Kernel Learning for Explainable Climate Science

Vidhi Lalchand [1]   Kenza Tazi[2,3]   Talay M Cheema[2]   Richard E. Turner[2]   Scott Hosking[3]

[1]Dept. of Physics, University of Cambridge, Cambridge, UK
[2]Dept. of Engineering, University of Cambridge, Cambridge, UK
[3]British Antarctic Survey, Cambridge, UK

Abstract

The Upper Indus Basin, Himalayas provides water for 270 million people and countless ecosystems. However, precipitation, a key component to hydrological modelling, is poorly understood in this area. A key challenge surrounding this uncertainty comes from the complex spatial-temporal distribution of precipitation across the basin. In this work we propose Gaussian processes with structured non-stationary kernels to model precipitation patterns in the UIB. Previous attempts to quantify or model precipitation in the Hindu Kush Karakoram Himalayan region have often been qualitative or include crude assumptions and simplifications which cannot be resolved at lower resolutions. This body of research also provides little to no error propagation. We account for the spatial variation in precipitation with a non-stationary Gibbs kernel parameterised with an input dependent lengthscale. This allows the posterior function samples to adapt to the varying precipitation patterns inherent in the distinct underlying topography of the Indus region. The input dependent lengthscale is governed by a latent Gaussian process with a stationary squared-exponential kernel to allow the function level hyperparameters to vary smoothly. In ablation experiments we motivate each component of the proposed kernel by demonstrating its ability to model the spatial covariance, temporal structure and joint spatio-temporal reconstruction. We benchmark our model with a stationary Gaussian process and a Deep Gaussian processes.

1 MOTIVATION

The Indus River is one of the longest rivers in Asia, sustaining the livelihoods of over 268 million people [Wester et al., 2019]. The river and its tributaries provide fresh water for drinking, domestic usage, industrial processes, and agriculture through the world's largest contiguous irrigation system [Basharat, 2019]. The river also delivers most of Pakistan's electricity through hydropower plants [Nie et al., 2021] and supports countless ecosystems and biodiversity hotspots [Xu et al., 2019].

Of all the rivers originating in the Himalayas, the Indus depends most strongly on water from snow and glacier found in the Upper Indus Basin (UIB) [Lutz et al., 2014]. Over 60% of the Indus' annual flow is attributed to the springtime melt of the snowpack and glaciers [Immerzeel et al., 2010]. As climate change progresses, these drivers are expected to be replaced by precipitation. This change will lead to on average less but more extreme variations in river flow and in turn more floods, landslides, and droughts [Huss et al., 2017].

With no robust adaption measures, financial losses and profound socio-economic consequences including food and water scarcity, mass migration and violent conflict are projected [Wester et al., 2019, Huss et al., 2017]. The scale of these ramifications is still unknown with the largest source of uncertainty attributed to precipitation. [Li et al., 2016, Wulf et al., 2016, Remesan and Holman, 2015, Meng et al., 2014, Andermann et al., 2011].

Previous research modelling and predicting precipitation in the Hindu Kush Karakoram Himalayan region are often qualitative or include assumptions and simplifications which cannot be resolved at lower resolutions [Dahri et al., 2016]. For example, the regional climates models from Coordinated Regional Climate
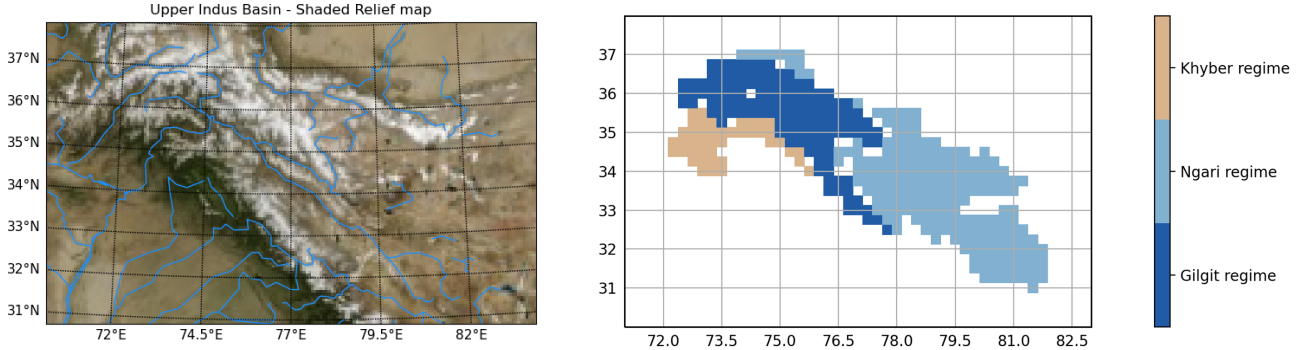
Figure 1: Left: Shaded-relief image of the topography of the Upper Indus Basin with natural water bodies overlaid. Right: Three precipitation regimes in the UIB identified through kmeans clustering. Regime names were chosen by the authors.

