MeteoGalEus: An Iberian Multilingual Weather Dataset in Galician, Euskera, and Spanish

A. Vivel-Couso[*1], N. Z. Pramata[*2], D. Robredo[1], A. Soroa[2], J.M. Alonso-Moral[1]
[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS),
Universidade de Santiago de Compostela, Santiago de Compostela, Spain
{ainhoa.vivel.couso, david.robredo.manuel, josemaria.alonso.moral}@usc.es
[2]Basque Center for Language Technology (HiTZ Center - Ixa),
University of the Basque Country UPV/EHU, Donostia, Spain
npramata001@ikasle.ehu.eus, a.soroa@ehu.eus

Abstract

This paper introduces MeteoGalEus, a multilingual weather dataset that combines meteorological observations from two Spanish regional agencies, Euskalmet and MeteoGalicia. The dataset contains daily records spanning 4 years and 3 months, with aligned observations for both sources. MeteoGalEus captures key meteorological variables including temperature, wind and state of the sky. The dataset is provided in a structured format, facilitating data analysis and integration, with textual forecasts available in the official languages for each region (i.e., Galician and Spanish for MeteoGalicia; Euskera and Spanish for Euskalmet). By merging and harmonizing data from two regional agencies, MeteoGalEus is a unique resource for cross-regional weather analysis and multilingual climate studies. This dataset is suited for tasks requiring high-quality, aligned, and standardized weather data across multiple languages and regions. We conducted baseline experiments using LLaMA-based models in both zero-shot and fine-tuned settings to illustrate the use of MeteoGalEus for natural language generation (NLG). Fine-tuning led to consistent improvements across all metrics, with BERTScore increasing from 0.68 to 0.79, ROUGE from 0.20 to 0.35, and BLEU from 0.02 to 0.17 in the best-performing model. The experiments show how MeteoGalEus can be taken as a benchmark for multilingual and cross-regional NLG tasks.

Keywords: Corpus, Evaluation Methodologies, Less-Resourced Languages, Multilinguality, Validation of LRs, Quality Assurance, Natural Language Generation

1. Introduction

High-quality, standardized multilingual weather data are essential for developing climate studies, meteorological modeling, and cross-regional analyses that can guide informed decisions and help mitigate undesired negative effects (Gowdy, 2024). However, weather observations of different regions are often collected and stored by multiple agencies in varying formats and languages, which complicates integration and large-scale studies. In the Iberian Peninsula, for example, regional meteorological agencies provide valuable datasets, but differences in data content, representation, language, and temporal coverage continue to pose challenges for researchers seeking comparable data. To address these challenges, we introduce the Iberian Multilingual Weather Dataset (MeteoGalEus), a curated dataset combining daily meteorological observations from the regional meteorological agency of Galicia (MeteoGalicia[1]) and the Euskera Agency of Meteorology (Euskalmet[2]). Thus, MeteoGalEus supports cross-regional and multilingual weather research, enabling applications in climate trend analysis, weather forecasting, and multilingual Natural Language Processing (NLP).

This paper presents the construction, structure, and content of MeteoGalEus. We highlight its potential as a resource for the research community. In addition, we provide baseline analyses and evaluation metrics to show coverage, consistency, and usability of the the dataset. Baselines are intended not only as reference points but also as a way to illustrate the complexity of cross-regional and multilingual weather data integration.

The MeteoGalEus dataset and the code used for analysis, training, testing, and evaluation are publicly available at: https://gitlab.citius.gal/ainhoa.vivel/meteogaleus/-/tree/LREC

2. Related Work

Weather forecasting and meteorological analysis are central topics in data-driven research, and several benchmark datasets have been introduced to support work in this area. Early efforts include SUMTIME (Sripada et al., 2003) and WEATHERGOV (Liang et al., 2009), which provide structured meteorological records paired with textual summaries for Natural Language Generation (NLG) tasks (Gatt and Krahmer, 2018). Both resources are released in English only.

[*]These authors contributed equally to this work.
[1]https://www.meteogalicia.gal/web/home
[2]https://www.euskalmet.euskadi.eus/hasiera

Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026), pages 9905–9919
11–16 May 2026.

More recently, region-specific resources have emerged. MeteoGalicia-ES (González-Corbelle et al., 2022) and MeteoGalicia-GL[3] provide 3033 records each of tabular meteorological observations with human-written forecasts in Spanish and Galician, respectively, curated by the official meteorological agency of Galicia. Even though MeteoGalicia-ES, MeteoGalicia-GL and MeteoGalEus share a common source, the datasets differ in both temporal coverage and content. While MeteoGalicia-ES/GL include only information about the state of the sky, MeteoGalEus includes additional variables about temperature and wind.

Similarly, Euskalmet provides structured observations in Spanish and Euskera for the Basque Country, which are incorporated and harmonized in MeteoGalEus to create a consistent cross-regional dataset. Meteorological observations from Euskalmet are publicly available through Open Data Euskadi[4], i.e., the open data portal developed by the Euskera Government. Notice that, Open Data Euskadi represents an important initiative for promoting open access to regional information and supporting research and innovation.

Despite these advances, existing datasets are often limited in temporal coverage, language diversity, or interoperability across regions. To our knowledge, no publicly available resource harmonizes weather records from multiple regional agencies while maintaining multilingual consistency. In particular, there are no meteorological datasets that include data in Euskera. Moreover, we are not aware of any other multilingual meteorological datasets incorporating languages with limited resources. To address this gap, we created a new dataset by extending MeteoGalicia with data from Euskalmet and adding content in Euskera. Accordingly, MeteoGalEus aligns and standardizes daily meteorological data from Galicia and the Basque Country in three languages (Spanish, Galician, and Euskera), providing a unique resource for cross-regional, multilingual, and multimodal research.

Meteorological data has traditionally been utilized for NLG tasks, especially with data-to-text template-based systems, where predefined templates are exploited by NLG pipelines to automatically produce reports based on structured data inputs (Reiter and Dale, 2000; Reiter, 2024). For example, MultiMeteo (Coch, 1998) is a system that generates weather forecasts in multiple languages. In addition, GALiWeather[5] is the first NLG system devoted to natural language forecasts in Galicia (Ramos-Soto et al., 2015).

However, with the recent advancements in Artificial Intelligence (AI), particularly in natural language processing, end-to-end NLG systems are becoming more and more popular (Dušek et al., 2020). Initiatives such as CLLMate (Li et al., 2024), ClimateLLM (Li et al., 2025), and SST-LLM (Hu et al., 2025) exemplify the growing integration of large language models (LLMs) in meteorological and climate science applications. These developments illustrate the shift from template-based approaches to more dynamic, LLM-driven approaches for processing and interpreting complex data in meteorology and climate science, highlighting the relevance and utility of meteorological data for NLG tasks in the context of forecasting, multilingual reporting, deploying AI-driven predictive systems, etc.

