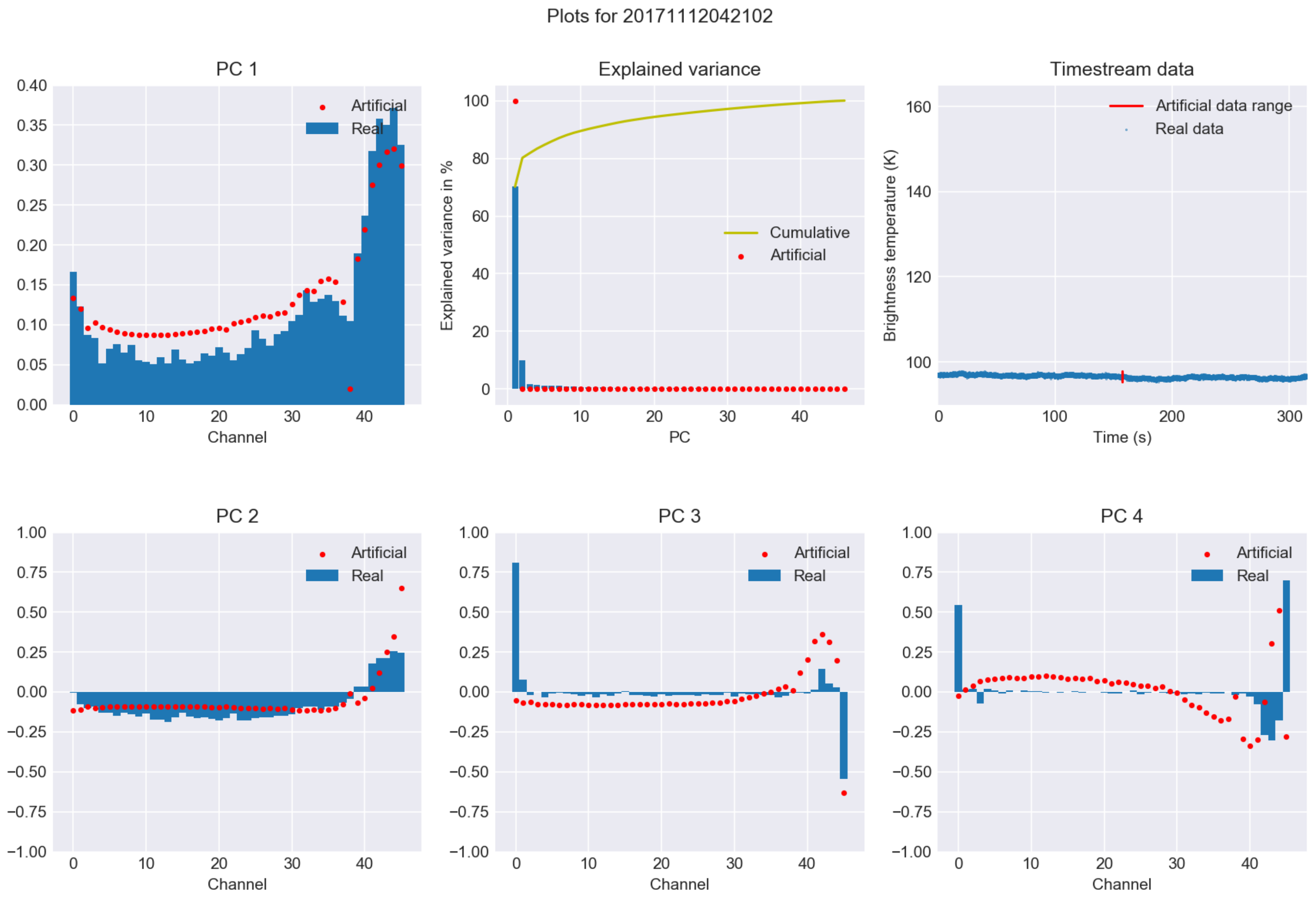

Figure 4.11: Plots of the principal component analysis of data set 20171112042102. This is an example of a PCA that exhibits a systematic error.

Chapter 5

Conclusion & recommendations

5.1 Conclusions

The goal of this research was to find the most dominant noise source during DESHIMA's still-sky measurements of T_{sky} by using a principal component analysis. It was expected that the atmosphere would have a large influence on the data. This was done by creating an artificial data set that only contained atmospheric noise data, with all constant variables except for a fluctuating PWV value. A principal component analysis on this artificial data revealed that the first two principal components represent the first and second-order terms of the Taylor series expansion of the T_{sky}-PWV relation, respectively. After performing a principal component analysis on DESHIMA data of still-sky measurements and comparing this to the artificial data PCA, it is found that the largest noise source for still-sky measurements also represents the first-order term of the T_{sky}-PWV relation, meaning that the most dominant noise source during still-sky observations is the fluctuating PWV value in the atmosphere. This is independent of the range and mean value of the PWV value. When the PWV fluctuation is significant, the second-order Taylor term of the T_{sky}-PWV relation becomes more dominant than the photon noise and the detector $1/f$ noise. When the fluctuation is small, the photon and detector generally become more dominant than the higher-order Taylor terms. The results also show some systematic errors, which require further research.

5.2 Further research

5.2.1 Improvements

Even though this study has given good results, the method can be improved. One improvement could be to use the actual bandwidths of the channels of the spectrometer, instead of only the peak frequencies. The overlapping bandwidths of the spectrometer channels cause the power of a signal to be shared between neighbouring channels. Including the bandwidths would yield more accurate results for the PCA.

Also, it has been established that the method used to estimate the PWV ranges is not very accurate. Improving this method will especially be useful for the comparisons between the real and artificial data PCA's.

5.2.2 Random noise level

The explained variance plots presented in the results showed the relative noise contribution of different noise sources. The noise in the real data is assumed to originate from three distinct noise mechanisms: the atmospheric noise, photon noise, and detector $1/f$ noise. The noise contribution of the atmosphere in a data set is seen to change depending on the magnitude of the PWV fluctuation. The photon noise and the detector $1/f$ noise, however, do not change that much for different observations. So, even though the relative contributions of these noise types may change a lot per observation, their T_{sky} contributions do not. Using this knowledge, a random noise level can be designed. This random noise level could be instrumental in the explained variance plots, as it would clarify whether a principal component represents random noise or some other physical mechanism.

5.2.3 Systematic error

In the previous chapter, the resulting PCA plots of one data set have been presented which do not agree with the theory described in this report. It is suggested that this may be an error in the PCA algorithm that was used, but this is not certain. Before moving on to the next step in this research, which would be to include scanning observations, this error must be thoroughly studied and clarified, as the analyses will only become more complicated and finding the cause of this error will, too.

5.2.4 Calibration method

This research serves the first step in determining the influence of the atmosphere on observation data. DESHIMA is obviously not designed to do still-sky observations only. When searching for unknown submm galaxies, the telescope must scan larger areas of the sky, changing its azimuth and elevation angles. These observations will, as opposed to the data sets used in this study, include not only noise data, but also actual signals from these new galaxies.

