

Tesis Doctoral

Visualización de datos geoespaciales aplicada a la meteorología

Diehl, Alexandra

2016-05-02

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Diehl, Alexandra. (2016-05-02). Visualización de datos geoespaciales aplicada a la meteorología. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Diehl, Alexandra. "Visualización de datos geoespaciales aplicada a la meteorología". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2016-05-02.



Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Departamento de Computación

Visualización de datos geoespaciales aplicada a la meteorología

Tesis presentada para optar al título de Doctor de la Universidad de
Buenos Aires en el área de Ciencias de la Computación

Alexandra Diehl

Directores de tesis: Dr. Claudio A. Delrieux

Dra. Marta E. Mejail

Consejera de estudios: Dra. Marta E. Mejail

Fecha de defensa: 2 de Mayo de 2016

Buenos Aires, 2016

Visualización de datos geoespaciales aplicada a la meteorología

Resumen. La visualización de datos geoespaciales cubre un gran espectro de técnicas visuales e interactivas para la representación, interacción y análisis de datos geoespaciales. Este tipo de datos se caracteriza por tener una referencia geográfica, también puede contener una componente temporal, y usualmente proviene de fuentes masivas y heterogéneas. El diseño de herramientas efectivas para asistir a los usuarios en el proceso de razonamiento analítico de datos complejos y dinámicos constituye un gran desafío. Un diseño efectivo comprende una combinación correcta de representaciones visuales, mecanismos interactivos, y procesamiento semiautomático. Una selección equilibrada de estos elementos depende del dominio específico de aplicación. Esto puede ser sólo alcanzado por un enfoque integral donde el usuario cumple un rol central, colaborando en cada paso de su diseño y evaluación. Diferentes factores influencian este proceso, por ejemplo, el flujo de trabajo de tareas específico del dominio de aplicación, los requerimientos del usuario, y la complejidad de los datos.

La presente tesis conduce al lector a través de diferentes enfoques que abordan el desafío de diseñar soluciones de visualización geoespacial eficientes, con aplicaciones específicas para la meteorología. Estos enfoques combinan distintas estrategias de visualización analítica para asistir a los pronosticadores en el análisis del pronóstico operativo del estado del tiempo. Los pronosticadores necesitan realizar un análisis rápido de la información, en el día a día. Proponemos un diseño novedoso que balancea la experiencia del usuario y el análisis semiautomático. Este diseño permite al usuario identificar tendencias y anomalías, analizar incertidumbre, y detectar errores del modelo numérico en forma ágil. También analizamos la integración de tecnologías de juegos serios y visualización científica. Nuestro diseño incorpora técnicas de visualización y tecnologías de juegos serios, en particular de simuladores de vuelo. El mismo demuestra los beneficios que esta in-

tegración puede brindar a la interacción con datos geoespaciales por medio de una aplicación de visualización 3D que explora ambientes topográficos y datos atmosféricos. Hemos evaluado la factibilidad y eficiencia de ambos enfoques por medio de diferentes casos de estudio diseñados en estrecha colaboración con meteorólogos, expertos del dominio. Esta evaluación fue realizada usando una metodología de diseño participativa e iterativa. Las conclusiones discutidas en este trabajo abrirán nuevas oportunidades de investigación para el diseño de soluciones de visualización de datos geoespaciales eficientes en el área de meteorología, en el análisis del estado del tiempo y soporte para la toma de decisiones.

Palabras clave: visualización científica, análisis visual interactivo, visualización analítica, datos geoespaciales, pronóstico del estado del tiempo, meteorología, ciencias de la atmósfera, juegos serios, simuladores de vuelo.

Visualizing geospatial data for meteorology

Abstract. Geospatial data visualization covers a huge scope of visual techniques and interactive mechanisms for the representation, interaction and analysis of geospatial data. This kind of data is characterized for having a geographic reference, it can have a temporal component and often comes from massive and heterogeneous sources. The design of effective tools to assist users on analytical reasoning on these complex and dynamic data constitutes a big challenge. Effective design entails a right combination of visual representations, interactive mechanisms and semi-automatic processing. A carefully chosen selection of these elements depends on the specific domain of application. This can be achieved by an integral approach where the user plays a central role, collaborating in each step of the design and evaluation. Different factors influence this process, for example, the specific domain task workflow, user's needs and data complexity.