3. Dataset

MeteoGalEus spans four years and three months, covering a total of 1557 days from November 25, 2020, to March 1, 2025. For each date, the dataset includes entries from two regions, ensuring balanced representation: half of the records originate from Galicia and half from the Basque Country, resulting in 3084 weather entries. The dataset includes only dates present in both sources, guaranteeing temporal alignment across records. Each entry contains key meteorological variables: temperature, wind speed and direction, and state of the sky for 33 to 36 cities in each of the two regions. Forecasts are provided in Spanish, together with one of the two co-official languages (Galician or Euskera), depending on the region of origin. By harmonizing and standardizing observations from two regional agencies, MeteoGalEus supports cross-regional and multilingual weather research, enabling applications in climate trend analysis, weather forecasting, and multilingual natural language processing.

3.1. Languages and Regions

The MeteoGalEus dataset spans two regions of Spain with distinct linguistic and cultural contexts. From MeteoGalicia, it includes forecasts in Spanish and Galician, covering the Galicia region, while from Euskalmet it provides information in Spanish and Euskera, representing the Basque Country. By incorporating the official language of Spain (Spanish) alongside the co-official language of each region, the dataset reflects the natural multilingual character of weather communication in these territories. While the forecasts are bilingual, the meteorological data are provided exclusively in Spanish, and all keys in the JSON files are standardized in English to ensure consistency and facilitate computational processing.

[3] https://gitlab.citius.gal/gsi-nlg/meteogalicia-es

[4] https://opendata.euskadi.eus/

[5] https://citius.usc.es/transferencia/demostradores-tecnoloxicos/GALiWeather

3.2. Data Structure

MeteoGalEus is distributed in JSON format (see examples in Appendix A), combining both meteorological data and linguistic content. It includes the date of observation and key meteorological variables such as state of the sky, temperature, and wind across multiple locations. Values correspond to daily predictions rather than actual measurements. Thus, MeteoGalEus contains both structured numerical and categorical data (e.g., temperature, wind, state of the sky) along with textual data (weather forecasts). Each entry in the dataset corresponds to a single day and contains both textual forecasts and meteorological measurements for a specific regional agency:

- `id`. Unique identifier that preserves trace to the original source (e.g., euskalmet_2020-11-25 or 3158). This way the dataset is endowed with traceability, allowing users to easily locate the raw data in the source agency.
- `from`. The source agency (MeteoGalicia or Euskalmet).
- `autonomous_community`. The geographic area (Galicia or the Basque Country).
- `date`. The date of the forecast in ISO 8601 format (e.g., 2020-11-25).
- `forecast`. The textual pieces of information written by human experts (i.e., professional meteorologists affiliated to each agency) in Spanish, Basque and Galician (keys ES, EU, GL, respectively).
- `meteorological_data`. A structured list of measurements for multiple locations within the corresponding autonomous community. For each city we store the subregion name (`zone`), the city name (`location`), daily maximum (`max`) and minimum (`min`) temperatures, wind `direction` and `speed` measurements for three time `periods` (morning, afternoon, night), and the state of the sky for a location (`state_of_the_sky`), including cloud coverage and precipitation type these periods.

3.3. Collection and Alignment

The MeteoGalEus dataset was created through a multi-step process aimed at ensuring consistency and comparability across sources. First, we collected daily meteorological data from two official regional agencies: MeteoGalicia (covering Galicia, with forecasts in Spanish and Galician) and Euskalmet (covering the Basque Country, with forecasts in Spanish and Basque). We discarded information that was not present in both datasets, and limited the dataset to days for which both agencies provided records, ensuring that the resulting dataset contains only synchronized observations from the two regions. After collecting the data, we harmonized the values from Euskalmet and MeteoGalicia into a common format. While Euskalmet provided hourly forecasts for several variables, MeteoGalicia reported data in three daily time slots: morning, afternoon, and night. Hence, we converted the hourly data from Euskalmet to time slots, taking into account daylight saving time changes[6].

In addition, MeteoGalicia provided observations for 33 locations throughout the Galicia region, while Euskalmet listed 319 locations across the Basque Country. It was important to identify a comparable number of locations from both datasets. We filtered out locations with no observational data and selected 36 locations from Euskalmet that represented all three main zones in the Basque Country. We also standardized the zone names in Euskalmet data, going from geographical names (e.g. Coast Zone) to provincial names (e.g. Bizkaia or Gipuzkoa, depending on the location). In the case of wind speed, the values were discretized into five categories[7]. As for the wind direction, Euskalmet follows 16-point compass system while MeteoGalicia follows 8-point cardinal system. We made harmonization to use the 8-point cardinal system for both datasets.

Finally, we merged all data into a single, structured JSON format, preserving all meteorological common variables as well as the multilingual forecast descriptions written by human meteorologists. The alignment process allows direct cross-regional and cross-lingual analysis, while maintaining the richness and authenticity of the original records.

3.4. Data Analysis

We performed an exhaustive analysis[8] of the dataset in order to find potential differences among records coming from the two different sources.

3.4.1. Measurements

As mentioned before, MeteoGalEus measurements can be divided into three categories: state-of-the-sky-related (making up 34.02% of the total

[6]Winter schedule uses morning 7:00 to 14:00, afternoon 14:00 to 20:00, and night 20:00 to 7:00, while summer schedule uses morning 6:00 to 14:00, afternoon 14:00 to 21:00, and night 21:00 to 6:00.

[7]Wind speed values: *calma* (≤ 5 km/h), *viento débil* (≤ 20 km/h), *viento moderado* (≤ 40 km/h), *viento fuerte* (≤ 70 km/h), and *viento muy fuerte* (> 70 km/h)

[8]Details are available at: https://gitlab.citius.gal/ainhoa.vivel/meteogaleus/-/tree/LREC/analysis

data), wind-related (44.71%), and temperature-related (21.27%). These proportions are not always consistent between MeteoGalicia and Euskalmet. For example, in MeteoGalicia the most common measurements are state-of-the-sky-related, making up 47.06% of the total, while both wind- and temperature-related measurements account for 26.47% each. However, records coming from Euskalmet exhibit a completely different behavior. In this case, the most common measurements are wind-related, making up 54.32% of the total, while state-of-the-sky-related measurements represent 27.16%, and temperature-related ones only 18.52%.

3.4.2. Missing Values

MeteoGalicia contains missing data and value inconsistencies, which are caused by two main reasons. On the one hand, some locations occasionally lack values for certain days. This was likely due to measurement stations being temporarily non-operational or failing to transmit data to the central server. On the other hand, not all measurement points are equipped with the same sensors, so some measurements (e.g., wind condition or temperature) are unknown for some locations.

Regarding wind conditions measurements, MeteoGalicia records missing values with -9999.0, which we decided to remove in the final JSON format. It is worth noting that in MeteoGalicia wind measurements are reported only in certain zones[9]. We also replaced the ambiguous and too general '*Fuera de Galicia*' into four zones, namely, *Bergantiños–Soneira*, *Ría de Muros e Noia*, *Ferrol Terra*, *Mariña Oriental*, and *Cíes*.