The next step would be to include observations which include scanning data, meaning that the telescope would have changing azimuth and elevation angles. In order to account for the changing angles, new variables have to be included in the analysis. This could serve as the next step in designing a calibration method that perfectly filters out the atmosphere.

Bibliography

[1] Caitlin M Casey, Desika Narayanan, and Asantha Cooray. Dusty star-forming galaxies at high redshift. *Physics Reports*, 541(2):45–161, 2014.

[2] G Lagache, M Cousin, and M Chatzikos. The [cii] 158 micron line emission in high-redshift galaxies. *arXiv preprint arXiv:1711.00798*, 2017.

[3] A Endo, JJA Baselmans, PP van der Werf, B Knoors, SMH Javadzadeh, SJC Yates, DJ Thoen, L Ferrari, AM Baryshev, YJY Lankwarden, et al. Development of deshima: a redshift machine based on a superconducting on-chip filterbank. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VI*, volume 8452, page 84520X. International Society for Optics and Photonics, 2012.

[4] Jochem Baselmans. Kinetic inductance detectors. *Journal of Low Temperature Physics*, 167(3-4):292–304, 2012.

[5] Jiansong Gao. *The physics of superconducting microwave resonators.* PhD thesis, California Institute of Technology, 2008.

[6] David M Slocum, Elizabeth J Slingerland, Robert H Giles, and Thomas M Goyette. Atmospheric absorption of terahertz radiation and water vapor continuum effects. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 127:49–63, 2013.

[7] Akira Endo, Kenichi Karatsu, Yoichi Tamura, Tai Oshima, Akio Taniguchi, Tatsuya Takekoshi, Shin'ichiro Asayama, Tom JLC Bakx, Sjoerd Bosma, Juan Bueno, et al. First light demonstration of the integrated superconducting spectrometer. *arXiv preprint arXiv:1906.10216*, 2019.

[8] Akira Endo, Kenichi Karatsu, Alejandro Pascual Laguna, Behnam Mirzaei, Robert Huiting, David J Thoen, Vignesh Murugesan, Stephen JC Yates, Juan Bueno, Nuri van Marrewijk, et al. Wideband on-chip terahertz spectrometer based on a superconducting filterbank. *Journal of Astronomical Telescopes, Instruments, and Systems*, 5(3):035004, 2019.

[9] A. Endo. Thz superconducting astronomical instrumentation, lecture 2 : Radiative transfer. THz Sensing Group, Dept. Microelectronics, Faculty of EEMCS, TU Delft, April 2019.

[10] ALMA. Atmosphere model. `https://almascience.nrao.edu/about-alma/atmosphere-model`. [Online; accessed 10-July-2019].

[11] Juan R Pardo, José Cernicharo, and Eugene Serabyn. Atmospheric transmission at microwaves (atm): an improved model for millimeter/submillimeter applications. *IEEE Transactions on antennas and propagation*, 49(12):1683–1694, 2001.

[12] Ian Jolliffe. *Principal component analysis.* Springer, 2011.

[13] Jonathon Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.

[14] James E Gentle. Matrix algebra. *Springer texts in statistics, Springer, New York, NY, doi*, 2:156–157, 2017.

[15] Crystal Brogan. Advanced calibration techniques. https://science.nrao.edu/science/meetings/2014/14th-synthesis-imaging-workshop/lectures-files/2014_AdvancedCalibration.pdf.

Appendix A

Results

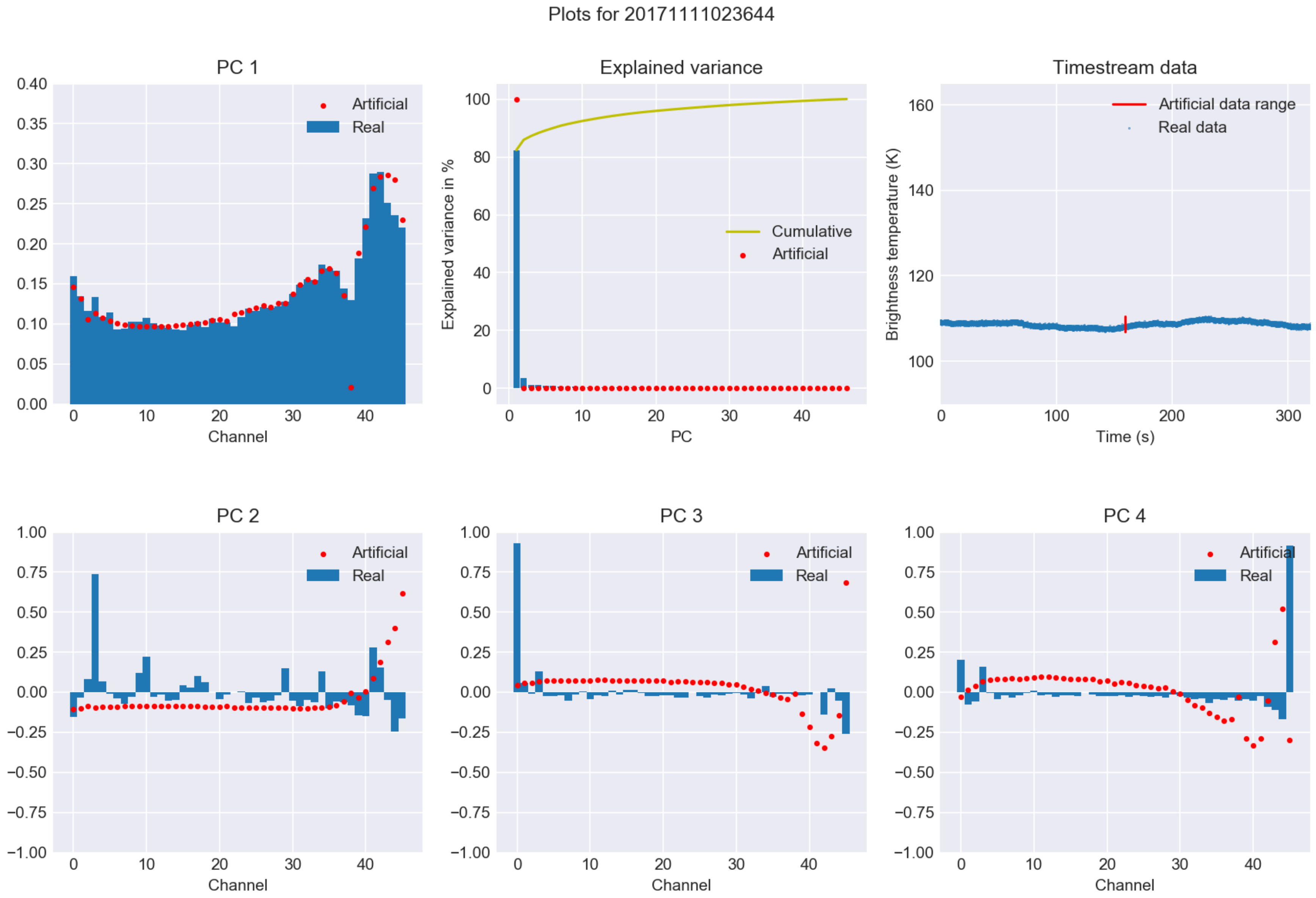

Figure A.1: Plots of the principal component analysis of data set of measurements when the telescope was pointed at the sky (blue). The same plots show a comparison with a PCA performed on the dummy data.

