



HAL
open science

Computing and Visualizing Agro-Meteorological Parameters based on an Observational Weather Knowledge Graph

Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel, Fabien Gandon,
Olivier Corby

► To cite this version:

Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel, Fabien Gandon, Olivier Corby. Computing and Visualizing Agro-Meteorological Parameters based on an Observational Weather Knowledge Graph. WWW 2023 - ACM Web Conference 2023 - International World Wide Web Conference, Apr 2023, Austin (TX), United States. ACM, Companion Proceedings of the ACM Web Conference 2023, 2023, WWW '23 Companion: Companion Proceedings of the ACM Web Conference 2023. 10.1145/3543873.3587357 . hal-04230234

HAL Id: hal-04230234

<https://hal.science/hal-04230234v1>

Submitted on 5 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Computing and Visualizing Agro-Meteorological Parameters based on an Observational Weather Knowledge Graph

Nadia Yacoubi Ayadi
University Côte d’Azur, Inria, CNRS,
I3S (UMR 7271)
Sophia Antpolis, France
nadia.yacoubi-ayadi@inria.fr

Catherine Faron
University Côte d’Azur, Inria, CNRS,
I3S (UMR 7271)
Sophia Antpolis, France
faron@i3s.unice.fr

Franck Michel
University Côte d’Azur, CNRS, Inria,
I3S (UMR 7271)
Sophia Antpolis, France
fmichel@i3s.unice.fr

Fabien Gandon
University Côte d’Azur, Inria, CNRS,
I3S (UMR 7271)
Sophia Antpolis, France
fabien.gandon@inria.fr

Olivier Corby
University Côte d’Azur, Inria, CNRS,
I3S (UMR 7271)
Sophia Antpolis, France
olivier.corby@inria.fr

ABSTRACT

Linked-data principles are more and more adopted to integrate and publish semantically described open data using W3C standards resulting in a large amount of available resources [7]. In particular, meteorological sensor data have been uplifted into public RDF graphs, such as WeKG-MF which offers access to a large set of meteorological variables described through spatial and temporal dimensions. Nevertheless, these resources include huge numbers of raw observations that are tedious to be explored and reused by lay users. In this paper, we leverage WeKG-MF to compute important agro-meteorological parameters and views with SPARQL queries. As a result, we deployed a LOD platform as a web application to allow users to navigate, consume and produce linked datasets of agro-meteorological parameters calculated on-the-fly.

CCS CONCEPTS

• **Information systems** → **Resource Description Framework (RDF)**; *Web applications*.

KEYWORDS

Knowledge graph, Semantic modelling, Observational data, Linked data, Agro-Meteorology

ACM Reference Format:

Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel, Fabien Gandon, and Olivier Corby. 2023. Computing and Visualizing Agro-Meteorological Parameters based on an Observational Weather Knowledge Graph. In *Companion Proceedings of the ACM Web Conference 2023 (WWW ’23 Companion)*, April 30-May 4, 2023, Austin, TX, USA. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3543873.3587357>

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

WWW ’23 Companion, April 30-May 4, 2023, Austin, TX, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-9419-2/23/04...\$15.00
<https://doi.org/10.1145/3543873.3587357>

1 INTRODUCTION

Integrating and interpreting meteorological data requires metadata about the features of interest (e.g. the air), their observed properties (e.g. air temperature or humidity), the specific location of a weather station, the time (instant or interval) at or during which the parameter was measured, and a variety of other information. In our previous works, we proposed a model for meteorological data and we built the WeKG-MF knowledge graph of observational weather data from historical archives of synoptic stations located in France and its overseas territories [10]. It is compliant with linked open data best practices [1] and accessible through a SPARQL endpoint¹. Since then, we enriched WeKG-MF with basic aggregate values of weather parameters (e.g., daily min/max temperatures), not initially included in Météo-France weather reports, based on official documentation published by the World Meteorological Organization (WMO) [9] and Météo-France. These aggregate values were semantically described and integrated as part of the WeKG-MF graph to serve different needs and studies. In specific domains such as agriculture, experts are interested in established and significant parameters such as the total monthly amount of precipitation, the number of days where the amount of precipitation is higher than a threshold over a given period, or the monthly average of daily maximum and minimum temperatures, and others. These agro-meteorological parameters play a crucial role in determining crop growth, development and yield and are essential for agricultural planning, management and decision-making.