Regarding temperature, we observed a similar trend. The number of records per location vary depending on whether the measurement points were capable of recording that type of data. We aggregated measurements from different points, resulting in a single value per zone (in fact, two values – one for maximum and another for minimum temperature). However, some zones had no measurement points capable of recording temperature, which led to 27 locations with temperature data (while 6 locations are with missing values) for each forecast.

In summary, the absence of certain measurements in specific locations arises from the fact that not all MeteoGalicia stations are capable of recording all variables. Thus, there are no missing values in the traditional sense. Instead, some variables are simply not estimated for particular locations, depending on the measurement capabilities and instrumentation available at each station.

[9] *A Limia*, *Fuera de Galicia*, *Ordes–Arzúa–Melide*, *Lugo–Sarria*, and *Rías de Pontevedra e Vigo*

Phenomena Groups	MeteoGalEus	Meteo Galicia	Euskalmet
Clear	58.64%	46.24%	69.96%
Rain	24.44%	29.61%	19.72%
Clouds	14.98%	20.50%	9.95%
Storm	0.29%	0.33%	0.27%
Snow	0.63%	1.21%	0.10%
Fog	1.01%	2.11%	0.00%

Table 1: Distribution of grouped meteorological phenomena across the MeteoGalEus dataset.

In the case of Euskalmet, we identified that there are four cities[10] with missing values for some days. All in all, they represent 9.6% of the data. We also found that only 8 cities have all of the complete forecasted values of temperature, wind, and state of the sky in the whole time range. Moreover, there is no more than 3.96% of missing variables per city and variable, except for Vitoria-Gasteiz (8.56%) where measurements are only available since 12 April 2021. Finally, in the Euskalmet dataset, some of the minimum and maximum temperature values are identical, which does not accurately reflect the actual meteorological conditions. This issue stems from missing temperature data in the original source, which was imputed by Euskalmet. These values were provided as such and were incorporated into MeteoGalEus without further modification.

3.4.3. State of the sky

Cloud coverage, precipitation, and, in general, state-of-the-sky measurements are heavily determined by geographical location. This consequently means that differences can certainly be found in the different portions of the data.

By analyzing the main state-of-the-sky phenomena across the entire MeteoGalEus dataset, we find that the most common phenomena are *Despejado* – Clear (27.99%); *Poco Nuboso* – Mostly Clear (16.86%); *Nubes y claros* – Clouds and Clear Intervals (8.95%); *Chubascos Débiles* – Light Showers (7.55%); and *Nublado* – Cloudy (7.47%).

In the case of MeteoGalicia, the most common state-of-the-sky phenomena are *Nubes y Claros* – Clouds and Clear Intervals (18.75%); *Despejado* – Clear (17.32%); and *Nublado* – Cloudy (15.66%). On the other hand, in the case of Euskalmet, the most common phenomena are *Despejado* – Clear (37.73%); *Poco Nuboso* – Mostly Clear (32.24%); *Lluvia Débil* – Light Rain (10.66%); and *Muy Nuboso* – Very Cloudy (9.95%).

[10] Sestao, Getxo, Bilbao, and Vitoria-Gasteiz

Year	MeteoGalEus	MeteoGalicia	Euskalmet
2020	8.50	8.17	8.75
2021	12.69	13.00	12.45
2022	13.68	13.75	13.62
2023	13.94	13.96	13.94
2024	13.56	13.66	13.48
2025	9.00	9.16	8.87

Table 2: Annual mean temperature (in °C) across the MeteoGalEus dataset and its regional subsets.

It is also worth taking into account that, due to the different origins of the data, we find distinct state-of-the-sky labels depending on the region. To perform a more balanced analysis, we aggregated several of these labels under broader categories. After aggregating the values, we recomputed the distribution analysis that is summarized in Table 1.

3.4.4. Wind conditions

Regarding the wind-related measurements, we analyze wind-speed-related data and wind-direction-related data. Notice that, wind direction is conditioned by factors such as geographical location and position with respect to the sea.

The most common wind-direction measurements in the MeteoGalEus dataset are *Noroeste* – Northwest (18.88%); *Suroeste* – Southwest (16.75%); *Oeste* – West (15.58%); *Sur* – South (14.30%); and *Noreste* – Northeast (11.98%). Regarding MeteoGalicia, the most common wind direction measurements are *Suroeste* – Southwest (18.70%); *Noroeste* – Northwest (17.15%); *Variable*, which is a wildcard category used when no clear direction is detected (16.51%); *Sur* – South (11.12%); and *Norte* – North (10.00%). On the other hand, the most common wind-direction-related measurements from Euskalmet are *Noroeste* – Northwest (21.60%); *Oeste* – West (17.43%); *Suroeste* – Southwest (16.25%); *Sur* – South (15.11%); and *Noreste* – Northeast (10.66%).

Moving on to the wind-speed-related measurements, we find that for MeteoGalEus, 55.35% of the total data corresponds to *Viento Débil* (Weak Wind), 19.63% to *Calma* (Calm), and 18.52% to *Viento Moderado* (Moderate Wind). This shows how weak wind speeds are observed for the majority of measurements. In the case of the MeteoGalicia, the most common measurements are *Viento Débil*, Weak Wind (39.12%); *Viento Moderado*, Moderate Wind (31.06%); and *Variable* (16.51%). Surprisingly, only 1.13% of the total refers to the *Calma* (Calm) category. On the other hand, in the case of Euskalmet, we find a smaller number of categories (No Variable Wind nor Very Strong wind classes are ever used): 59.52% of the total data corresponds to *Viento Débil* (Weak Wind), 24.38% to *Calma* (Calm), 15.30% to *Viento Moderado* (Moderate Wind), and the remaining 0.81% corresponds to *Viento Fuerte* (Strong Wind).

Autonomous Community	Province	Mean Temperature
Galicia	A Coruña	14.48
	Lugo	12.11
	Ourense	13.15
	Pontevedra	15.73
Euskadi	Araba/Álava	11.73
	Bizkaia	14.06
	Gipuzkoa	13.50

Table 3: Mean temperature (in °C) by province across the MeteoGalEus dataset.

3.4.5. Temperature

We pay attention to mean temperature for each year. Something worth taking into account is that for 2020 and 2025, the mean temperature is much lower than in the rest of the years as shown in Table 2. This is because, for these two particular years, we have incomplete data from only the last months and the first months of the year, respectively. Accordingly, the available data from these years correspond to the cold season only.

Interestingly, the mean temperature depends on the geographical area as shown in Table 3. We can see some variation not only between autonomous communities but also within each territory (with the lowest temperatures in Euskadi).

3.4.6. Texts

To examine MeteoGalEus textual data, an analysis on text references was performed, where state-of-the-sky, wind, temperature, geographical-space, and time-related words were counted using a thesaurus (see Table 4).