Downscaling Experiment for South Asia regularly overestimates historical precipitation in the Himalayas by over 100% for both winter and summer [Sanjay et al., 2017]. These models are also computationally expensive, lack error propagation, and generate large model-dependent uncertainty.

Gaussian processes (GPs) offer a versatile and interpretable way of studying and predicting precipitation in this area. Crucially, they offer two unique advantages which preclude several other modelling tools. First, GP estimates provide principled uncertainty quantification by design. One can derive concrete prediction intervals through the posterior predictive distribution; predictions accompanied by robust uncertainty estimates can be critical for downstream decision making. Secondly, GPs are a flexible prior of functions and allow one to encode specific inductive biases through kernel construction. For instance, one can encode properties like periodicity, smoothness or spatial heterogeneity through careful specification of the covariance kernel.

This work concerns interpretable kernel constructions for precipitation modelling in the UIB. We account for non-stationarity and spatial heterogeneity through a spatio-temporal kernel. In the next section, we give a brief overview of the GP framework along with non-stationary kernels. Section 3 presents experimental results from the case-study where we leverage spatio-temporal and non-stationary kernels.

2 BACKGROUND

GPs are a powerful probabilistic and non-parametric tool for modelling functions. They are fully specified by a mean and covariance function where the lat-

ter controls the inductive bias and support of functions under the prior. The choice of the covariance function (alternatively, kernel or kernel function) and in turn selecting the hyperparameters of the covariance function is jointly referred to as the *model selection* problem [Rasmussen and Williams, 2006] in GPs. A large cross-section of Gaussian process literature uses universal kernels like the squared exponential (SE) kernel along with automatic relevance determination (ARD) in high-dimensions. The SE-ARD kernel is a translation-invariant stationary kernel which gives infinitely smooth and differentiable samples in function space. The ARD framework operates by pruning away extraneous dimensions through contracting their inverse-lengthscales. The SE-ARD framework is the most commonly reported baseline for Gaussian process regression tasks. The SE-ARD kernel is given by,

$$k(\boldsymbol{x}, \boldsymbol{x}')_{\text{SE-ARD}} = \sigma_f^2 \exp\left\{ -\frac{1}{2} \sum_{d=1}^{D} \frac{(x_d - x'_d)^2}{\ell_d^2} \right\} \quad (1)$$

where $\boldsymbol{x}, \boldsymbol{x}' \in X$ are high-dimensional inputs $X \equiv \{\boldsymbol{x}_i\}_{i=1}^{N}$, each $\boldsymbol{x}_i \in \mathbb{R}^D$ and $\{\ell_d\}_{d=1}^{D}$ denotes a scalar lengthscale per dimension and σ_f^2 is usually a scalar amplitude.

Given observations $(X, \boldsymbol{y}) = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N}$ where y_i are noisy realizations of some latent function values f corrupted with Gaussian noise, $y_i = f(\boldsymbol{x}_i) + \epsilon_i$, $\epsilon_i \in \mathcal{N}(0, \sigma_n^2)$, let $k_\theta(\boldsymbol{x}_i, \boldsymbol{x}_j)$ denote a positive definite covariance function parameterised with hyperparameters θ. The generative model governing the data is given by,

$$f \sim \mathcal{GP}(0, k_\theta(., .))$$

$$p(\boldsymbol{y}|f) = p(y_i|f_i) = \prod_{i=1}^{N} \mathcal{N}(f_i, \sigma_n^2) \quad (2)$$

A Gaussian noise setting which we assume through out yields a closed-form marginal likelihood $p(\boldsymbol{y}|\theta)$. In the standard set-up the negative log marginal likelihood serves as the loss function against which θ is optimised, we call this procedure ML-II. The SE-ARD kernel with ML-II inference serves as the baseline model to benchmark performance for the variants proposed in this work. Learning occurs through adaptation of the hyperparameters (θ) of the covariance function. In the case of the SE-ARD kernel described above, $\theta = \{\sigma_f^2, \{\ell_d\}_{d=1}^D, \sigma_n^2\}$.

$$p(\boldsymbol{y}|\theta) = \int p(\boldsymbol{y}|f)p(\boldsymbol{f}|\theta)df \qquad (3)$$

$$= \int \mathcal{N}(0, K_f)\mathcal{N}(f, \sigma_n^2 \mathbb{I})df \qquad (4)$$

$$= \mathcal{N}(0, K_f + \sigma_n^2 \mathbb{I})$$

The matrix K_f denotes the kernel $k_\theta(\cdot, \cdot)$ evaluated at inputs X.