In this paper, we present our work to formalize and implement a set of SPARQL queries to compute agro-meteorological parameters based on WeKG-MF, and to deploy a LOD platform allowing users to navigate, consume, visualize and produce linked datasets of agro-meteorological parameters calculated on-the-fly on WeKG-MF. We also provide an updated presentation of WeKG-MF and its data model. Section 2 provides an overview of WeKG-MF and highlights its design principles. Section 3, shows how WeKG-MF enables the calculation of relevant agro-meteorological parameters based on a set of SPARQL queries. Section 4 describes the LOD platform that enables users to calculate and visualize on-the-fly different agro-meteorological parameters.

¹<https://weakg.i3s.unice.fr/sparql>

2 WEKG-MF OBSERVATIONAL WEATHER KNOWLEDGE GRAPH

In this section we summarize the description of WeKG core model for linked weather data presented in [10], focusing here on its enrichment to represent aggregate meteorological features based on spatio-temporal slices of observations and we present the last version of the WeKG-MF knowledge graph.

2.1 WeKG Semantic Model

The WeKG model re-uses and extends domain ontologies, among which the Sensor, Observation, Sample, and Actuator (SOSA) [6] and Semantic Sensor Network (SSN) [5] ontologies for sensors and observations, the GeoSPARQL ontology for spatial features and their relations [3], the Time ontology [4] for temporal entities, and the RDF Data Cube ontology [8] for aggregation and multidimensionality features. The WeKG model addresses the semantic modelling issues related to meteorological observations and provides a generic model that can be adopted by weather data providers to publish their dataset as linked data. It captures the semantics of atomic and fine-grained weather observations by reusing and extending SOSA/SSN classes. It meets the specificity of meteorological observations, including the measurement periods for various weather properties. In this respect, OWL temporal classes are defined and reused to describe the measurement period of time of meteorological observations related to specific observable properties. For instance, observations related to rainfall indicate precipitations amounts recorded during the last 3, 6, 9, 12 or 24 hours of a specific synoptic hour². In WeKG model, depending on the observable property, it is possible to describe the temporal aspects of a meteorological observation using different temporal entities either to represent the specific synoptic hour of a result or the duration of measurement interval. At least one temporal entity is attached to each fine-grained atomic observation. Observations produced by sensors can rapidly reach enormous volumes. Aggregating them to create suitable meteorological indicators is a recurrent need in different domains. In this respect, we reuse the RDF Data Cube Vocabulary (DCV) [8] to describe multi-dimensional data structures. In our model, we define spatio-temporal slices of observations by fixing the spatial and temporal dimensions: the spatial dimension refers to the weather station, while the temporal dimension corresponds to a calendar interval. An observation (instance of `qb:Observation`) in a spatio-temporal slice is represented by a set of measures (instances of `qb:MeasureProperty`) each linked to an observable property with property `qb:concept`. Although SOSA/SSN ontologies are commonly used to represent knowledge about sensor data across domains, the definition of observable properties and features of interest, as well as their alignment with existing controlled vocabularies, are delegated to the community of interest. We established a SKOS vocabulary, guided by the Value Sets ontology design pattern³ (VS), to represent the potential observable properties and features of interest in the meteorological field.

²Main synoptic hours are 00, 06, 12, 18 UTC. Intermediate synoptic hours are 03, 09, 15, 21 UTC

³<https://www.w3.org/TR/swbp-specified-values/>

2.2 WeKG Thesaurus of Agro-Meteorological Parameters

Agro-meteorological parameters calculation leverages both fine-grained observations and aggregate values calculated over a period of time. These parameters include various indicators such as extreme monthly temperatures, Growing Degree Days (*GDD*), cumulative weekly/monthly *GDD*, minimum, maximum and average Growing Season Temperatures (*GST_{min}*, *GST_{max}* and *GST_{avg}*), cumulative (weekly/monthly) rainfall, the number of frost/summer/wet/windy days, average soil temperatures. All of these parameters are calculated over the growing season period which vary from one crop to another. Temperatures and precipitations are both key factors in the viability and growth of crops as well as the yield quantity and quality. For instance, during the growing season, it is important for experts to assess the heat accumulation based on the *GST* and *GDD* parameters. For most plants, phenological development stages from seeding to maturity is related to the accumulation of heat or temperature units above a threshold or base temperature below which no growth occurs. For instance, in viticulture, during the growing season that ranges from April to October (Northern Hemisphere), *GDD* and *GST* are calculated according to the base temperature of 10°C. Furthermore, frost events caused by extreme minimum temperatures can cause damage to emerging buds and shoots, especially if they occur during the first two months of the growing season. Also, high rainfall during June can have a negative impact on grapevine flowering and subsequent grape yield. While all of the previous assumptions are made for viticulture, changing to another crop such as wheat implies the same calculation methods even if thresholds may vary depending on the development stages defined for a certain crop. We extended the WeKG thesaurus of meteorological observable properties and features of interest to include new SKOS concepts that represent the agro-meteorological parameters. They are formally defined, documented and related to previously defined concepts.