4. Experimental Settings

The evaluation protocol for MeteoGalEus measures both text generation quality and multilingual consistency across Spanish, Galician, and Euskera. The dataset is split into non-overlapping training and test sets with aligned observations. Baselines are established using LLMs in zero-shot and fine-tuning settings. Models are evaluated per language and jointly across languages using standard NLP metrics that assess similarity, semantic adequacy,

Reference Type	MeteoGalEus	Meteo Galicia	Euskalmet
State of the sky	2.74	3.21	2.27
Wind	2.11	2.83	1.39
Temperature	3.36	3.86	2.86
Space	5.36	6.34	4.37
Time	1.40	1.43	1.36

Table 4: Average number of references per forecast across the MeteoGalEus dataset.

and fluency. This setup ensures transparent, reproducible, and cross-lingual evaluation.

To establish reference performance on MeteoGalEus, we adopt the LLaMA-3.1-8B-Instruct model (Touvron et al., 2023) as a baseline system. We consider two transfer learning techniques to explore different levels of supervision and adaptation. Namely, we explored zero-shot and fine-tuned approaches to evaluate the model's ability to generate weather forecasts in multiple languages. These two approaches allow us to benchmark model performance across different learning regimes, and to compare how well LLMs can handle forecast generation and multilingual text production without extensive domain-specific engineering.

The dataset was split into training and test sets in such a way that, for any given date, observations from both Galicia and the Basque Country were assigned to the same split, ensuring consistency across regions within each partition. A randomized 70/30 split at the date level was then applied, with all records from a given day assigned consistently to either training or test. For validation during training, an 80/20 split was established.

The training set (MeteoGalEus_train.json) spans from 25 November 2020 to 27 February 2025, covering 1079 distinct dates with a total of 2158 entries over a time span of 1555 days (approx 4.26 years). The test set (MeteoGalEus_test.json) spans from 29 November 2020 to 1 March 2025, comprising 463 distinct dates with a total of 926 entries over a time span of 1553 days (approx 4.25 years). This procedure guarantees that no temporal leakage occurs between the splits while preserving temporal diversity in both training and test datasets.

We designed experiments to evaluate both regional and linguistic generalization, considering within-region and cross-region performance and cross-lingual evaluation:

- **In-domain Evaluation.** The model is tested on the same geographical region as the training data is associated with. This measures the model's ability to reproduce familiar patterns and regional language usage.
- **Cross-domain Evaluation.** The model is tested on a different geographical region than it was trained on. This assesses the model's cross-regional transfer capabilities.

These experimental settings allow us to systematically evaluate cross-region transfer, regional generalization, and the effectiveness of learning in a multilingual weather forecasting scenario.

Regarding computational resources, the transfer learning experiments using MeteoGalEus and Euskalmet data were conducted with 2 GPUs type A100 (80GB), while only one of the two GPUs was needed for the text generation task. For the same experiments associated to MeteoGalEus and MeteoGalicia we used 2 GPUs type H100 (80GB).

4.1. Transfer Learning

We considered two transfer learning strategies for exploring how LLMs can exploit the MeteoGalEus dataset:

- **Zero-Shot.** The pre-trained model generates forecasts directly from task instructions, without any exposure to MeteoGalEus examples.
- **Fine-Tuning.** The model is fully trained on the MeteoGalEus training split, updating its parameters to produce specialized forecasts.

This setup allows us to compare the model's ability to generate multilingual weather forecasts under no adaptation (zero-shot) and full adaptation (fine-tuning). These strategies enable the evaluation of the balance between generalization and task-specific adaptation, highlighting the benefits of lightweight prompting versus full fine-tuning in a multilingual, structured-data-to-text setting. Moreover, they provide a reference point for assessing future models on MeteoGalEus.

4.2. Baseline Models

In this study, we evaluate several baseline configurations for LLaMA-3.1-8B-Instruct[11] (Touvron et al., 2023) on the MeteoGalEus dataset, covering both monolingual and multilingual training scenarios. Table 5 summarizes the different setups and their corresponding train and test splits.

The monolingual baselines trained exclusively in Spanish provide a reference point for language-specific performance across three evaluation scenarios: training and testing on the same region, training and testing on different regions, and training and testing on the entire dataset. For the monolingual baselines in Galician or Euskera, both training and evaluation are conducted within the same

[11] https://huggingface.co/meta-llama/Llama-3.1-8B-Instruct

ID	Language	Train Split	Test Split
0	ES	MeteoGalEus	MeteoGalEus
1		Euskalmet	MeteoGalicia
2		Euskalmet	Euskalmet
3		MeteoGalicia	Euskalmet
4		MeteoGalicia	MeteoGalicia
5	GL	MeteoGalicia	MeteoGalicia
6	GL+ES	MeteoGalicia	MeteoGalicia
7	EU	Euskalmet	Euskalmet
8	EU+ES	Euskalmet	Euskalmet

Table 5: Baseline configurations for LLaMA-3.1-8B-Instruct on MeteoGalEus. Each baseline is evaluated under zero-shot and fine-tuning setups.

region. For the multilingual baselines, we consider two language combinations: GL+ES and EU+ES. In these cases, training is performed in both languages, but evaluation is carried out only in the co-official language of the region. Each multilingual model is trained and evaluated within the same region, either Galicia or the Basque Country. Overall, these baselines serve as a foundation for comparing zero-shot and fine-tuned performance across languages, providing insights into both monolingual and multilingual capabilities of LLaMA-3.1-8B-Instruct on meteorological data.

Based on the train/test combinations outlined in Table 5, we performed a total of 7 model trainings to cover all fine-tuning scenarios across the different languages and regions. Specifically, we considered 9 fine-tuning evaluations (one per model/train split combination) and 5 zero-shot evaluations, where the pre-trained LLaMA model is evaluated directly on the test splits without additional training.

4.3. Evaluation

We evaluated the quality of generated weather descriptions using standard metrics from NLP. Since MeteoGalEus provides multilingual textual forecasts, we paid attention to both lexical and semantic similarity between generated and reference texts. We used the following metrics in the experiments:

- BLEU (Papineni et al., 2002) measures n-gram overlap between generated and reference texts, providing a surface-level assessment of similarity.
- ROUGE (Lin, 2004) computes recall-oriented overlap of n-grams and the longest common subsequences, capturing content coverage and the preservation of important information. Namely, we used ROUGE-Lsum as it measures the longest common subsequence at the summary level, rewarding sentence-level structure, coherence, and content coverage, which is important for multi-sentence forecasts.
- BERTScore (Zhang et al., 2020) uses contextual embeddings to measure semantic similarity between generated and reference texts, thus capturing meaning beyond exact word matches. We decided to report BERTScore F1 mean because it evaluates semantic similarity in terms of contextual embeddings, allowing us to capture meaning equivalence even when the model paraphrases or uses alternative phrasing.

All metrics were computed using the `evaluate` library from Hugging Face (Von Werra et al., 2022), which provides standardized implementations for these metrics across multiple languages. For clarity and brevity, we report only these metrics in the next section. However, a complete set of metrics is provided in the accompanying repository for reproducibility and further analysis.