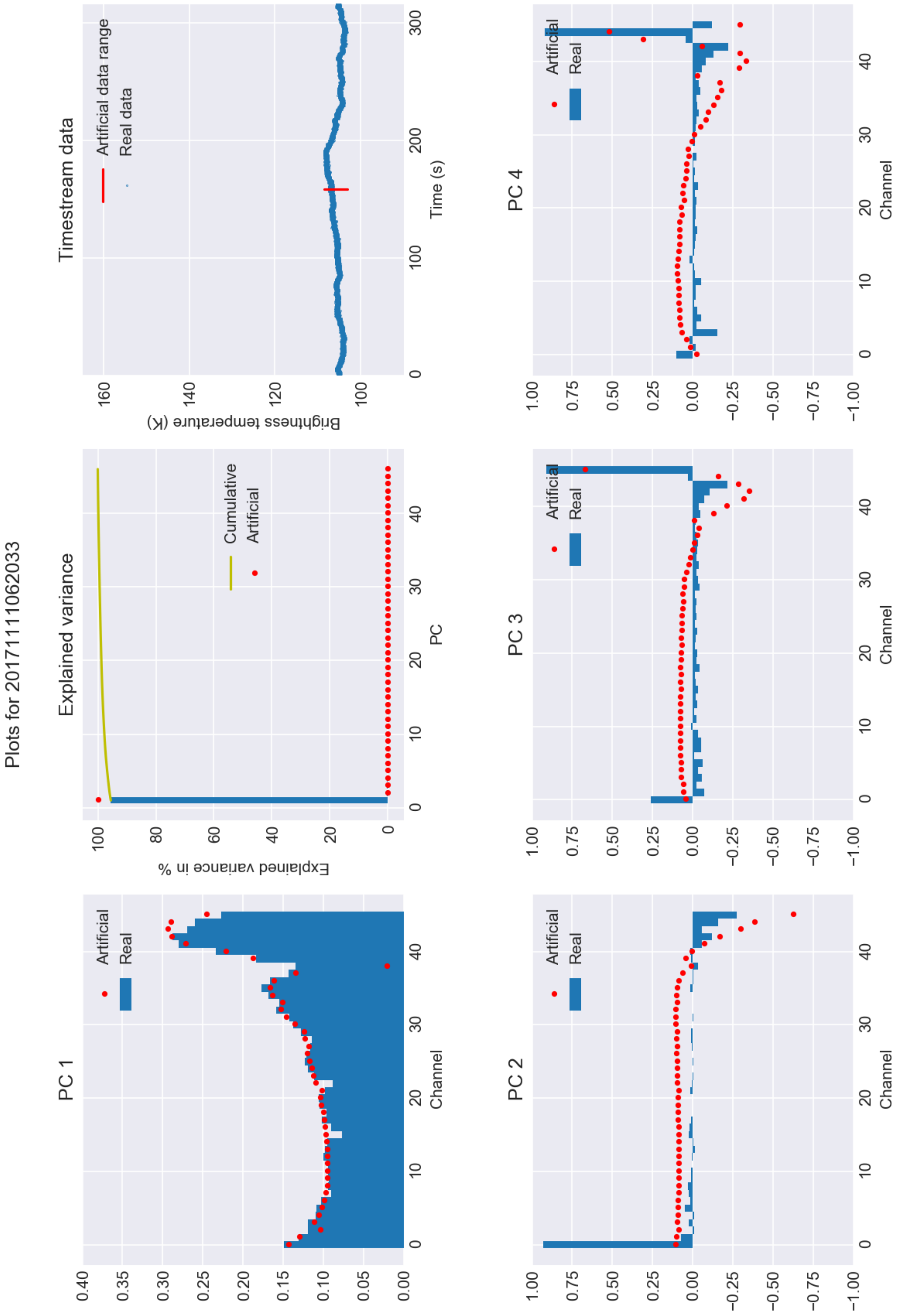
Plots for 20171111062033
PC 1
Artificial
Real
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Explained variance
Cumulative
Artificial
100
80
60
40
20
0
Explained variance in %
0
10
20
30
40
PC
Timestream data
Artificial data range
Real data
160
140
120
100
Brightness temperature (K)
0
100
200
300
Time (s)
PC 2
PC 3
PC 4
Artificial
Real
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
10
20
30
40
Channel

Figure A.2

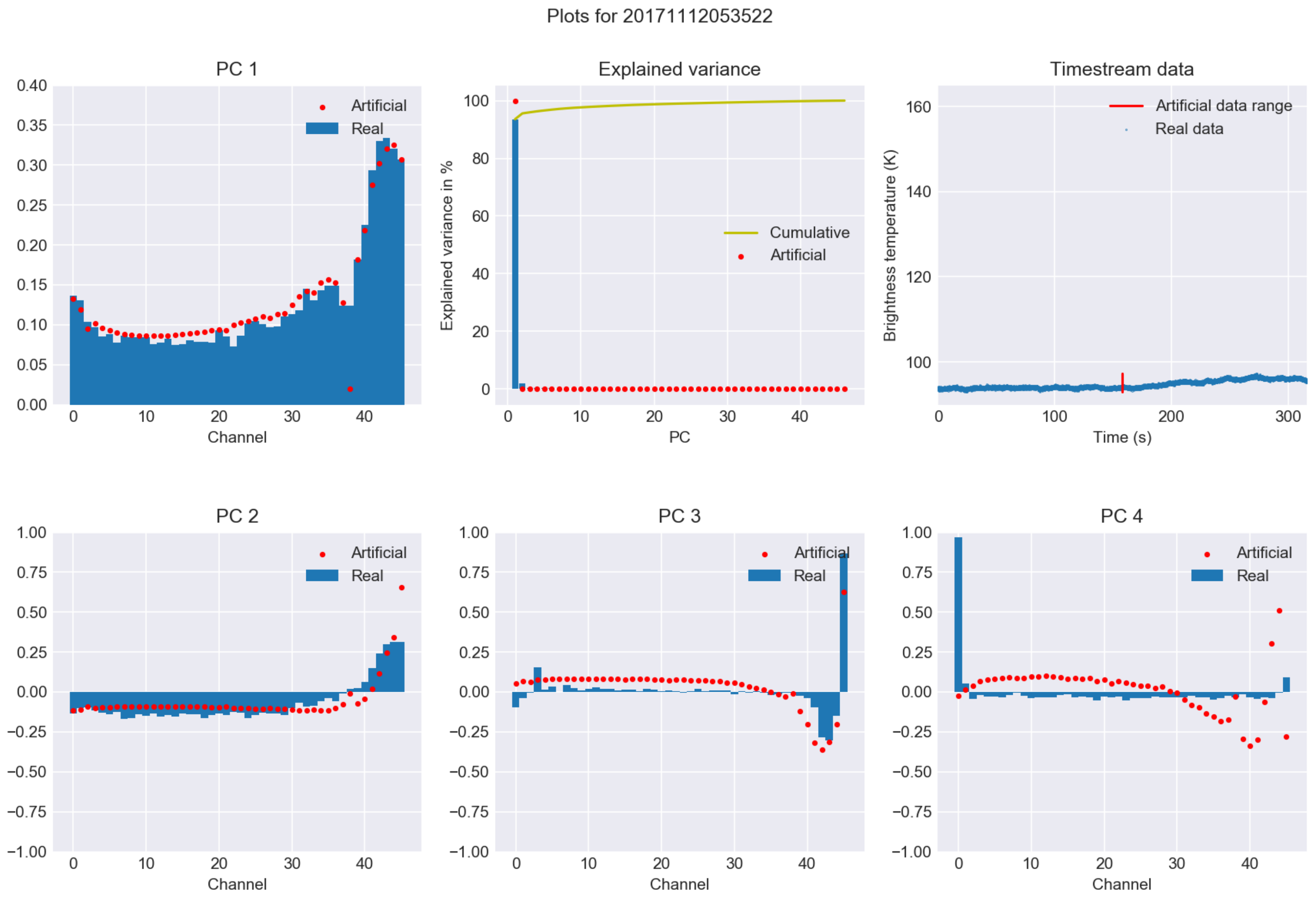
Plots for 20171112053522
PC 1
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Artificial
Real
Explained variance
Explained variance in %
100
80
60
40
20
0
PC
Cumulative
Artificial
Timestream data
Brightness temperature (K)
160
140
120
100
0
100
200
300
Time (s)
Artificial data range
Real data
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Channel
Artificial
Real