2.3 WeKG-MF Knowledge Graph

WeKG-MF is a FAIR knowledge graph compliant with the WeKG semantic model and built from open weather data published by Météo-France. It now includes 10 years (2012-2022) of historical weather observations, initially downloadable as CSV files⁴. The RDF files are loaded into a Docker-based Virtuoso OS server. The WeKG-MF generation pipeline is reproducible and the mapping rules can be customized to meet the needs of other weather data providers. A documentation is available on our Github repository⁵.

WeKG-MF allows users to query atomic weather observations as well as aggregate values pre-calculated and stored in RDF Data Cubes. Atomic observations and pre-calculated RDF slices are loaded in different named graphs. An example SPARQL query retrieving daily average temperatures in the french region 'Nouvelle Aquitaine' during 2021 is presented in listing 1. An additional RDF Turtle file describes the WeKG-MF graph using the DCAT⁶, VOID⁷ and SPARQL-SD⁸ vocabularies.

⁴https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=90&id_rubrique=32

⁵<https://github.com/Wimmics/weather-kg/tree/main/meteo>

⁶<https://www.w3.org/TR/vocab-dcat/>

⁷<https://www.w3.org/TR/void/>

⁸<https://www.w3.org/TR/sparql11-service-description/>

Listing 1: Query to retrieve avg. daily temp. computed from the atomic observations in WeKG-MF recorded by weather stations located in “Nouvelle Aquitaine” French region.

```
SELECT distinct (STR(?label_region) as ?RegionName) ?stationName
?date ?avg_temp WHERE {
VALUES ?year {"2021"^^xsd:gYear}
# Query weather stations located in "Nouvelle Aquitaine".
?statURI a weo:WeatherStation; rdfs:label ?stationName;
dct:spatial [wdt:P131 [rdfs:label ?label_region; wdt:P2585 '75']] .
# Query slices for each statURI.
?slice a qb:Slice ; wes-dimension:station ?statURI;
wes-dimension:year ?year;
qb:observation [a qb:Observation;
wes-attribute:observationDate ?date;
wes-measure:avgDailyTemperature ?avg_temp]. }
ORDER BY ?date
```

3 COMPUTATION OF AGRO-METEOROLOGICAL PARAMETERS

The WeKG-MF graph is intended to serve different use case scenarios in several domains, including agriculture, biodiversity and climate studies. In the context of the D2KAB French project⁹, we identified several competency questions (CQ) that reflect the needs of experts to understand and compare weather conditions across years and to study their potential impact on crop growth and plant disease advent and management. These CQ have guided us to identify a set of agro-meteorological parameters that can be computed on WeKG-MF.

The first step towards the computation of agro-meteorological parameters is to pre-calculate and store aggregate values as RDF data cubes. We adopted a strategy which consists in combining several SPARQL CONSTRUCT queries, whose results sets are integrated into the spatio-temporal slice of daily maximum, minimum and mean air temperatures as well as daily total precipitation. Note that the WMO documentation [9] indicates methods to calculate daily total rainfall for the day R_d as the accumulated amount of precipitation recorded at a specific day d from 6 : 00 UTC till 6 : 00 UTC of the following day $d+1$. The same time interval is considered to calculate the maximum daily air temperatures $Tmax_d$ whereas minimum daily temperature $Tmin_d$ are calculated over the time interval spanning from 18 : 00 UTC the previous day $d - 1$ till 18 : 00 UTC of the day d .

We implemented these calculation methods in SPARQL and stored query results as part of the WeKG-MF graph. This allowed us to find a balance between calculating on-the-fly agro-meteorological parameters and pre-calculating and storing significant SPARQL queries results in order to manage the workload of the WeKG-MF SPARQL endpoint.

As an example, Listing 2 illustrates the SPARQL query to calculate GDD , GST_{avg} and total rainfall recorded in French weather stations (in the Northern hemisphere, $lat \geq 0$) during 2021. All of these agro-meteorological parameters are calculated for the viticulture's growing season and GDD are calculated based on the 10°C base temperature. A benchmark of 20 SPARQL queries serving as examples to illustrate how WeKG-MF can be queried is available¹⁰.