5. Results

Table 6 summarizes the zero-shot results of our LLaMA-based models on the MeteoGalEus dataset, while Table 7 presents the fine-tuning results.

In the zero-shot setting, BERTScore ranges from 0.603 to 0.694, indicating moderate semantic fidelity, while ROUGE scores remain low (0.047-0.200) and BLEU values are very small (0.004–0.016), reflecting the difficulty of achieving exact lexical overlap without task-specific training. Among these, models 5 and 6 achieve the best overall performance, with a BERTScore of 0.694, ROUGE of 0.189, and BLEU of 0.016, representing the strongest zero-shot configuration. Notably, models 7 and 8, which use only Euskalmet data, exhibit the lowest scores, suggesting that dealing with examples in Euskera may be more challenging. It is worth noting that the results reported in Table 6 are the same for those models that were tested with the same test split (i.e., 1 and 4, 2 and 3, 5 and 6, 7 and 8, as seen in Table 5) because the zero-shot setting does not care of the given training split.

Fine-tuning leads to an improvement across all metrics, with model 4 obtaining the highest performance overall: BERTScore reaching up to 0.788, ROUGE up to 0.345, and BLEU up to 0.167. This demonstrates that task-specific adaptation significantly enhances semantic alignment and surface-level similarity with reference forecasts. Despite these gains, models 7 and 8 remain with relatively low performance, indicating the complexity of

ID	BERTScore	ROUGE	BLEU
MeteoGalEus			
0	0.682	0.195	0.013
MeteoGalicia			
1	0.683	0.200	0.015
4	0.683	0.200	0.015
5	**0.694**	**0.189**	**0.016**
6	**0.694**	**0.189**	**0.016**
EuskalMet			
2	**0.677**	**0.188**	**0.010**
3	**0.677**	**0.188**	**0.010**
7	0.603	0.047	0.004
8	0.603	0.047	0.004

Table 6: Results for LLaMA model in the zero-shot setting, grouped according to the test set. Best models results bold.

ID	BERTScore	ROUGE	BLEU
MeteoGalEus			
0	0.769	0.298	0.133
MeteoGalicia			
1	0.721	0.179	0.029
4	**0.788**	**0.345**	**0.167**
5	0.752	0.199	0.031
6	0.750	0.180	0.024
EuskalMet			
2	**0.742**	**0.233**	**0.074**
3	0.717	0.174	0.036
7	0.626	0.002	0.000
8	0.628	0.002	0.000

Table 7: Results for LLaMA model in the fine-tune setting, grouped according to the test set. Best models results bold.

generating weather forecasts in Euskera and the model's limited generalization to such cases.

Overall, these results highlight that fine-tuning substantially improves the text quality, particularly in terms of semantic fidelity as captured by BERTScore. The combination of semantic and surface-level metrics ensures a balanced and comprehensive evaluation of the models' performance.

5.1. Qualitative Comparison of Generated and Expert Meteorological Texts

To complement the quantitative evaluation with automatic metrics, we present a qualitative analysis comparing LLM-generated meteorological reports with reference forecasts from expert meteorologists. This assessment highlights factual accuracy and the presence of hallucination, statements not supported by the input data.

We present a representative example from Galicia, comparing a reference forecast with one generated by the LLM, both manually translated into English by native Galician speakers. The original texts and additional examples are provided in Appendix B, , and can also be found in the JSONL files located at `meteogaleus/scripts/modelling/finetune_llama/output/generated/`. Using the provided ID, the corresponding input data can be accessed in the MeteoGalEus dataset.

The following example (`ID: 4640`) illustrates a weather forecast generated by Model 5, which was fine-tuned exclusively on Galician data. The forecast corresponds to the autonomous community of Galicia on December 16, 2024. The manually translated text in English is:

> *Today Tuesday Galicia will remain under the influence of high pressure, with winds from the south. Thus, we will start with morning fog banks in interior areas and isolated fog banks, which will give way to partly cloudy or clear skies. Minimum temperatures have experienced a slight decrease, while maximum temperatures will rise. The wind will blow lightly from the south, with moderate intervals along the coast and in higher areas.*

and the reference text:

> *Today Galicia remains under anticyclonic influence. This way, we will have clear skies, but with the entry of mid- and high-level clouds that will not block the sun. Temperatures will not change significantly. The wind will blow lightly from the south, with moderate to strong intervals offshore toward the end of the day between Louro and Fisterra.*

The monolingual, region-specific model effectively captures the overall meteorological context for Galicia, correctly describing state of the sky and winds. However, it occasionally introduces hallucinations, such as wrongly specifying the day of the week, which is not present in the source data. Minor phenomena, like morning fog, are sometimes overemphasized, while more significant details, such as cloud type and offshore wind intensity, may be simplified. The generated text remains fluent, coherent, and regionally appropriate. Native Galician speakers did not identify any linguistic errors. Overall, the model is suitable for producing readable, regionally realistic forecasts in Galician. Post-editing or verification against official data is essential to prevent hallucinations or inaccuracies in the predictions.

We present the forecast generated by the zero-shot version of Model 5-6 for the same entry.

> *Forecast for Galicia for today:\n\nToday, the region of Galicia will be characterized*

by stable weather, with temperatures ranging between 10ºC and 17ºC, with the highest maximum temperatures in coastal areas and the lowest minimum temperatures in inland areas.\n\nThe state of the sky will be predominantly cloudy, with high clouds throughout the day, although there will be some periods of clouds and clear skies, especially in the afternoon and at night.\n\nThe wind will be light to moderate, from the southeast and south, becoming stronger at night.\n\nIn summary, a summer-like day with pleasant temperatures, but with cloudy weather and a light wind.

The zero-shot output differs substantially from both the expert reference and the fine-tuned model generation. While the model correctly captures some general patterns from the input data (such as the presence of southerly winds, cloud coverage, and temperatures roughly around coastal values) it introduces several inaccuracies.

The model overestimates cloudiness by describing the sky as "predominantly cloudy", whereas the dataset mainly reports high clouds and several clear or partly cloudy periods. The expert forecast reflects this more accurately, describing clear skies with mid- and high-level clouds that do not block sunlight. In addition, the generated temperature range (10 - 17ºC) ignores colder inland conditions, where minimum temperatures in the dataset reach -4ºC. The model produces a simplified summary that does not fully represent regional variability.