This thesis guides readers through different approaches that address the challenge of designing efficient geospatial visualization solutions with specific applications in meteorology. Our approaches face different visual analytical strategies to assist forecasters in the analysis of operational weather forecasts. Forecasters need to perform quick analysis of the data on a daily basis. We propose novel designs that balance user experience and semi-automated analysis. They allow the user to identify trends and anomalies, to analyze uncertainty, and to detect numerical model errors in an agile process. We also analyze the integration of serious games technologies and scientific visualization. Our design combines visualization techniques and serious games technologies, in particular flight simulators. It puts forward the benefits that this integration can bring to the interaction with geospatial data by means of a specific data model design and a 3D visualizer of topographic environments and weather information. We evaluated the feasibility and efficiency of our approaches by means of several case studies

designed in close collaboration with meteorologists, who are the domain experts. This evaluation was performed using an iterative and participatory methodology. Conclusions discussed in this work will open new research opportunities to enhance the design of efficient geospatial visualization tools for meteorology, weather analysis and decision-making support.

Keywords: scientific visualization, interactive visual analysis, visual analytics, geospatial data, operational weather forecasting, meteorology, atmospheric sciences, serious games, flight simulators.

Dedicado a

*Jorge, que me señaló el cielo y me acompañó a volar, y a mi mamá,
Delicia, que con su gran amor fortaleció mis alas.*

Agradecimientos

Gracias a mis directores de doctorado, el Prof. Dr. Claudio Delrieux y la Prof. Dra. Marta Mejail, por darme esta gran oportunidad y acompañarme durante este trayecto con gran solicitud y paciencia.

También quiero agradecer a los profesores e investigadores que confiaron en mí, y me apuntalaron durante la realización de mi tesis: Gracias ¡Oh Meeeester!, queridísimo Prof. Dr. Meister Eduard Gröller, por su gran generosidad, por darme la oportunidad y el honor de participar de un grupo de tanta excelencia, y compartir su sabiduría no sólo científica sino también de vida. Gracias queridísimo Stefan, Prof. Dr. Stefan Brucker, por tu gran generosidad de compartir también tu enorme conocimiento y genialidad. Que gran honor trabajar contigo!

Gracias Juan, Dr. Juan Ruiz, otro científico genial, generoso, que honor y agradecimiento siento al trabajar con personas tan grandes en el ámbito científico. Gracias Celeste, Prof. Dra. Celeste Saulo, me abriste las puertas al fascinante mundo de la meteorología, tan generosamente, gracias por este honor tan grande y por compartir tu tan valioso tiempo y dedicación. Gracias Yanina, Dra. Yanina García Skabar, y todo tu equipo del Servicio Meteorológico Nacional de la Argentina por su gran apoyo y colaboración.

También muchas gracias a los profesores que me acompañaron durante todos estos años con una gran escucha y corazón: Prof.Dr. Julio Jacobo Berllés, a la Prof. Dra. Anita Ruedín, y a la Prof. Dra. Patricia Borensztein.

Gracias a mis compañeros de oficina: Dra. María Elena Buemi, Dr. Sebastián Ubalde, Dr. Francisco “Pachi” Gómez Fernández, Ismael Orozco,

Florencia Iglesias, Dr. Norberto Goussies, Dra. Roxana Matuk, Mailén Alsina, y Dr. Daniel Acevedo. Gracias al Dr. Pablo De Cristóforis y todo su equipo de robótica. Gracias a mi compañero Pablo Haramburu por las ideas y tiempo compartido.

Gracias a mis compañeros de Algoritmos I que me han ayudado tanto durante estos años de duro trabajo: a Gabriela Di Piazza, Sebastian Galimba, Emiliano Hoess, Rodrigo Castaño, Facundo Carrizo, y todos los que pude haber omitido.

Gracias al personal del departamento de computación, que siempre me asistieron en todo lo que necesitaba con gran corazón: muy especialmente a Aída Interlandi, Mariana De Martino, Diana Costa, Lara Rosemberg, Guillermo Alfaro, Mateo Pinna, Esteban Mocskos, Andrés Juarez, Sebastián Naury, y Alejandro Nieva.

Gracias a mis grandes amigos y colegas de la facu, como los quiero!! Mercedes Perez Millán, Francisco Soulignac, Marina Grouhaus, Juan Pablo Puppo, Nicolás Botbol, Fernando José Hernandez Gómez, y Viviana Cotik. Gracias a mis colegas de Algoritmos y Estructuras de Datos I, por tantos años de colaboración y compañerismo. Gracias a mi amigo y colega Guillermo Frank, por la ayuda y oreja. Gracias a mis colegas y amigos de CITEFA: Vicky, Javier, Lucas, y Nicolás. Gracias al Ing. Horacio Abbate por la generosidad de participarme de su proyecto, por su dirección, consejos, e ideas. Gracias a la Prof. Dra. Juliana Gambini por el tiempo de trabajo que compartimos juntas.