Listing 2: SPARQL query to calculate GDD , GST_{avg} and total rainfall over the Growing Season in 2021.

```
SELECT (SAMPLE(STR(?regionLabel)) as ?region) ?stationName ?year
(SUM(?gdd) as ?GDD) (AVG(?temp_avg) as ?GST_avg)
(SUM(?precipitation) as ?rainfall)
(SUM(?precip_daily) as ?rainfall_slice)
FROM NAMED <http://ns.inria.fr/meteo/observationslice/2021>
FROM NAMED <http://ns.inria.fr/meteo/weatherstation>
FROM NAMED <http://ns.inria.fr/meteo/observation/2021>
WHERE {
GRAPH <http://ns.inria.fr/meteo/weatherstation> {
?station rdfs:label ?stationName.
OPTIONAL {
?station dct:spatial/wdt:P131 ?region.
?region rdfs:label ?regionLabel; wdt:P2585 ?regionCode. }
}
{GRAPH <http://ns.inria.fr/meteo/observationslice/2021> {
?s a qb:Slice ; wes-dimension:station ?station ;
qb:observation [a qb:Observation ;
wes-attribute:observationDate ?date ;
wes-measure:avgDailyTemperature ?temp_avg ]
BIND( IF( ?temp_avg > 10 , ?temp_avg - 10, 0) as ?gdd)
FILTER (MONTH(?date) > 3 && MONTH(?date) < 11) }
}
UNION
{GRAPH <http://ns.inria.fr/meteo/observationslice/2021> {
?s a qb:Slice ; wes-dimension:station ?station ;
qb:observation [a qb:Observation ;
wes-attribute:observationDate ?date ;
wes-measure:rainfall24h ?precip_daily ]
FILTER (MONTH(?date) < 11) }
}
UNION
{GRAPH <http://ns.inria.fr/meteo/observation/2021> {
?obs a weo:MeteorologicalObservation;
sosa:observedProperty wevp:precipitationAmount;
sosa:hasSimpleResult ?precipitation;
sosa:phenomenonTime [a weo:Interval24h ];
sosa:resultTime ?datetime;
wep:madeByStation ?station.
BIND (xsd:date(SUBSTR(STR(?datetime), 1,10)) as ?date)
FILTER (CONTAINS (STR(?datetime), "T06:00:00"))
FILTER (MONTH(?date) < 11) }
}
BIND(YEAR(?date) AS ?year)
FILTER (?lat > 0) # only northern hemisphere
}
GROUP BY ?station ?stationName ?year
ORDER BY ?stationName ?year
```

4 VISUALISATION OF AGRO-METEOROLOGICAL PARAMETERS

In our previous work [2], we developed a Web application which demonstrates an interactive visualization to explore atomic observations and slices included in the WeKG-MF graph at different levels of temporal and spatial details. Besides navigating through the WeKG-MF graph, we developed new visualisations in order to allow lay users such as experts in agriculture to easily compute, visualize and compare agro-meteorological parameters (online demo¹¹).

From a graphical point-of-view, we have developed several features to enrich the WeKG-MF users' experience with the calculation and visualisation of agro-meteorological parameters. First, an interactive map augmented with markers indicating geo-spatial locations of weather stations is displayed (Figure 1a). The user is able to select a weather station by clicking on its marker and a time period by selecting two calendar dates indicating the start and end

⁹<https://www.d2kab.org>

¹⁰<https://github.com/NadiaYA2019/DemoKGViz-1/blob/main/WeKG-MF-SPARQL-Queries.ipynb>

¹¹<https://nadiaya2019.github.io/DemoKGViz-1/>

SPARQL endpoint → https://weakg.i3s.unice.fr/sparql			
Number of RDF Triples	190.037.088	Weather stations	62
Total weather Observations	24.778.231	Weather properties	32
Number of slices	363	Meteorological features	6
		Agro-meteorological parameters	14

Table 1: Key statistics of the WeKG-MF Knowledge Graph.

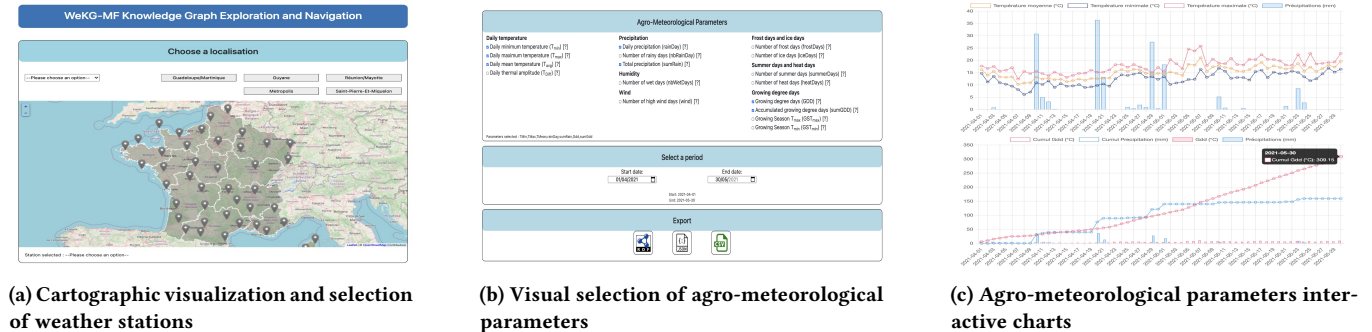


Figure 1: WeKG-MF-based Web application for computing and visualizing agro-meteorological parameters

dates of the growing season. Then, she is able to select one or more agro-meteorological parameters of interest.