The model also introduces unsupported interpretations (e.g., describing the day as "summer-like", despite the forecast corresponding to December and including sub-zero temperatures). It also suggests that winds become stronger at night, although the input data shows mixed wind patterns without a consistent nighttime increase. Additionally, the generated text contains lexical interference from Spanish, using "viento" instead of the Galician "vento", "sureste" instead of "sueste" and "interioranas" instead of "do interior". It also incorrectly uses "hoy" instead of "hoxe", although the correct form appears in the following sentence. The example contains several stylistic issues. The generated text starts with a template-like forecast header ("Forecast for Galicia for today:"), a structure typical of LLM-generated summaries but not used in the reference forecasts. It also introduces line breaks ("\n"). Although this specific output does not include them, other zero-shot generations in our experiments frequently contain additional formatting artifacts, such as bold markers ("**") or bullet lists ("-") that are never used in this type of text. These elements reflect the model's general formatting tendencies but differ from the natural narrative style typically used in expert-written forecasts. Overall, this example illustrates some limitations of zero-shot generation: although the model verbalizes structured meteorological data, it struggles with content prioritization, stylistic conventions, and factual consistency. The model succeeds in conveying structured information but it has limitations in replicating human-like text conventions. Accordingly, fine-tuning is important to produce natural, fluent, and regionally accurate forecasts. Comparing fine-tuned and zero-shot outputs shows the improvements in style, coherence, and adherence to human conventions achieved through fine-tuning.

6. Conclusion

In this work, we introduced MeteoGalEus, a multilingual dataset of weather forecasts covering Spanish, Galician, and Euskera, designed to support research in multilingual NLP, with special attention to data-to-text NLG modeling. The dataset provides aligned, high-quality forecasts across multiple languages, enabling fair and consistent evaluation of monolingual and multilingual pre-trained models.

To demonstrate its usefulness, we established baseline results using LLaMA models in both zero-shot and fine-tuned settings. Our experiments show that fine-tuning substantially improves semantic fidelity and surface-level similarity to human-written forecasts, highlighting the dataset's value for training and evaluating modern language models.

Fine-tuning led to an average improvement of BERTScore 0.055 ± 0.028, ROUGE 0.019 ± 0.066, and BLEU 0.043 ± 0.057, confirming consistent performance gains across evaluation metrics. Qualitative analysis further shows that fine-tuned models can generate fluent and regionally appropriate forecasts, although occasional hallucinations and simplifications of meteorological details still occur. Despite these gains, certain examples, particularly in Euskera, remain challenging, suggesting that MeteoGalEus can also serve as a benchmark for testing model robustness and generalization in scenarios associated to low-resource languages.

We expect MeteoGalEus to facilitate further research in multilingual NLG, model evaluation, and cross-lingual understanding, and we encourage the community to explore additional modeling approaches, data augmentation strategies, and evaluation metrics. Beyond the presented baselines, the dataset opens avenues for exploring cross-lingual adaptation and multimodal forecasting.

7. License

MeteoGalEus is licensed under the terms and conditions of the Creative Commons Attribution-

ShareAlike 4.0 International (CC BY-SA 4.0)[12].

8. Acknowledgements

Thanks to the Euskalmet and MeteoGalicia Agencies for providing data under a Creative Commons license. Ainhoa Vivel Couso is a PhD Researcher supported by the Galician Ministry of Education, Science, Universities and Professional Training (ED481A-2024-059). Nella Zabrina Pramata was supported financially by the Erasmus Mundus Masters Programme in Language and Communication Technologies, EU grant no. 2019-1508. All authors recognize the support of the Galician Ministry of Education, Science, Universities and Professional Training (CiTIUS 2024-2027 ED431G2023/04, GSI 2022-2025 ED431C2022/19) and the European Union (European Regional Development Fund - ERDF). This work is also supported by the Spanish Ministry of Science, Innovation and Universities (MCIN/AEI/10.13039/501100011033/) through MAIXAI4TRUST PID2024-157680NB-I00, DeepThought PID2024-159202OB-C21 and humanAIze AIA2025-163322-C61.

9. Bibliographical References

J. Coch. 1998. System demonstration interactive generation and knowledge administration in MultiMeteo. In *Natural Language Generation*, Niagara-on-the-Lake, Ontario, Canada. Association for Computational Linguistics.

O. Dušek, J. Novikova, and V. Rieser. 2020. Evaluating the state-of-the-art of end-to-end natural language generation: The e2e nlg challenge. *Computer Speech & Language*, 59:123–156.

A. Gatt and E. Krahmer. 2018. Survey of the State of the Art in Natural Language Generation: Core Tasks, Applications and Evaluation. *Journal of Artificial Intelligence Research*, 61:65–179.

J. González-Corbelle, J.M. Alonso-Moral, A. Bugarín Diz, and J. Taboada. 2022. Dealing with hallucination and omission in neural Natural Language Generation: A use case on meteorology. In *Proceedings of the 15th International Conference on Natural Language Generation*, pages 121–130, Waterville, Maine, USA and virtual meeting. Association for Computational Linguistics.

Trey Gowdy. 2024. Data-driven climate solutions. Nicholas Institute for Energy, Environment and Sustainability: Annual Report (Highlights from FY 2024).

Y. Hu, H. Liao, M. Wu, and L. Yuan. 2025. SST-LLM: time series forecasting based on large language models. In *International Symposium on Artificial Intelligence Innovations (IS-AII)*, volume 13681, pages 198–206. SPIE.

H. Li, Z. Wang, J. Wang, A.K.H. Lau, and H. Qu. 2024. CLLMate: A Multimodal LLM for Weather and Climate Events Forecasting. *arXiv preprint arXiv:2409.19058*.

S. Li, W. Yang, P. Zhang, X. Xiao, D. Cao, Y. Qin, X. Zhang, Y. Zhao, and P.l Bogdan. 2025. ClimateLLM: Efficient Weather Forecasting via Frequency-Aware Large Language Models. *arXiv preprint arXiv:2502.11059*.

P. Liang, M. Jordan, and D. Klein. 2009. Learning semantic correspondences with less supervision. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 91–99, Suntec, Singapore. Association for Computational Linguistics.

C.-Y. Lin. 2004. ROUGE: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.

K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu. 2002. BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318. ACL.

A. Ramos-Soto, A. Bugarin, S. Barro, and J. Taboada. 2015. Linguistic Descriptions for Automatic Generation of Textual Short-Term Weather Forecasts on Real Prediction Data. *IEEE Transactions on Fuzzy Systems*, 23:44–57.

E. Reiter. 2024. *Natural Language Generation*. Springer.

E. Reiter and R. Dale. 2000. *Building Natural-Language Generation Systems*. Cambridge University Press.

S. Gowri Sripada, E. Reiter, J. Hunter, and J. Yu. 2003. Exploiting a parallel TEXT - DATA corpus.

H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, and G. Lample. 2023. LLaMA: Open and Efficient Foundation Language Models.

[12] https://creativecommons.org/licenses/by-sa/4.0

L. Von Werra, L. Tunstall, A. Thakur, S. Luccioni, T. Thrush, A. Piktus, F. Marty, N. Rajani, V. Mustar, and H. Ngo. 2022. Evaluate & evaluation on the hub: Better best practices for data and model measurements. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 128–136, Abu Dhabi, UAE. Association for Computational Linguistics.