Gracias a mis colegas y amigos del Vis-Group Vienna: Johanna Schmidt, Alexey Karimov, Artem Amirkhanov, Gabriel Mistelbauer, Mathieu Le Muzic, Ivan Viola, Viktor Vad, Andrej Varchola, y Peter Mindek. Gracias a Anita Mayerhofer, que genia, y a mi amiga Tammy Ringhofer. Gracias al Prof. Kresimir Matkovic por confiar en mí, participarme en sus proyectos y compartir sus ideas conmigo. Gracias a mis amigos de Viena: Tamás Binker, "Gibs" (Brigitte Grüber), Irina Stör, Anna Paulin, y Natalia.

Gracias Prof. Dr. Turner Whitted, por confiar en mí para sus proyectos, y transmitirme tanta pasión por la computación gráfica. Gracias por ser siempre tan amable, generoso, y considerado. Asimismo, gracias Prof.

Dr. Curtis Wong por confiar en mí para su proyecto World Wide Telescope. Gracias Prof. Joachim Denzler por abrirme las puertas de su grupo en la Universidad de Jena, por su generosidad, y afecto. Gracias a Jaime Puente de Microsoft por tenerme en cuenta y ser tan bondadoso y amable conmigo, y por sus consejos profesionales. Gracias a mis colegas y amigos de Jena: Xiaoyan Jiang, Eric Rodner, Björn Fröhlich, Paul Bodesheim, Alexander Freytag, Mahesh Venkata Krishna, Michael Koch, Seana Duggan, Riya Menezes, y Ivana Sumanovac.

Gracias a Prof. Manganiello por su amistad y generosos consejos, por su gran apoyo, y el entusiasmo y amor que me inspiró por las ciencias.

A mis tesistas: Leandro Pelorosso y Rodrigo Pelorosso, que me acompañaron y acompañan con sudor y lágrimas en este proceso laborioso, pero tan recompensador.

Gracias a Emi Höss y Pablo Del Sel por confiar en mí para que dirija su tesis, denle gas, es el último esfuerzo!!

Gracias al P. José Bevilacqua por su cariño paterno, su acompañamiento espiritual, y su gran amistad y afecto.

Gracias a la Lic. Mariana Kos, que me ayudó a salir adelante y dar este paso tan grande, con mucho profesionalismo, humanidad, y afecto.

Gracias a todos los amigos que Dios me puso en el camino, y que me acompañaron durante este largo trayecto: Mariel Musso (Dra. al cuadrado), María Mercedes Sánchez, Natalia Croce, Florencia Ramírez, Gabriela Martinitto, y Adriana Magallan. A mi amiga y madrina Elizabeth Díaz, por sus rezos, apoyo, y ayuda continua.

Gracias miles a mi amiga Rita Ayoub, que puso el lomo conmigo para el “proof reading” de la tesis. Y también a mi amiga María Canepa de Weil, por su generosa colaboración.

Gracias a mis amigos de Torcuato y Regina que me bancaron todos los fines semana, y me mimaron con cafecitos “Lavazza”: Paola, Florencia, Tincho, Luis, César, Noelia, y Pedro.

Gracias a mi hermosa familia: a mi hermano, Rodrigo Diehl, por su gran paciencia y buen humor, a mi tío Aristóbulo Gómez por su apoyo incondi-

cional, y a mi primo Hernán Solimán por enseñarme una gran lección de vida, y ayudarme a reconocer que la vida no-es-fácil.

Gracias Jorge Anchividart a quien dedico esta tesis, porque remó conmigo, voló conmigo, a la par, desde el principio hasta el final, un gran compañero de vida que Dios puso en mi camino. Gracias a la familia de Jorge, especialmente a su mamá Elena Carli, que aguantó interminables tardes y noches en su casa, siempre me apoyó, tiró ondas positivas y rezó por mi. A Oscar, a Susana y Diego, por el aguante!!

Gracias a mi mamá, Delicia Gómez, a quien también dedico esta tesis, por todo, pero especialmente porque siempre me alentó a estudiar, dejó el todo por el todo por sus hijos, y dio resultados. Aleluya!!