This works considers the non-stationary counterpart of the squared-exponential kernel, the Gibbs kernel [Gibbs, 1998] for the task of spatial precipitation modelling. In contrast to the SE-ARD kernel proposed above, the Gibbs kernel introduces input dependent lengthscales in each dimension. Instead of a single-point estimate ℓ_d per dimension we can consider $\ell_d(\boldsymbol{x})$, a lengthscale function. One can choose a parametric form for the dependence of the lengthscale on the inputs in each dimension but very often this dependence is not known a priori. The Gibbs kernel for multi-dimensional inputs is given by,

$$k_{\text{Gibbs}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \prod_{d=1}^D \sqrt{\frac{2\ell_d(\boldsymbol{x}_i)\ell_d(\boldsymbol{x}_j)}{\ell_d^2(\boldsymbol{x}_i) + \ell_d^2(\boldsymbol{x}_j)}} \times \qquad (5)$$

$$\exp\left\{ -\sum_{d=1}^D \frac{(x_i^{(d)} - x_j^{(d)})^2}{\ell_d^2(\boldsymbol{x}_i) + \ell_d^2(\boldsymbol{x}_j)} \right\}$$

Previous work, Heinonen et al. [2016] considers a formulation where the lengthscale function is modelled non-parameterically with a latent Gaussian process defined on the same inputs.

$$\hat{\ell} = \log(\ell_d) \sim \mathcal{GP}(0, k_\ell(., .)) \qquad (6)$$

$$\hat{\ell}(\boldsymbol{x}) = \log(\ell_d(\boldsymbol{x})) \sim \mathcal{N}(0, K_\ell) \qquad (7)$$

The hierarchical formulation makes the posterior and marginal likelihood analytically intractable. However, in the Gaussian likelihood setting one can consider a maximum-a-posteriori (MAP) solution [Kersting et al., 2007, Heinonen et al., 2016] by maximising,

$$\ell_{\text{MAP}} = \underset{\ell}{\text{argmax}} \log p(\boldsymbol{y}|\hat{\ell})p(\hat{\ell}) \qquad (8)$$

$$= \underset{\ell}{\text{argmax}} \log \mathcal{N}(\boldsymbol{y}|0, K_f + \sigma_n^2 \mathbb{I})\mathcal{N}(\hat{\ell}|0, K_\ell)$$

where ℓ_{MAP} is a vector of the size of the training inputs X (note that it is typical to work with the log of the marginal likelihood to avoid numerical underflow). Extrapolating the non-parametric lengthscale to test inputs entails estimating $p(\ell_\star|\ell_{\text{MAP}}, \boldsymbol{y})$ which we approximate by $\mathbb{E}(\ell_\star|\ell_{\text{MAP}}) = K_{\star\ell}K_\ell^{-1}\ell_{\text{MAP}}$ (expectation of a conditional Gaussian). Once the MAP solution has been learnt at the training inputs, the posterior predictive $p(\boldsymbol{f}^\star|\boldsymbol{y}, \ell_{\text{MAP}}, \ell_\star) \sim \mathcal{N}(\mu^\star, \Sigma^\star)$ is given by,

$$\mu^\star = K_{\star f}^T(K_f + \sigma_n^2 \mathbb{I})^{-1}\boldsymbol{y}$$

$$\Sigma^\star = K_{\star\star} - K_{\star f}^T(K_f + \sigma_n^2 \mathbb{I})^{-1}K_{f\star} \qquad (9)$$

where K_f is based on the evaluation of eq. (5) on the training inputs (using ℓ_{MAP}) and $K_{\star f}$ is based on the evaluation of k_{Gibbs} on test and training inputs using ℓ_{MAP} and $\mathbb{E}(\ell_\star|\ell_{\text{MAP}})$ respectively.

The latent GP parameterisation allows for extremely flexible modelling where samples from the posterior function space 'adapt' to the varying spatial dynamics inherent in the data.

3 CASE-STUDY: ACCURATE PRECIPITATION MODELLING IN THE UIB

We conduct three experiments to highlight different features of the kernel composition. In the spatial regression task, we benchmark the non-stationary spatial covariance kernel based on the Gibbs construction against stationary baselines. The temporal extrapolation task uses a locally periodic kernel to fit the univariate precipitation dynamics at single spatial points over time. The spatio-temporal task considers an additive kernel with components acting on the spatial and temporal slice of the inputs to model dynamics over space and time. The models are implemented using `GPyTorch` [Gardner et al., 2018] and data is plotted using `xarray` [Hoyer and Hamman, 2017].

Data: in this case study, we use the 5th ECMWF Reanalysis (ERA5) dataset [Hersbach et al., 2020]. Reanalysis blends historical observations from surface, sonde and satellite measurements with numerical weather forecasting models. Through this data assimilation, the models create a past record of historical climate at high temporal and spatial resolution. ERA5 runs from 1959 to the present day over 0.25° grid and assimilates data from a large number of sources. Each datapoint represents the average monthly precipitation in mm/day over a gridbox.