In the current version of our Web application, we defined and implemented 14 different agro-meteorological in SPARQL. As the user selects parameters, different visualisations are proposed (Figure 1b). We provide different kinds of visualisations. A first visualisation shows different line charts, where the x-axis represents time, the y-axis represents the temperatures and the plots are the timeseries of daily min/max/avg temperatures. If the user selects daily total precipitation, the previously chart is updated automatically. This visualisation allows users to study daily temperatures and rainfall during the growing season period (top chart in Figure 1c). It is automatically updated if the user selects another time period. A second visualisation allows users to compare accumulated GDD and rainfall (bottom chart in Figure 1c).

5 CONCLUSION AND FUTURE WORK

In this paper, we presented (1) WeKG-MF, an RDF KG generated from the meteorological data observations available on Météo-France’s portal, based on a reusable, formal model exploiting and extending a network of well-known ontologies; (2) a base of SPARQL queries to compute a set of aggregators and agro-meteorological parameter and provide users with synthesized views of the huge number of measurements included in WeKG-MF. Use cases in the agronomy and agriculture domains are identified for which WeKG-MF will be integrated with other knowledge sources in order to meet expert requirements.

ACKNOWLEDGEMENTS

This work was carried out within the project D2KAB "From Data to Knowledge in Agronomy and Biodiversity" financed by the French National Research Agency (ANR-18-CE23-0017).

REFERENCES

- [1] Dean Allemang, James A Hendler, and Fabien Gandon. 2020. *Semantic web for the working ontologist*. ACM Press.
- [2] Nadia Yacoubi Ayadi, Damien Graux, and Catherine Faron. 2022. Multi-level Visual Tours of Weather Linked Data. In *Proc. 7th International Workshop on the Visualization and Interaction for Ontologies and Linked Data at 21st International Semantic Web Conference (ISWC 2022), Hangzhou, China, October 23, 2022 (CEUR Workshop Proceedings, Vol. 3253)*. CEUR-WS.org, 52–57.
- [3] Robert Battle and Dave Kolas. 2012. Enabling the geospatial Semantic Web with Parliament and GeoSPARQL. *Semantic Web* 3, 4 (2012), 355–370.
- [4] S. Cox and C. Little. 2020. *Time Ontology in OWL*. W3C Candidate Recommendation 26 March 2020. W3C Organism. <https://www.w3.org/TR/owl-time/>
- [5] Armin Haller, Krzysztof Janowicz, Simon J. D. Cox, Maxime Lefrançois, Kerry Taylor, Danh Le Phuoc, Joshua Lieberman, Raúl García-Castro, Rob Atkinson, and Claus Stadler. 2019. The modular SSN ontology: A joint W3C and OGC standard specifying the semantics of sensors, observations, sampling, and actuation. *Semantic Web* 10, 1 (2019), 9–32.
- [6] Krzysztof Janowicz, Armin Haller, Simon J. D. Cox, Danh Le Phuoc, and Maxime Lefrançois. 2019. SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *J. Web Semant.* 56 (2019), 1–10. <https://doi.org/10.1016/j.websem.2018.06.003>
- [7] Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, and Franck Michel. 2023. IndeGx: A Model and a Framework for Indexing RDF Knowledge Graphs with SPARQL-based Test Suits. *Journal of Web Semantics* (Jan. 2023). <https://doi.org/10.1016/j.websem.2023.100775>
- [8] Dave Reynolds and Richard Cyganiak. 2014. *The RDF Data Cube Vocabulary*. W3C Recommendation. W3C. <https://www.w3.org/TR/2014/REC-vocab-data-cube-20140116/>.
- [9] World Meteorological Organization. 2009. Handbook on CLIMAT and CLIMAT TEMP Reporting.
- [10] Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel, Fabien Gandon, and Olivier Corby. 2022. A Model for Meteorological Knowledge Graphs: Application to Météo-France Observational Data. In *Proc. International Conference On Web Engineering (ICWE 2022)*. 283–299.