Muchísimas gracias al Prof. Dr. Werner Purgathofer, el Prof. Dr.Pablo Minnini, y el Prof. Dr.Marco Mora Cofré por el gran honor de tener a tan excelsos científicos como miembros de mi jurado.

Gracias a todos, Gracias, Gracias, a Dios!!!

Contents

I Thesis Overview	18
1 Resumen en español	19
1.1 Introducción	19
1.1.1 Visión general	19
1.1.2 Desafíos	22
1.1.3 Planteo del problema	25
1.1.4 Metas	27
1.1.5 Contribuciones	28
1.1.6 Publicaciones	30
1.1.7 Estructura de la tesis	32
1.2 Resumen parte II	33
1.3 Resumen parte III	33
1.4 Conclusiones	34
1.4.1 Visión general	34
1.4.2 Principales hallazgos	36
1.4.3 Limitaciones	38
1.4.4 Trabajo futuro y problemas abiertos	38
1.4.5 Perspectivas y visión a futuro	38
2 Introduction	40
2.1 Overview	40
2.1.1 Challenges	43
2.2 Problem statement	45

2.2.1	Visual analytics for operational weather forecasting	45
2.2.2	Serious games applied to geovisualization	47
2.3	Goals	49
2.4	Contributions	49
2.5	Publications	50
2.6	Thesis outline	53
II	Visual analytics for operational weather forecasting	54
1	Related work	55
1.1	Overview	55
1.2	Visualization for Weather and Climatology	55
1.3	Comparative visualization	57
1.4	Uncertainty visualization	58
1.5	Analysis of time-series data	58
1.6	Analysis of spatio-temporal patterns	59
2	Visual Interactive Dashboard (VIDa)	61
2.1	Overview	61
2.2	Components	64
2.2.1	Semillón	64
2.2.2	Albero	80
2.2.3	Hornero	88
2.3	Data sources	88
2.3.1	Semillón	88
2.3.2	Albero	89
2.4	Implementation	90
2.4.1	Architecture and technologies	90
2.4.2	GPU visualization pipeline	90
3	Case studies	98
3.1	Overview	98
3.2	Semillón	99
3.2.1	Temporal trend analysis	99

3.2.2 Forecast verification	101
3.2.3 Forecast uncertainty analysis	103
3.2.4 Lessons learned	104
3.3 Albero	107
3.3.1 Extreme precipitation analysis	108
3.3.2 Albero for Technique Optimization	110
3.3.3 Lessons learned	112
3.4 Conclusions	113
 III Serious games applied to geovisualization	 114
 1 Related work	 115
1.1 Overview	115
1.2 Game engines	117
1.3 Flight simulators	119
1.3.1 FlightGear	120
1.3.2 Microsoft Flight Simulator X (FSX)	121
1.3.3 X-Plane	121
1.3.4 OpenFlight	122
1.4 OpenSceneGraph (OSG)	122
 2 3D Geovisualizer	 124
2.1 Overview	124
2.2 Design	124
2.2.1 Data sources	125
2.2.2 Data model	126
2.3 Implementation	128
2.4 Results	128
2.5 Limitations and future work	130
 IV Conclusions	 131
 1 Reflections on our work	 132

1.1 Overview	132
1.2 Main findings	133
1.3 Limitations	134
1.4 Future work	135
1.5 Open problems and perspective	135
1.6 Summary	137
V Appendices	138
1 Appendix A	139
1.1 Abbreviations	139
1.2 Nomenclature	141
Bibliography	141

List of Figures

1	Relación entre modelo de usuario y modelo de diseño (basado en Tory y Möller [121]).	20
2	Subconjunto de temas y subtemas presentados en la agenda de investigación de la ICA, en 2009 [129].	23
3	Modelo de generación de conocimiento para visualización analítica.	27
4	Relationship between user model and design model.	41
5	Topics of the 2009 ICA research agenda.	44
6	Knowledge generation model for visual analytics.	47
7	VIDa's component overview.	63
8	Semillón task workflow.	65
9	Screenshot of the current CIMA's website.	66
10	Screenshot of Semillón.	67
11	Semillón linked-views.	69
12	Curve-pattern analysis process.	71
13	Minimap timeline overview.	72
14	Meteogram linked-view.	75
15	Click-based sketching using the curve-pattern selector.	77
16	Curve-Pattern Δ metric	79
17	Visualization outputs of precipitation probabilistic forecasts at the National Weather Services.	82
18	Probabilistic forecasts: current automated workflow.	83