Figure A.3

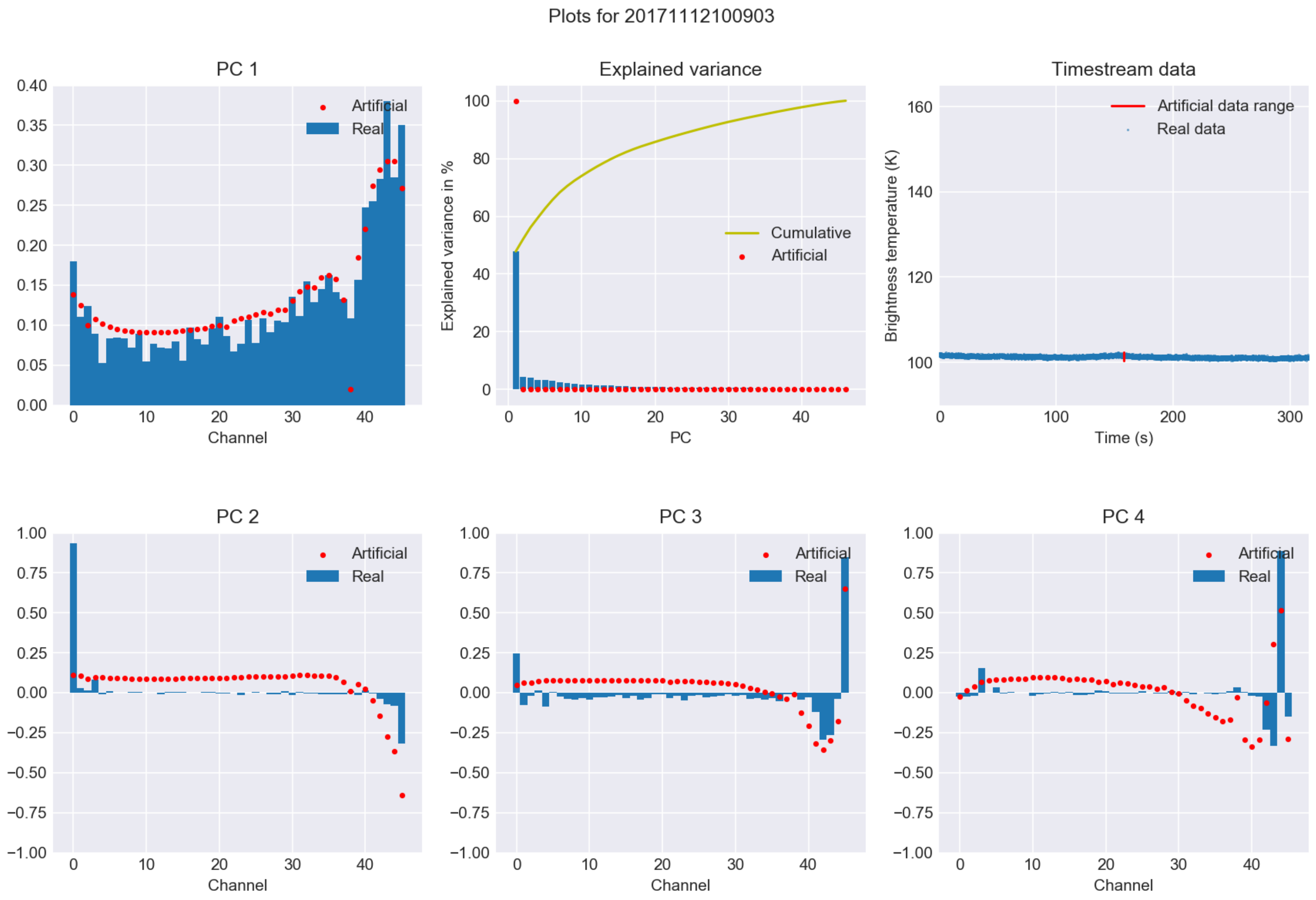
Plots for 20171112100903
PC 1
Artificial
Real
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Explained variance
Explained variance in %
100
80
60
40
20
0
Cumulative
Artificial
PC
Timestream data
Brightness temperature (K)
160
140
120
100
Artificial data range
Real data
0
100
200
300
Time (s)
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Channel