Precipitation in the UIB is complex and this is reflected in the ERA5 dataset. In this area, precipitation is driven by two major atmospheric events: the Indian
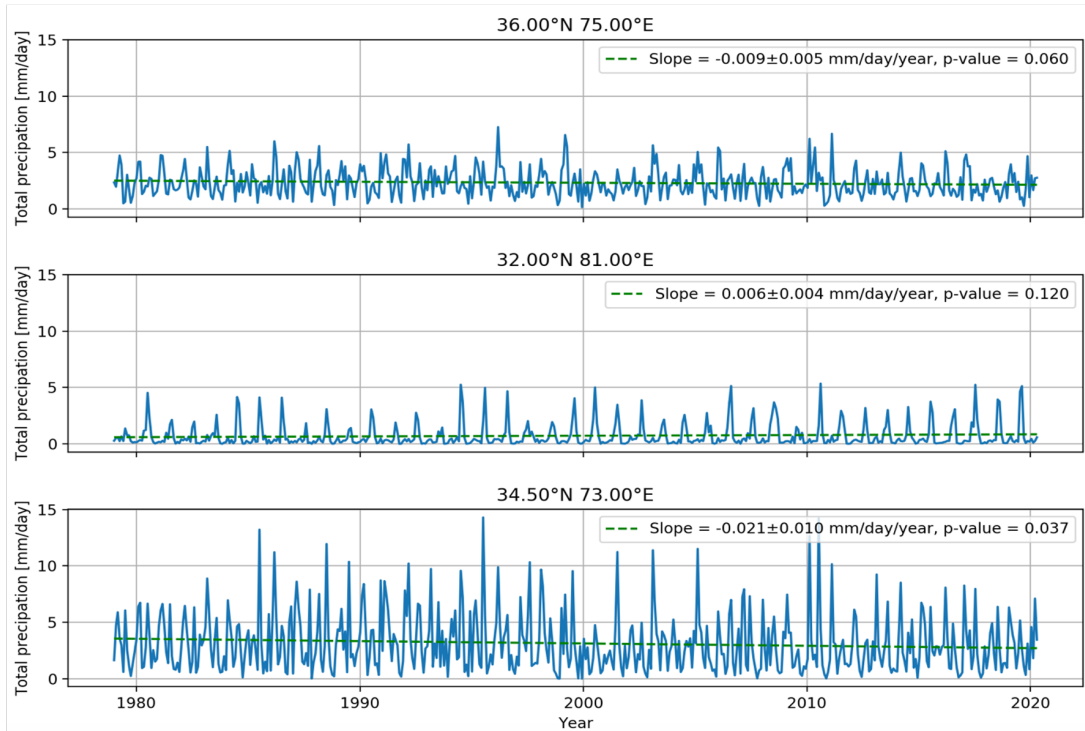
Figure 2: Timeseries drawn from each precipitation regime shown in Figure 1 (top: Gilgit, middle: Ngari, bottom: Khyber). Note that there are no obvious trends in average or extreme precipitation during this period.

Summer Monsoon (ISM) and the western fronts (Westerlies). The ISM brings rain from June to September. The ISM reaches the south-eastern UIB first and has a decreasing contribution to the annual rainfall in the North-West direction.

The Westerlies are strongest in the winter from December to April, peaking in March [Dahri et al., 2016]. The relative contribution of westerly fronts increases from the South-East to the North-West of the basin. As a consequence, the eastern UIB receives up to 70% of its annual rainfall from the summer whereas the western UIB receives 40-60% of its precipitation during the winter [Dahri et al., 2016].

The complexity of this distribution is illustrated in Figure 1 and 2 where three timeseries are sampled from the three characteristic precipitation regimes identified through K-means clustering.

3.1 SPATIAL REGRESSION

We model precipitation in the UIB at a single point in time focusing on accurately capturing the spatial variation of precipitation. The stationary model and the deep GP uses a 2D SE-ARD kernel learning stationary hyperparameters (with ML-II inference) which are constant across space. In the Gibbs formulation (eq. (5)) we learn a lengthscale per spatial point using

MAP estimation (as described in section 2). Figure 4 shows the best mean prediction from 10 splits for each of the baselines. The 2-layer deep GP [Damianou and Lawrence, 2013] with skip connections performs worse than the standard baseline in terms of prediction error. The high-precipitation areas in the Khyber zone are best captured by the Gibbs kernel. It is interesting to note that the predictive intervals for the DGP yield higher predictive densities for test data despite inferior predictive means. The Gibbs formulation gives the best trade-off in terms of reconstruction error assessed against ground-truth and quality of predictive uncertainty.