19	Albero: a new interactive workflow for probabilistic forecasting based on Analogs.	83
20	Screen-shot of Albero.	84
21	Übermap overview.	85
22	Interactive Forecast Layer close-up.	86
23	Analog Viewer description.	87
24	VIDa's architecture.	91
25	GPU visualization pipeline.	92
26	Transformation from LCC to Web-Mercator.	94
27	Sequence diagram of the multithreaded server.	95
28	Diurnal and nocturnal temperature development.	100
29	Trend analysis using Semillón.	101
30	Forecast verification using Semillón.	103
31	Forecast uncertainty analysis using Semillón.	104
32	Forecast subtraction analysis using Semillón.	105
33	Visual summary of case studies for Semillón.	106
34	Albero composition for the 1st of April, 2013.	109
35	Albero Analog Viewer showing analogs, observations and statistical summaries.	109
36	Interactive Forecast Layer showing different information from forecasts, observations and errors.	110
37	Albero screen-shot showing the extreme precipitation which happened during April, 2nd, 2013.	111
38	Albero: distribution of the Mean Squared-Error.	112
39	Typical areas of application of serious games.	116
40	Generic game application model.	118
41	Simplified class diagram of an OpenFlight specification.	123
42	Topographic enviroment data model of our 3D Geovisualizer.	126
43	Objects in the scene represented in a class diagram.	127
44	3D Geovisualizer first- and third-person views.	129
45	3D Geovisualizer first-person overview.	130

Part I

Thesis Overview

Resumen en español

1.1 Introducción

1.1.1 Visión general

Durante la última década, la disciplina de visualización ha experimentado un cambio gradual [61] en el foco de atención, antes con mayor énfasis en las técnicas de visualización, y ahora en la incorporación de los aspectos humanos de la visualización.

Este cambio se debe parcialmente a la intrínseca y creciente conexión entre la disciplina de visualización y áreas de investigación en percepción visual, así como al gradual entendimiento de las lecciones aprendidas durante la interacción entre los expertos de dominio y usuarios.

Ya en el 2004, Tory y Möller [120] enfatizaron la importancia que tiene considerar los factores humanos en el diseño de algoritmos de visualización. Los autores subrayaron cómo la percepción del usuario y la interacción con visualizaciones pueden influenciar una mayor compresión de los datos y del sistema completo.

El mismo año, Tory y Möller [121] propusieron una taxonomía que clasifica los algoritmos de visualización basados en modelos de diseño fuertemente influenciados por factores humanos. Los autores diferencian el “modelo de diseño” del “modelo de usuario”. El modelo de usuario se construye sobre la idea que el usuario tiene en mente acerca de la visualización,

mientras que el modelo de diseño se construye utilizando las técnicas de visualización y representaciones visuales que el diseñador selecciona para tratar de representar fielmente el objeto de estudio. La figura 1 muestra un diagrama de la interrelación entre usuarios, diseñadores y desarrolladores del sistema de visualización.