Figure A.4

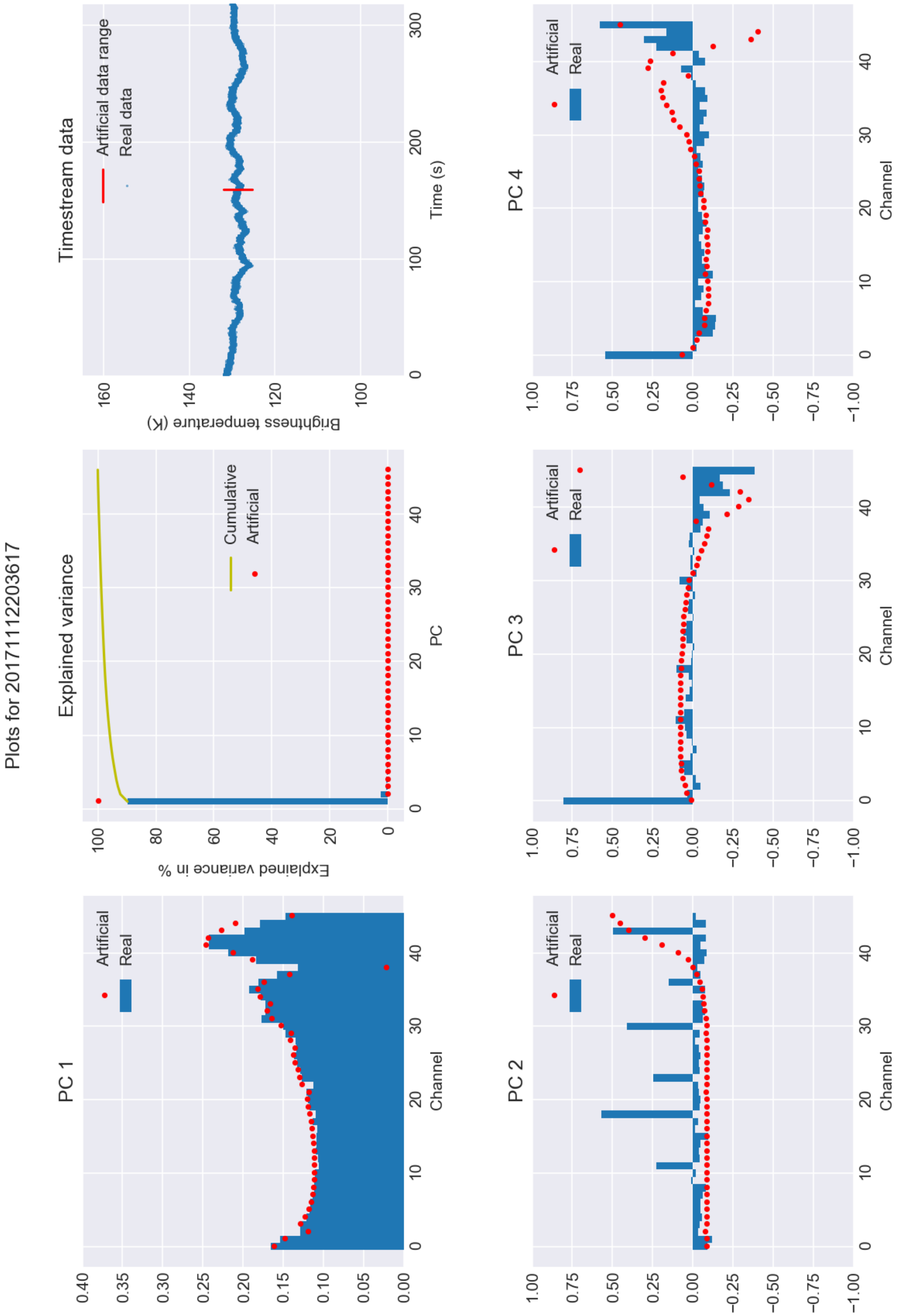
Plots for 201711112203617
PC 1
Artificial
Real
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Explained variance
Cumulative
Artificial
Explained variance in %
100
80
60
40
20
0
PC
Timestream data
Artificial data range
Real data
Brightness temperature (K)
160
140
120
100
Time (s)
300
200
100
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Figure A.5

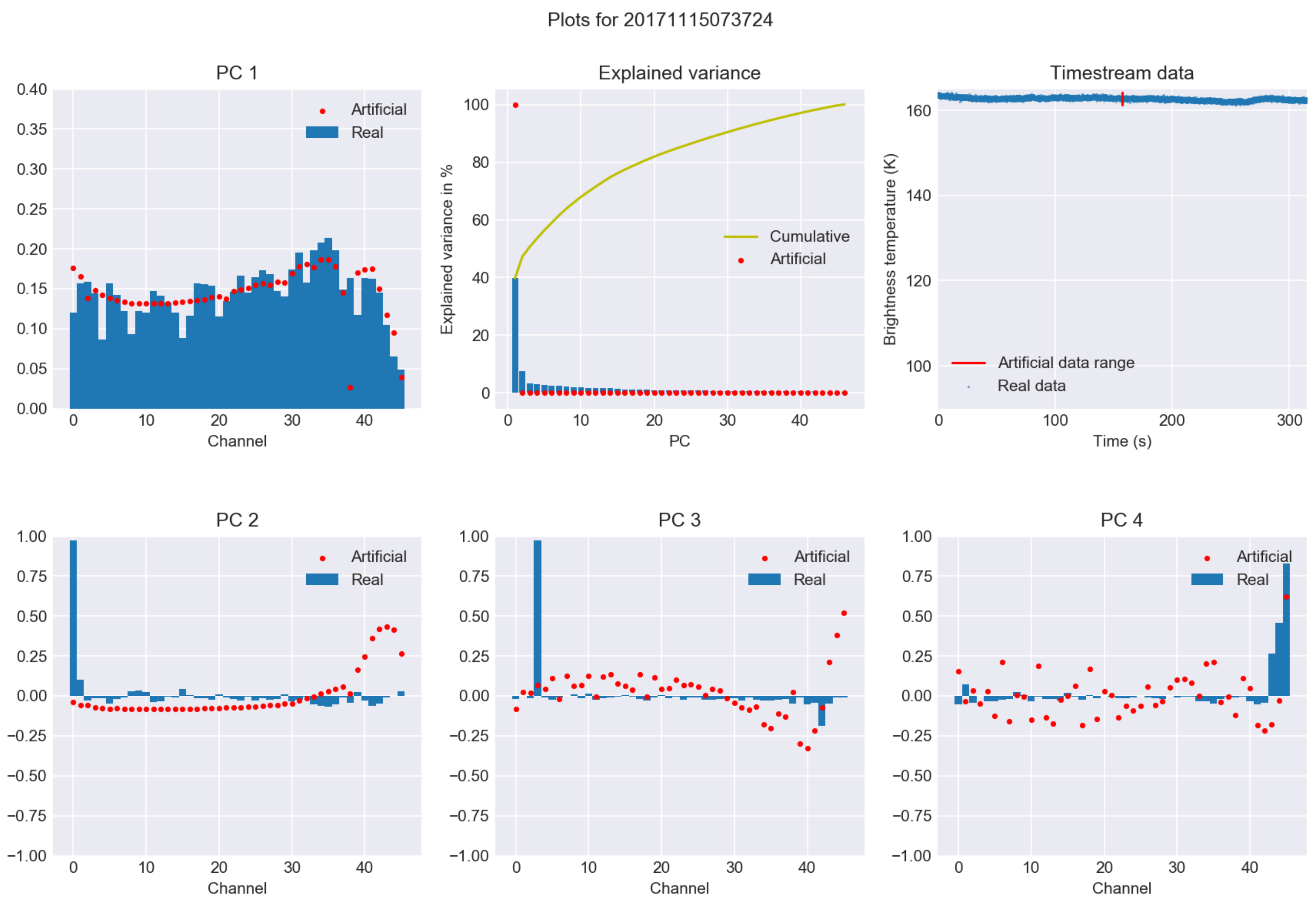
Plots for 20171115073724
PC 1
Artificial
Real
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Explained variance
Explained variance in %
100
80
60
40
20
0
Cumulative
Artificial
PC
Timestream data
Brightness temperature (K)
160
140
120
100
Artificial data range
Real data
0
100
200
300
Time (s)
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Channel