Metric	SE-ARD	DGP (L2)	Gibbs
RMSE	0.353 ± 0.014	0.603 ± 0.018	$\mathbf{0.271 \pm 0.014}$
NLPD	0.406 ± 0.063	$\mathbf{0.010 \pm 0.011}$	$\mathbf{0.019 \pm 0.016}$

Table 1: 2D Spatial regression on precipitation data over the entire UIB over the month of Jan, 2000. We report Root Mean Squared Error (RMSE) and Negative Log Predictive Density (NLPD) on held-out test data over 10 splits.
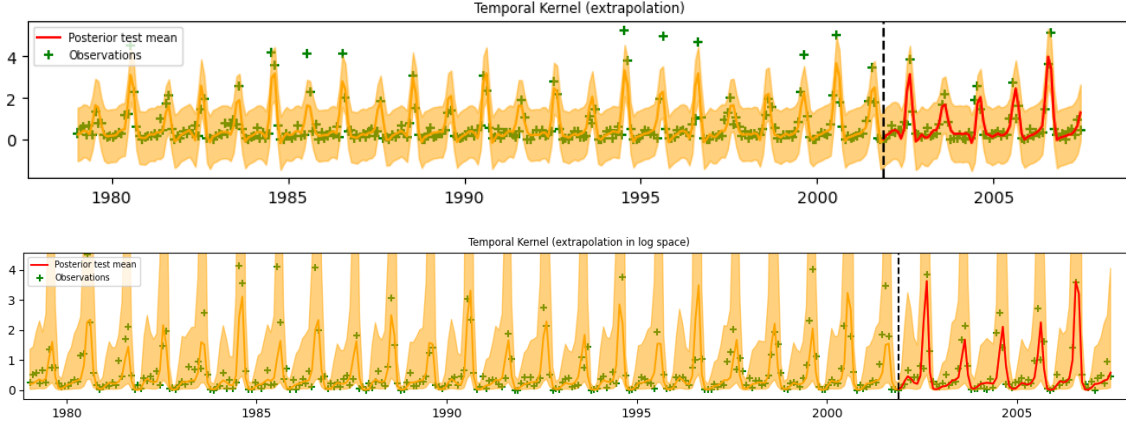
Figure 3: Predicting the precipitation trend with a locally periodic kernel. The dotted line demarcates the training and test regimes. The error bars denote 95% confidence intervals around the predicted mean. Some outliers remain uncaptured by the Gaussian prediction intervals, indicating mild non-Gaussianity in the targets. In the bottom plot we regress the outputs in log-space and visualise the 1d log-normal distribution with the respective quantiles.

3.2 TEMPORAL EXTRAPOLATION

We regress on univariate time inputs (monthly observations) to capture the local periodicity exhibited in the precipitation dynamics. In order to moderate the changing amplitudes of the peaks we add flexibility by multiplying with a SE-ARD kernel (acting on univariate time inputs). The periodic kernel we use is given by,

$$k_{\text{PER}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma_f^2 \exp\left\{-\frac{2\sin^2(\pi|\boldsymbol{x}_i - \boldsymbol{x}_j|/p)}{\ell^2}\right\} \tag{10}$$

Figure 3 depicts the prediction performance over train and test where we train on years 1979-2002. The predictions under the log-normal distribution ensure positivity (as desired), further the heavy-tailed non-Gaussian prediction intervals capture the outlying values which elude the symmetric Gaussian intervals. The test RMSE and NLPD for the fits are given by (0.5536 vs. 0.5328) and (1.3248 vs. 0.9721).

3.3 SPATIO-TEMPORAL REGRESSION

In this experiment we consider the task of predicting dynamics across space and time (3D inputs) for a 1 month ahead forecast. We train on January to April 2000 and test on the month of May. We wanted to constrain the training data set to a moderate size so as to execute exact GP inference for the baseline case.

For the stationary (shallow) model, the spatio-temporal kernel is formed by adding together the spatial and temporal components with one important in-novation. In the earlier sections the temporal component acted solely on the time dimension while the spatial component acted solely on the spatial dimensions (latitude and longitude). Denoting the time and spatial coordinates as $\boldsymbol{x}^{\text{t}}, \boldsymbol{x}^{\text{lat}}, \boldsymbol{x}^{\text{lon}}$ respectively, the kernel for the stationary model is given by,

$$k_{\text{stat.}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \\ \underbrace{k_{\text{SE-ARD}}((\boldsymbol{x}_i^{\text{lat}}, \boldsymbol{x}_i^{\text{lon}}), (\boldsymbol{x}_j^{\text{lat}}, \boldsymbol{x}_j^{\text{lon}})) \times k_{\text{PER}}(\boldsymbol{x}_i^t, \boldsymbol{x}_j^t)}_{\text{temporal}} \\ + \underbrace{k_{\text{SE-ARD}}((\boldsymbol{x}_i^{\text{lat}}, \boldsymbol{x}_i^{\text{lon}}), (\boldsymbol{x}_j^{\text{lat}}, \boldsymbol{x}_j^{\text{lon}}))}_{\text{spatial}} \tag{11}$$

We weave the spatial dimensions into the temporal kernel by making the RBF component dependent on the latitude and longitude, this allows the amplitude/scale of the periodic component to vary for different spatial regions. Note that this is still a stationary kernel. For the non-stationary extension in this setting we keep the temporal component identical to the stationary formulation eq. (11) but replace the spatial component with the Gibbs kernel eq. (5).