T. Zhang, V. Kishore, F. Wu, K.Q. Weinberger, and Y. Artzi. 2020. BERTScore: Evaluating Text Generation with BERT.

A. Examples of MeteoGalEus Data

This appendix provides two representative examples of the MeteoGalEus dataset, illustrating meteorological predictions estimated on November 25, 2020 in stations located in Galicia and the Basque Country. The purpose of this examples is to offer concrete reference for the structure, content, and format of the data used throughout this work.

For clarity and brevity, the `meteorological_data` field shows only three cities for each region. `![...]` is used to indicate where data have been omitted.

```
[
  {
    "id": "euskalmet_2020-11-25",
    "from": "Euskalmet",
    "autonomous_community": "Basque
    ↪  Country",
    "date": "2020-11-25",
    "forecast": {
      "ES": "Durante lo que resta de día
      ↪  tendremos el cielo poco nuboso
      ↪  en todas partes salvo en el
      ↪  Valle del Ebro y comarcas
      ↪  cercanas, donde se espesará de
      ↪  nuevo la niebla. El viento del
      ↪  sureste dejará algunas rachas
      ↪  muy fuertes en zonas expuestas
      ↪  y hará que las temperaturas
      ↪  nocturnas vuelvan a subir,
      ↪  sobre todo en el interior.",
      "EU": "Gainerako orduetan zerua
      ↪  hodei gutxirekin egongo da
      ↪  denean, Ebro bailaran izan
      ↪  ezik; han behe-hodeiak eta
      ↪  lainoa izango dira nagusi.
      ↪  Hego-ekialdeko haizea bizi
      ↪  ibiliko da, bolada oso
      ↪  gogorrekin haizeguneetan.
      ↪  Tenperatura minimoek gora
      ↪  egingo dute berriro, batez ere
      ↪  barnealdean, kostaldeko
      ↪  baxuenek berdintsu jarraituko
      ↪  dute ziurrenik."
    },
    "meteorological_data": [
      {
        "zone": "Bizkaia",
        "location": "Bilbao",
        "temperature": {
          "max": 11.6,
          "min": 11.6
        },
        "wind": [
          {
            "period": "mañana",
            "direction": "sur",
            "speed": "calma"
          },
          {
            "period": "tarde",
            "direction": "sur",
            "speed": "calma"
          },
          {
            "period": "noche",
            "direction": "sur",
            "speed": "calma"
          }
        ],
        "state_of_the_sky": {
          "morning": "despejado",
          "afternoon": "despejado",
          "night": "despejado"
        }
      },
      {
        "zone": "Araba/Álava",
        "location": "Amurrio",
        "temperature": {
          "max": 8.2,
          "min": 8.2
        },
        "wind": [
          {
            "period": "mañana",
            "direction": "sur",
            "speed": "calma"
          },
          {
            "period": "tarde",
            "direction": "sur",
            "speed": "calma"
          },
          {
            "period": "noche",
```

```
            "direction": "sur",
            "speed": "calma"
          }
        ],
        "state_of_the_sky": {
          "morning": "despejado",
          "afternoon": "despejado",
          "night": "despejado"
        }
      },
      {
        "zone": "Gipuzkoa",
        "location": "Donostia/San
        ↪  Sebastian",
        "temperature": {
          "max": 12.6,
          "min": 12.6
        },
        "wind": [
          {
            "period": "mañana",
            "direction": "sur",
            "speed": "calma"
          },
          {
            "period": "tarde",
            "direction": "sur",
            "speed": "calma"
          },
          {
            "period": "noche",
            "direction": "sur",
            "speed": "calma"
          }
        ],
        "state_of_the_sky": {
          "morning": "despejado",
          "afternoon": "despejado",
          "night": "despejado"
        }
      },
      ![...]
    ]
  },
  {
    "id": 3158,
    "from": "MeteoGalicia",
    "autonomous_community": "Galicia",
    "date": "2020-11-25",
    "forecast": {
      "ES": "Hoy en Galicia será un día
      ↪  de aire frío en capas altas de
      ↪  la atmósfera. De este modo,
      ↪  los cielos estarán
      ↪  parcialmente cubiertos con
      ↪  precipitaciones más frecuentes
      ↪  en el este de Lugo y Ourense .
      ↪  Las temperaturas mínimas
      ↪  experimentaron un moderado
      ↪  ascenso, mientras que las
      ↪  máximas sufrirán un ligero
      ↪  descenso. El viento soplará de
      ↪  componente norte, con rachas
      ↪  fuertes en el litoral norte.",
      "GL": "Hoxe en Galicia será un día
      ↪  de aire frío en capas altas da
      ↪  atmosfera. Deste xeito, os
      ↪  ceos estarán parcialmente
      ↪  cubertos con precipitacións
      ↪  máis frecuentes no leste de
      ↪  Lugo e Ourense . As
      ↪  temperaturas mínimas
      ↪  experimentaron un moderado
      ↪  ascenso, mentres que as
      ↪  máximas sufrirán un lixeiro
      ↪  descenso. O vento soprará de
      ↪  compoñente norte, con refachos
      ↪  fortes no litoral norte."
    },
    "meteorological_data": [
      ![...]
      {
        "zone": "Lugo",
        "location": "A Fonsagrada",
        "state_of_the_sky": {
          "morning": "nublado con
          ↪  chubasco",
          "night": "nublado con
          ↪  chubasco",
          "afternoon": "nublado con
          ↪  chubasco"
        }
      },
      {
        "zone": "Ourense",
        "location": "A Limia",
        "temperature": {
          "max": 7.0,
          "min": 7.0
        },
        "wind": [
          {
            "period": "mañana",
            "direction": "variable",
            "speed": "variable"
          },
          {
            "period": "noche",
            "direction": "variable",
            "speed": "variable"
          },
          {
            "period": "tarde",
            "direction": "variable",
            "speed": "variable"
          }
        ],
        "state_of_the_sky": {
          "morning": "nublado con
          ↪  chubasco",
          "night": "nublado",
          "afternoon": "nublado con
          ↪  chubasco"
        }
      },
      {
        "zone": "A Coruña",
        "location": "Bergantiños -
        ↪  Soneira",
```

```
        "temperature": {
          "max": 9.666666666666666,
          "min": 7.0
        },
        "wind": [
          {
            "period": "mañana",
            "direction": "norte",
            "speed": "viento moderado"
          },
          {
            "period": "noche",
            "direction": "noreste",
            "speed": "viento moderado"
          },
          {
            "period": "tarde",
            "direction": "noreste",
            "speed": "viento fuerte"
          }
        ],
        "state_of_the_sky": {
          "morning": "chubascos
          ↪  débiles",
          "night": "nublado con
          ↪  chubasco",
          "afternoon": "nublado con
          ↪  chubasco"
        }
      },
      ![...]
    ]
  }
]
```


B. Model Outputs and Reference Texts in Source Languages

This appendix provides examples of model-generated texts together with reference texts authored by meteorological experts. All texts are presented in their original languages: Galician, Basque, and Spanish.