Figure A.6

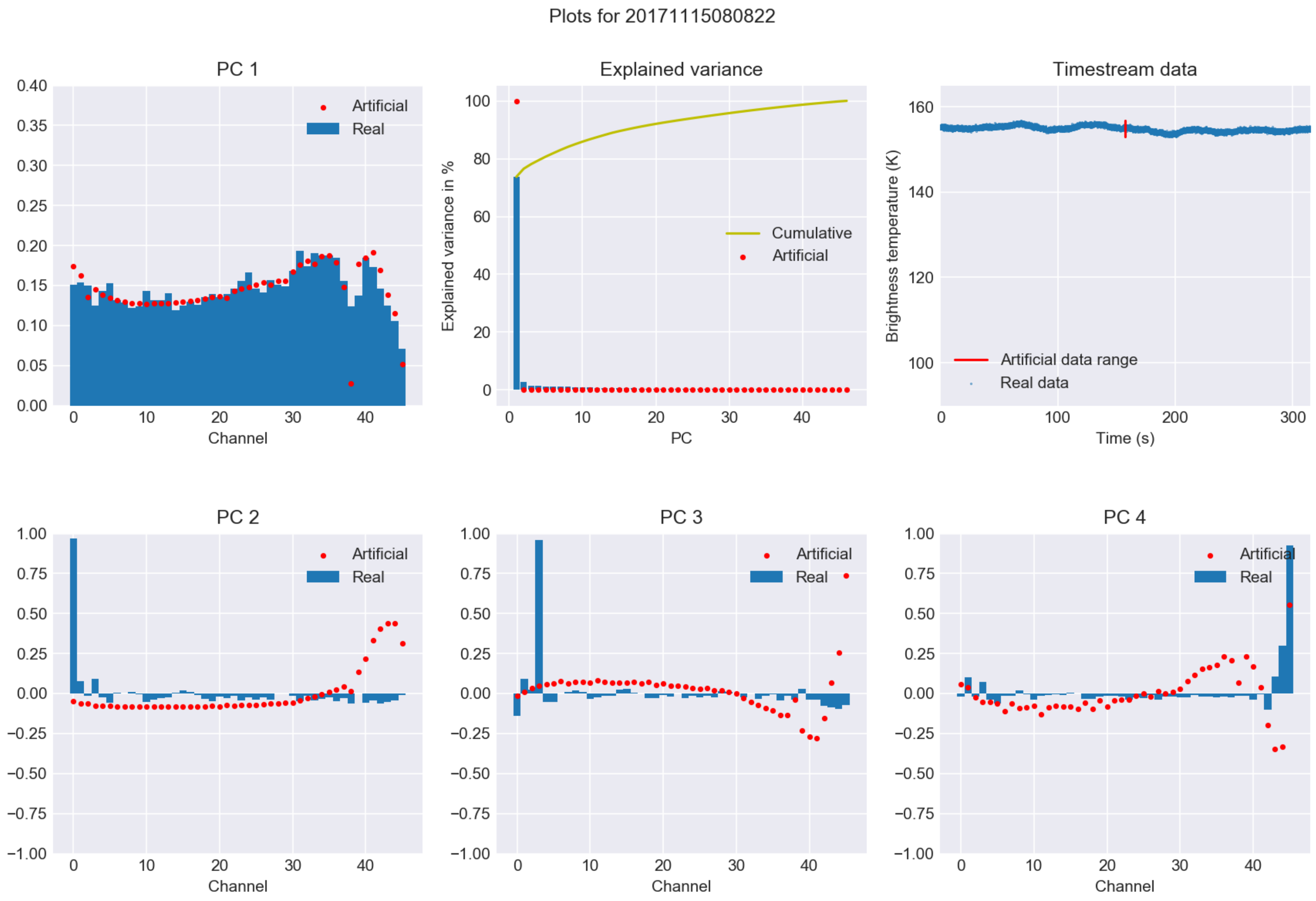
Plots for 20171115080822
PC 1
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Artificial
Real
Explained variance
Explained variance in %
100
80
60
40
20
0
PC
Cumulative
Artificial
Timestream data
Brightness temperature (K)
160
140
120
100
0
100
200
300
Time (s)
Artificial data range
Real data
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Channel
Artificial
Real

Figure A.7

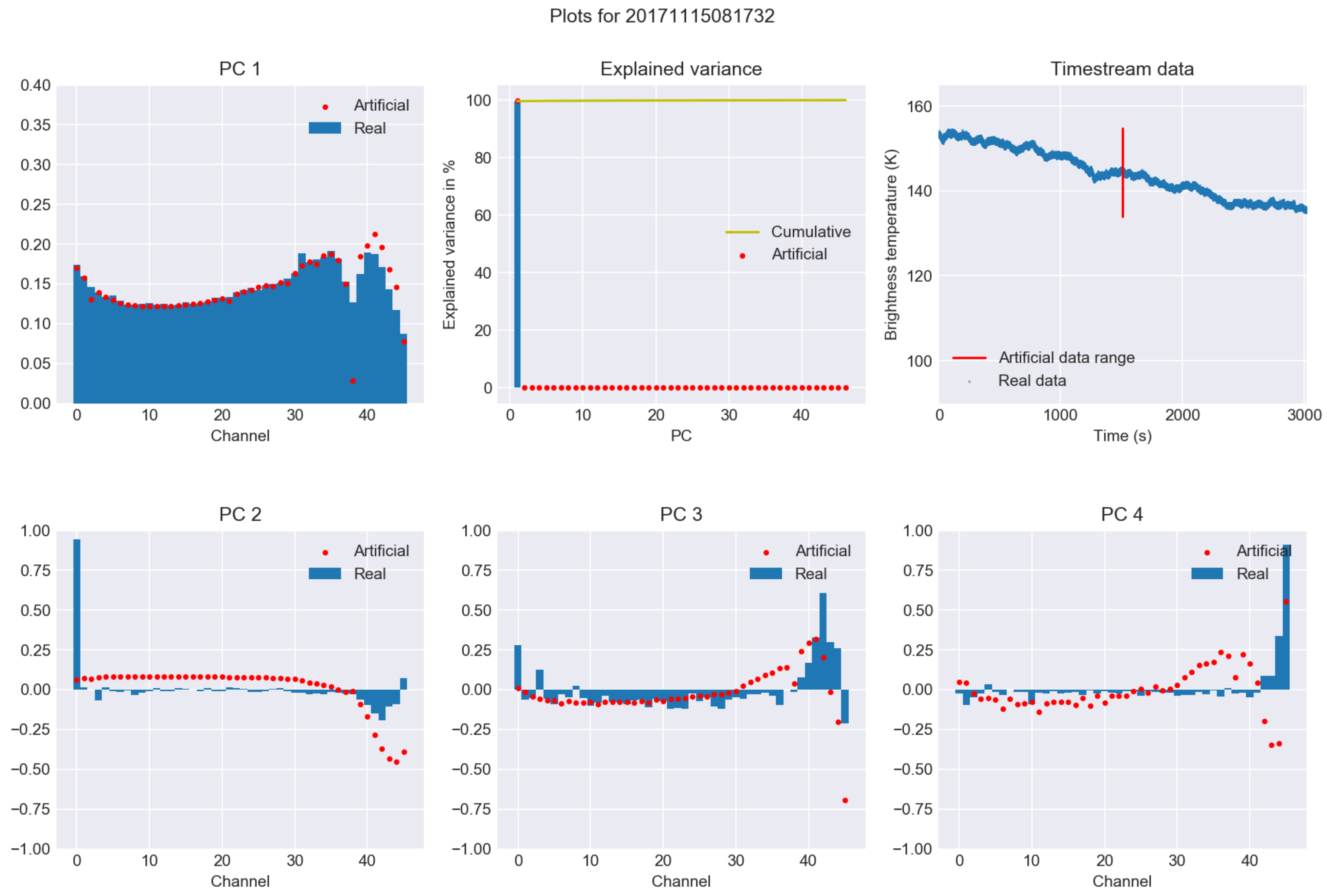
Plots for 20171115081732
PC 1
Artificial
Real
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Explained variance
Explained variance in %
100
80
60
40
20
0
Cumulative
Artificial
PC
Timestream data
Brightness temperature (K)
160
140
120
100
Artificial data range
Real data
0
1000
2000
3000
Time (s)
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Channel