$$k_{\text{non-stat.}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \\ \underbrace{k_{\text{SE-ARD}}((\boldsymbol{x}_i^{\text{lat}}, \boldsymbol{x}_i^{\text{lon}}), (\boldsymbol{x}_j^{\text{lat}}, \boldsymbol{x}_j^{\text{lon}})) \times k_{\text{PER}}(\boldsymbol{x}_i^t, \boldsymbol{x}_j^t)}_{\text{temporal}} \\ + \underbrace{k_{\text{Gibbs}}((\boldsymbol{x}_i^{\text{lat}}, \boldsymbol{x}_i^{\text{lon}}), (\boldsymbol{x}_j^{\text{lat}}, \boldsymbol{x}_j^{\text{lon}}))}_{\text{spatial}} \tag{12}$$

where k_{Gibbs} is parameterised by 2 non-parametric lengthscale processes - $\ell_{\text{lat}}(\boldsymbol{x}^{\text{spatial}})$ and $\ell_{\text{lon}}(\boldsymbol{x}^{\text{spatial}})$ acting on the 2d spatial dimensions as inputs. We enclose the regression results for the 1-month ahead fore-
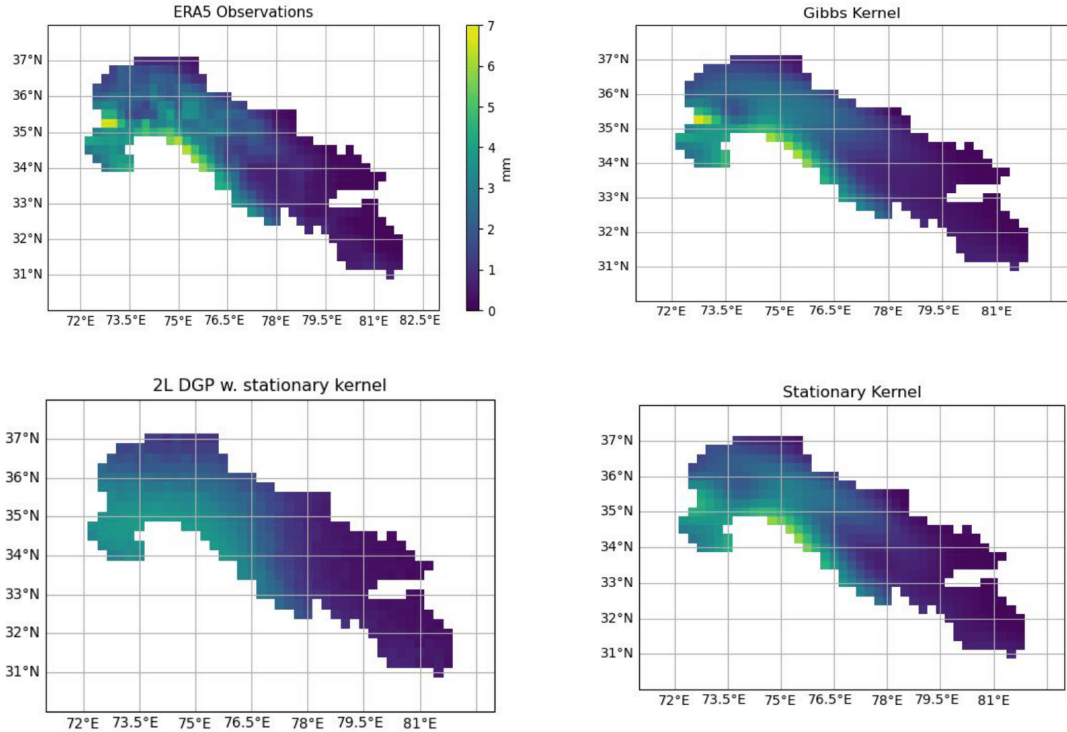
Figure 4: ERA5 data and 2D precipitation modelling over inputs (lat, lon) for the non-stationary Gibbs kernel, and two-layer DGP and shallow GP with stationary kernels. Table 1 provides test predictive performance metrics over 10 splits of the data.

cast based on the baseline and non-stationary construction in fig. 5.

We do not present results for the deep GP in this sections as they were found to yield substandard results in section 3.1.

4 CONCLUSIONS AND FURTHER WORK

We demonstrate kernel composition in modelling the spatial and temporal precipitation dynamics over the UIB. In the spatial setting, the non-stationary kernel does much better than a deep GP and the stationary baseline. More comprehensive and complex experiments will provide answers to how these methods can be best be used in 'real world' settings and, scientifically, what drives precipitation in the UIB.

Further investigation will seek to better understand historical precipitation distribution in this area. In particular, a detailed analysis of the Gibbs hyperparameter functions are needed over longer training and testing intervals. This would require making the non-stationary formulation compatible with sparse GPs relying on inducing points [Titsias, 2009]. Future iterations will also consider more features that are pre-

dictive of precipitation such as elevation, slope and large scale atmospheric variables. The Gibbs hyperparamaters will give us insight into where and how these features are influencing rain and snowfall in basin. Model comparison metrics like BIC (Bayesian information criterion) will help understand more complex non-stationary constructions in light of model complexity.