B.1. Galician

These examples illustrate the text generated by (1) the zero-shot Model 5-6 and (2) the fine-tuned Model 5 for the entry with ID `4640` in the MeteoGalEus dataset. The fine-tuned model was trained exclusively on Galician data from MeteoGalicia. The reference text was:

> *Hoxe Galicia continúa dentro da influencia anticiclónica. Deste xeito, teremos o ceo despexado pero con entrada de nubes de tipo medio e alto que non impedirán o paso do sol. As temperaturas non terán cambios significativos. O vento soprará frouxo de compoñente sur con intervalos de moderado a fortes en mar aberto ao final do día entre Louro e Fisterra.*

B.1.1. Zero-shot Model 5-6

The generated text was:

> *Pronóstico para Galicia para hoy:\n\nO día de hoxe, a rexión de Galicia estará caracterizada por un clima estabilizado, con temperaturas que oscilarán entre os 10°C e os 17°C, sendo as temperaturas máximas máis altas nas zonas costeiras e as mínimas máis baixas nas zonas interioranas.\n\nO estado do ceo será predominantemente nubrado, con nubes altas durante todo o día, aínda que haxa algúns períodos de nubes e claros, especialmente na tarde e na noite.\n\nO viento será débil a moderado, con dirección sureste e sur, sendo máis forte na noite.\n\nEn resumo, un día de verán con temperaturas agradables, pero con un clima nubrado e un viento suave.*

B.1.2. Fine-tuned Model 5

The generated text was:

> *Hoxe martes Galicia continuará na influencia das altas presións, con ventos de compoñente sur. Así, comezaremos con bancos de néboa matinais en zonas de interior e bancos de néboa illados que darán paso a ceos pouco nubrados ou despexados. As temperaturas mínimas sufriron un lixeiro descenso, mentres que as máximas ascenderán. O vento soprará frouxo de compoñente sur, con intervalos moderados no litoral e zonas altas.*

B.2. Euskera

These examples illustrate the text generated by (1) the zero-shot Model 7-8 and (2) the fine-tuned Model 7 for the entry with ID `euskalmet_2024-06-24` in the MeteoGalEus dataset. The fine-tuned model was trained exclusively on Euskera data from Euskalmet. The reference text was:

> *Giro eguzkitsua eta lasaia arratsaldean izan duguna; ez da beroegia izan, iparraldeko haizeari esker. Ilunabarretik aurrera haizea ahul eta aldakor ibiliko da eta orduak pasa ahala barnealdeko zenbait tokitan lainoguneak sor daitezke. Zerua hodei gutxirekin egongo da, behehodeien hondarrekin mendi inguruetan. Gaueko tenperatura minimoak jaits daitezke.*

B.2.1. Zero-shot Model 7-8

The generated text was:

> Pronóstico para el País Vasco:\n\nHoy, el tiempo en la región estará marcado por temperaturas suaves y vientos débiles. La temperatura máxima esperada es de alrededor de 22-23°C en la mayoría de las localidades, con máximas más altas en la zona occidental (Bastida, Moreda, Oion y Lapuebla) y máximas más bajas en la costa (Zumaia y Zarautz). La temperatura mínima será de alrededor de 12-14°C.\n\nEl viento será débil en la mayoría de las localidades, con direcciones noreste y este, excepto en algunas zonas donde se esperan calmas. La dirección del viento será noreste en la mañana y tarde, y se desplazará al este en la noche.\n\nEl cielo estará despejado en la mañana y tarde, con algunas nubes dispersas, pero se nublará en la noche en la mayoría de las localidades. La región se mantendrá seca, sin lluvia esperada.\n\nEn resumen, un día soleado y cálido en el País Vasco, con temperaturas suaves y vientos débiles.

B.2.2. Fine-tuned Model 7

The generated text was:

> *Gaur, osteguna, eguraldiak ildo beretik jarraituko du. Goizean goiz hodeiak ugarituko dira, batez ere Kantauri isurialdean; balio altuenak 20 °C ingurukoak izango dira. Hegoaldean, aldiz, egunak aurrera egin ahala berdintsu jarraituko dute. Hegoaldean, aldiz, egunak aurrera egin ahala berdintsu jarraituko dute.*

B.3. Spanish

For Spanish, we want to showcase two scenarios. First, using the entire MeteoGalEus dataset with examples from both Galicia and the Basque Country to illustrate the model's full capabilities. Second, evaluating cross-region performance by training in one region and testing in the other. In both cases, examples are in Spanish. These examples illustrate the text generated by (1) the zero-shot Model 0 for the entry with ID `3989` and (2) the fine-tune Model 3 for the entry with ID `euskalmet_2024-06-18` in the MeteoGalEus dataset.

B.3.1. Zero-shot Model 0

The generated text was:

> *Para la región de Galicia, se espera un día con temperaturas moderadas, con máximas entre 10°C y 17°C y mínimas entre 0°C y 10°C. El cielo estará cubierto de nubes y chubascos débiles en la mayoría de las localidades, con una tendencia a la nubosidad y chubascos más intensos por la noche.\n\n El viento será débil en la mayoría de las localidades, con direcciones variables y velocidades que no superarán los 30 km/h. Sin embargo, en algunas zonas, como la Ría de Muros e Noia y las Rías de Pontevedra e Vigo, se esperan vientos moderados por la tarde.\n\n **Resumen del pronóstico:**\n\n- Temperatura: máximas entre 10°C y 17°C, mínimas entre 0°C y 10°C.\n - Cielo: nubes y chubascos débiles en la mayoría de las localidades, con nubosidad y chubascos más intensos por la noche.\n - Viento: débil en la mayoría de las localidades, con direcciones variables y velocidades que no superarán los 30 km/h.*

And the reference text was:

> *Hoy, la llegada a Galicia de una nueva frente poco activa por el sudoeste traerá una día de cielos nubosos y con lluvias débiles, más frecuentes en la mitad sur y más aislados y ocasionales al norte. Las temperaturas en ascenso, que fue moderado en las mínimas, será ligero en las máximas, más acusado en la mitad sur. El viento soplará de componente sur, flojo en el interior y moderado en la costa.*

B.3.2. Fine-tuned Model 3

The generated text was:

> *Hoy, jueves, se prevé la entrada de las altas presiones que nos traerán una nueva jornada de tiempo seco y de cielos con pocas nubes por la mañana, creciendo nubes de evolución por la tarde, sin descartar algún chubasco aislado más probable en la provincia de Ourense. Las temperaturas mínimas sufrieron cambios significativos, mientras que las máximas experimentarán un ligero ascenso. El viento soplará flojo de componente oeste.*

And the reference text was:

> *Durante las próximas horas el tiempo será inestable. Se pueden desarrollar nubes de tormenta en cualquier momento y descargarán chubascos que en algunos*

puntos pueden ser intensos e ir acompañados de granizo. Predominará el viento del norte y en áreas de tormenta es posible que se intensifique.