Figure A.8

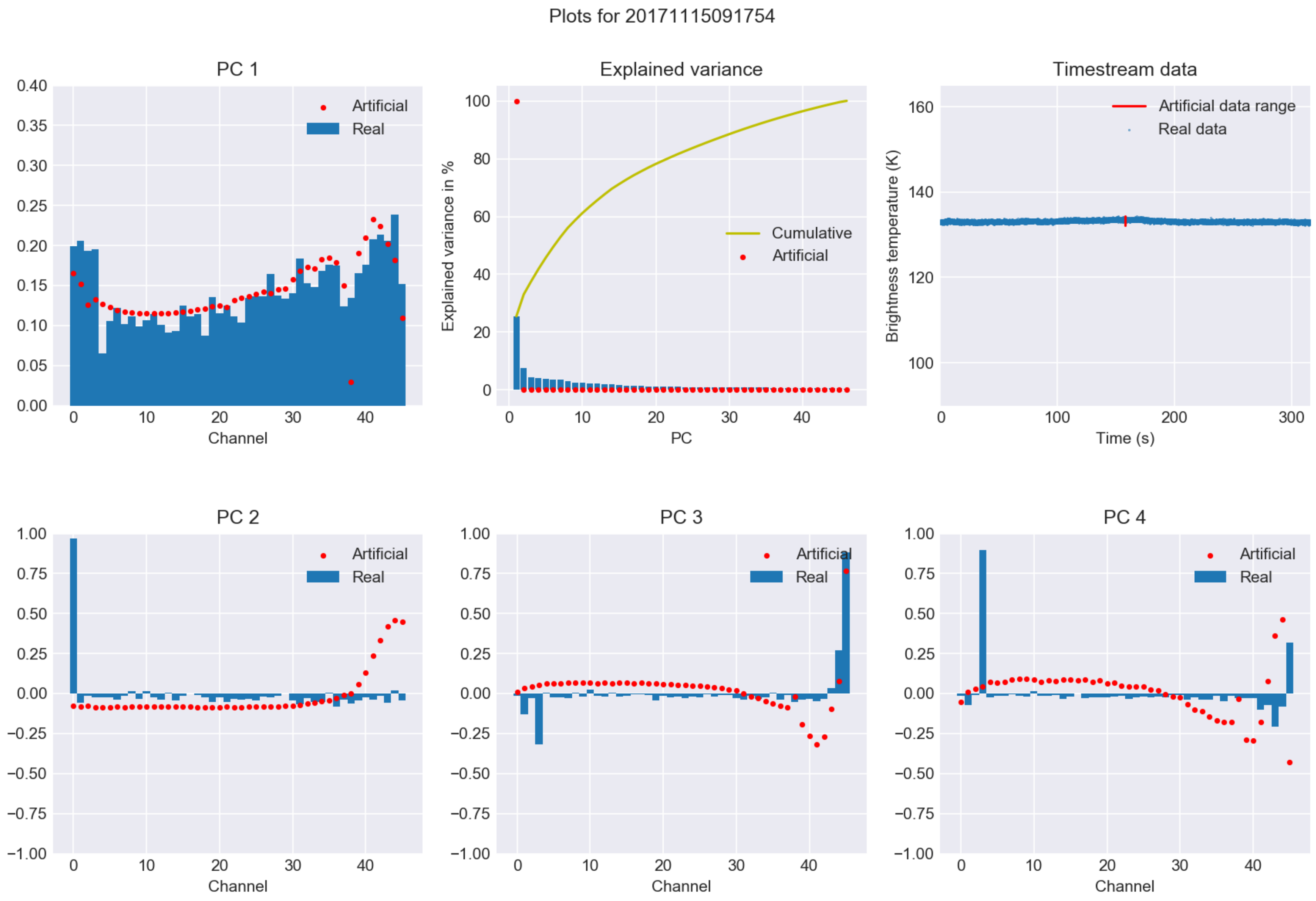
Plots for 20171115091754
PC 1
Artificial
Real
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Explained variance
Cumulative
Artificial
Explained variance in %
100
80
60
40
20
0
PC
Timestream data
Artificial data range
Real data
Brightness temperature (K)
160
140
120
100
0
100
200
300
Time (s)
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Figure A.9