Next, precipitation change under different climate scenarios will be considered. ERA5 climatic variables will be swapped with global climate model outputs. Although not explored in this paper, the non-stationarity of precipitation with respect to climate change will become more important in the future. This non-stationarity should be in part captured by the large scale atmospheric variables previously mentioned. However, a non-stationary kernel could also have a role to play here.

Finally, the Bayesian nature of these models will give a more accurate and informed quantification of uncertainty than regional climate models that uses ensemble spread as proxy for uncertainty. In the non-stationary setup, the hyperaparemeter distributions could also point scientists to the most significant drivers of prediction variability.
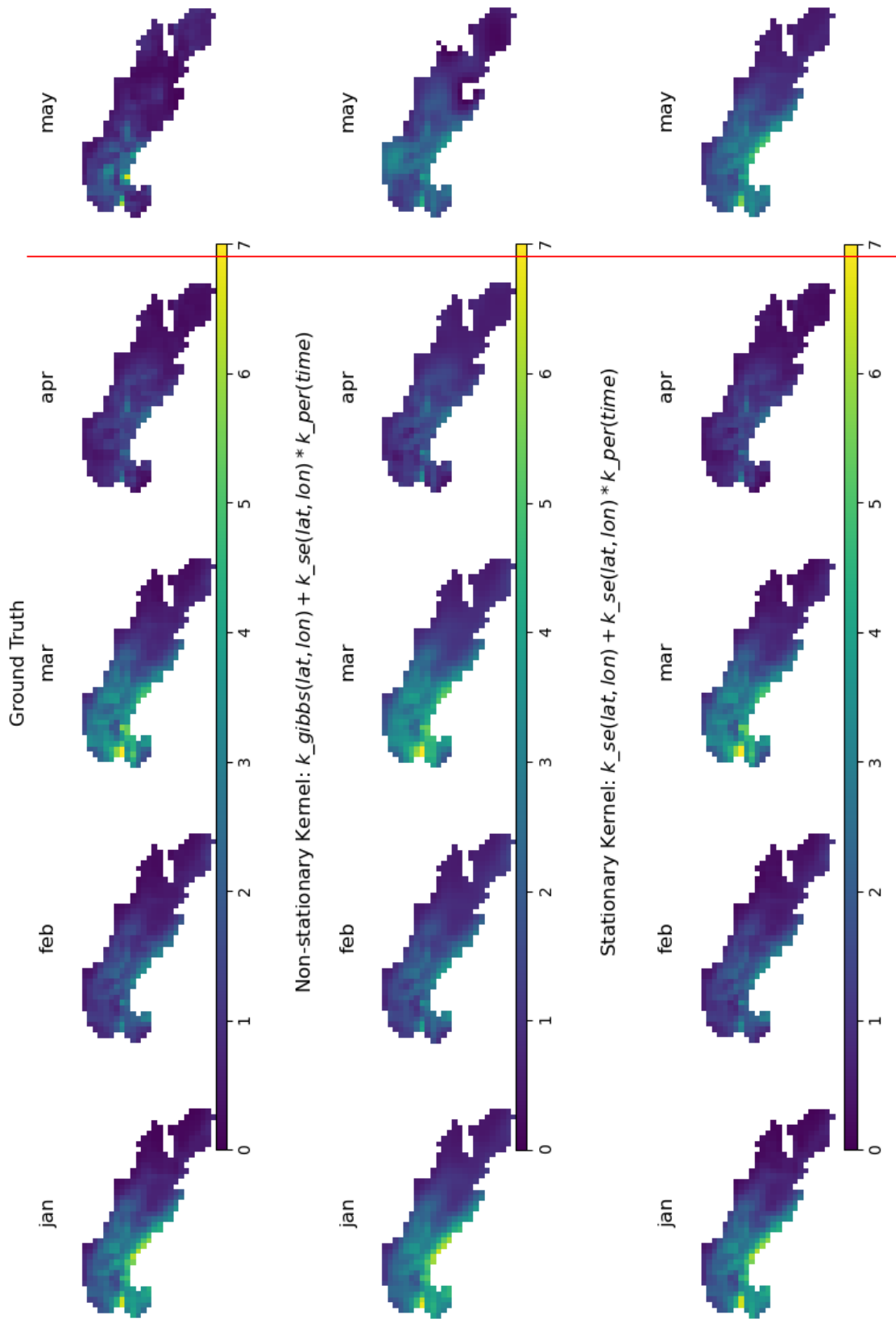
Figure 5: 3D regression for spatio-temporal precipitation modelling. The red line demarcates the training and test months. The non-stationary kernel yielded a an average test RMSE of 0.9426 across three runs vs. 1.1086 for the stationary kernel.

References

Cristoff Andermann, Stéphane Bonnet, and Richard Gloaguen. Evaluation of precipitation data sets along the himalayan front. *Geochemistry, Geophysics, Geosystems*, 12(7), 2011.