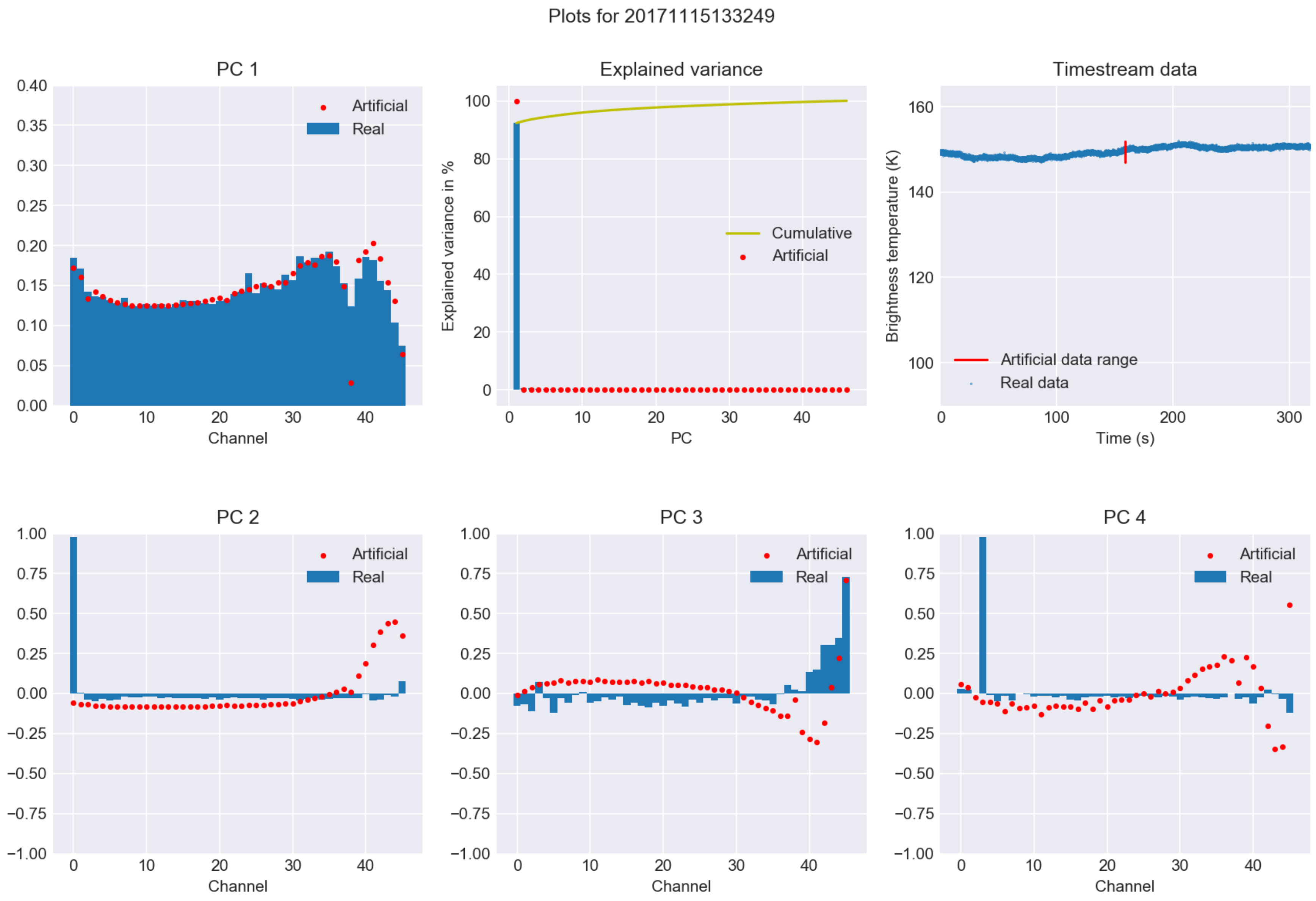
Plots for 20171115133249
PC 1
Artificial
Real
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Channel
Explained variance
Explained variance in %
100
80
60
40
20
0
Cumulative
Artificial
PC
Timestream data
Brightness temperature (K)
160
140
120
100
Artificial data range
Real data
0
100
200
300
Time (s)
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Channel

Figure A.10

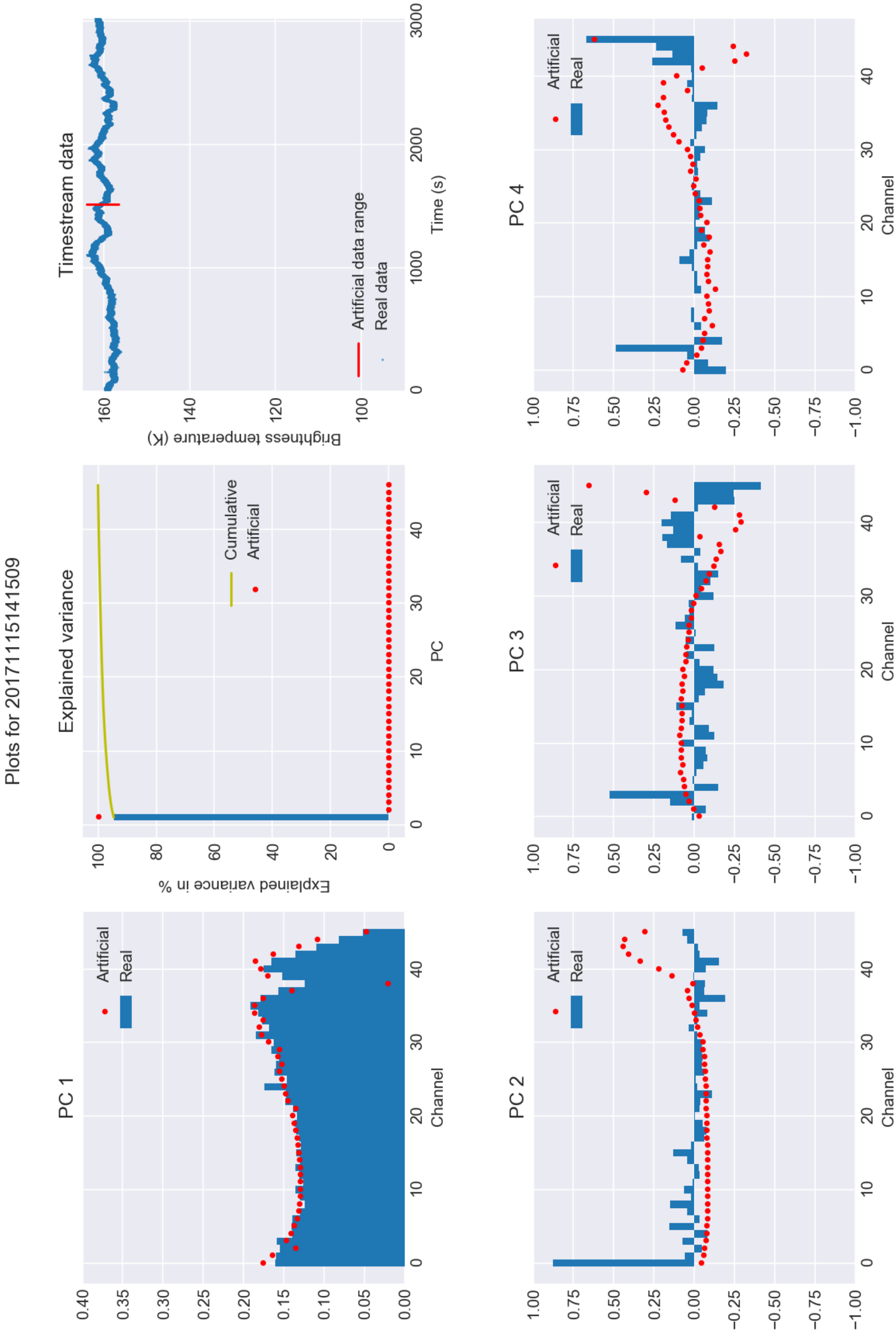
Plots for 20171115141509
PC 1
Artificial
Real
Channel
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Explained variance
Cumulative
Artificial
Explained variance in %
PC
100
80
60
40
20
0
Timestream data
Artificial data range
Real data
Brightness temperature (K)
Time (s)
160
140
120
100
1000
2000
3000
PC 2
PC 3
PC 4
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Figure A.11