Muhammad Basharat. Water management in the indus basin in pakistan: challenges and opportunities. In *Indus River Basin*, pages 375–388. Elsevier, 2019.

Zakir Hussain Dahri, Fulco Ludwig, Eddy Moors, Bashir Ahmad, Asif Khan, and Pavel Kabat. An appraisal of precipitation distribution in the high-altitude catchments of the indus basin. *Science of the Total Environment*, 548:289–306, 2016.

Andreas Damianou and Neil D Lawrence. Deep gaussian processes. In *Artificial intelligence and statistics*, pages 207–215. PMLR, 2013.

Jacob R Gardner, Geoff Pleiss, David Bindel, Kilian Q Weinberger, and Andrew Gordon Wilson. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. In *Advances in Neural Information Processing Systems*, 2018.

Mark N Gibbs. *Bayesian Gaussian processes for regression and classification*. PhD thesis, Citeseer, 1998.

Markus Heinonen, Henrik Mannerström, Juho Rousu, Samuel Kaski, and Harri Lähdesmäki. Nonstationary gaussian process regression with hamiltonian monte carlo. In *Artificial Intelligence and Statistics*, pages 732–740. PMLR, 2016.

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

S. Hoyer and J. Hamman. xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5(1), 2017. doi: 10.5334/jors.148. URL https://doi.org/10.5334/jors.148.

Matthias Huss, Bodo Bookhagen, Christian Huggel, Dean Jacobsen, Raymond S Bradley, John J Clague, Mathias Vuille, Wouter Buytaert, Daniel R Cayan, Gregory Greenwood, et al. Toward mountains without permanent snow and ice. *Earth's Future*, 5(5):418–435, 2017.

Walter W Immerzeel, Ludovicus PH Van Beek, and Marc FP Bierkens. Climate change will affect the asian water towers. *science*, 328(5984):1382–1385, 2010.

Kristian Kersting, Christian Plagemann, Patrick Pfaff, and Wolfram Burgard. Most likely heteroscedastic gaussian process regression. In *Proceedings of the 24th international conference on Machine learning*, pages 393–400, 2007.

Baofu Li, Yaning Chen, Zhongsheng Chen, Heigang Xiong, and Lishu Lian. Why does precipitation in northwest china show a significant increasing trend from 1960 to 2010? *Atmospheric Research*, 167:275–284, 2016.

AF Lutz, WW Immerzeel, AB Shrestha, and MFP Bierkens. Consistent increase in high asia's runoff due to increasing glacier melt and precipitation. *Nature Climate Change*, 4(7):587–592, 2014.

Lei Meng, Di Long, Steven M Quiring, and Yanjun Shen. Statistical analysis of the relationship between spring soil moisture and summer precipitation in east china. *International Journal of Climatology*, 34(5):1511–1523, 2014.

Yong Nie, Hamish D Pritchard, Qiao Liu, Thomas Hennig, Wenling Wang, Xiaoming Wang, Shiyin Liu, Santosh Nepal, Denis Samyn, Kenneth Hewitt, et al. Glacial change and hydrological implications in the himalaya and karakoram. *Nature reviews earth & environment*, 2(2):91–106, 2021.

Carl Edward Rasmussen and Christopher KI Williams. Gaussian processes for machine learning (adaptive computation and machine learning), 2006.

Renji Remesan and Ian P Holman. Effect of baseline meteorological data selection on hydrological modelling of climate change scenarios. *Journal of Hydrology*, 528:631–642, 2015.

Jayanarayanan Sanjay, Raghavan Krishnan, Arun Bhakta Shrestha, Rupak Rajbhandari, and Guo-Yu Ren. Downscaled climate change projections for the hindu kush himalayan region using cordex south asia regional climate models. *Advances in Climate Change Research*, 8(3):185–198, 2017.

Michalis Titsias. Variational learning of inducing variables in sparse gaussian processes. In *Artificial intelligence and statistics*, pages 567–574. PMLR, 2009.

Philippus Wester, Arabinda Mishra, Aditi Mukherji, and Arun Bhakta Shrestha. *The Hindu Kush Himalaya assessment: mountains, climate change, sustainability and people*. Springer Nature, 2019.

Hendrik Wulf, Bodo Bookhagen, and Dirk Scherler. Differentiating between rain, snow, and glacier contributions to river discharge in the western himalaya

using remote-sensing data and distributed hydrological modeling. *Advances in Water Resources*, 88: 152–169, 2016.

Jianchu Xu, Ruchi Badola, Nakul Chettri, Ram P Chaudhary, Robert Zomer, Bharat Pokhrel, Syed Ainul Hussain, Sunita Pradhan, and Rebecca Pradhan. Sustaining biodiversity and ecosystem services in the hindu kush himalaya. In *The Hindu Kush Himalaya Assessment*, pages 127–165. Springer, 2019.