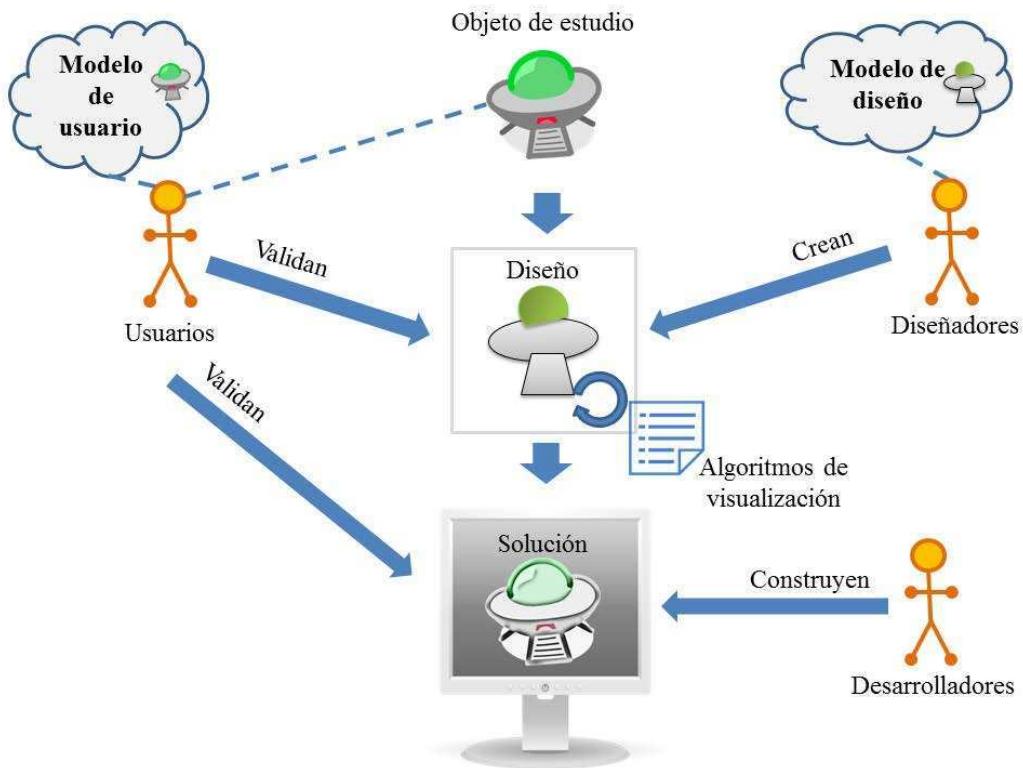


Figura 1: Relación entre modelo de usuario y modelo de diseño (basado en Tory y Möller [121]).

Diez años más tarde, Tory [119] enfatiza en su trabajo como los estudios de usuario están adoptándose cada vez más como una práctica usual en la disciplina de visualización. También Laramee y Kosara [61] identificaron los desafíos y los problemas abiertos asociados con los factores humanos y la visualización centrada en los factores humanos, denominando este enfoque “Visualización centrada en el usuario”. Los autores identificaron tres categorías de desafíos: centrados en el usuario, técnicos, y financieros. Los mismos se presentan sintéticamente en la tabla 1.

Estos desafíos se encuentran alineados con las áreas de “Geovisualización” (Visualización Geográfica o Geoespacial), “Visualización Analítica” [54],

Categorías	Desafíos
Aspectos centrados en el usuario	<p>Mejorar la comunicación y transferencia de conocimiento entre los expertos de dominio y los científicos del área de ciencias de la computación.</p> <p>Evaluar la usabilidad de las visualizaciones.</p> <p>Encontrar metáforas visuales efectivas.</p> <p>Seleccionar niveles óptimos de abstracción desde el punto de vista del usuario.</p> <p>Promover visualizaciones colaborativas entre expertos del dominio.</p> <p>Promover interacciones efectivas e intuitivas, especialmente en el campo de la realidad virtual.</p> <p>Representar la incertidumbre de los datos.</p>
Técnicos	<p>Aplicar técnicas escalables y de administración de grandes volúmenes de datos.</p> <p>Enfrentar la problemática del manejo de datos complejos (multidimensionales y multivariados).</p> <p>Crear técnicas efectivas de filtrado de datos y algoritmos para agregación de esos datos.</p>
Financieros	<p>Decidir qué tipo de tecnologías utilizar.</p> <p>Introducir estándares.</p> <p>Transformar los resultados de investigación en buenas prácticas y contribuciones para la sociedad.</p> <p>Crear visualizaciones que sean independientes de la plataforma.</p>

Tabla 1: Desafíos y problemas abiertos de la visualización de datos (adaptado de Laramee y Kosara [61]).

y en particular están alineados con el reciente campo de "Geovisualización Analítica" (Visualización Analítica Geoespacial) [5].

1.1.2 Desafíos

El área de investigación en geovisualización surge como una rama del área de las geotecnologías con el propósito específico de proveer métodos y herramientas para exploración visual, análisis, representación, y síntesis de la información geográfica.

Hace 15 años atrás, el foco de la geovisualización se centraba en cuatro puntos importantes:

- representación de la información geoespacial,
- integración de métodos computacionales y de visualización de datos geográficos,
- creación de interfaces efectivas para estas herramientas,
- usabilidad de estas herramientas.

En el año 2001, la International Cartographic Association (ICA) estableció, a través de su comisión de visualización y ambientes virtuales, una agenda de investigación en geovisualización con foco en la dinámica e interacción con datos geoespaciales. En esta agenda expresó que "todavía, la disciplina de visualización no está aprovechando el máximo potencial de los datos geoespaciales" [65]. Esta frase puede expresarse también como una pregunta: ¿Cómo explotar el máximo potencial de los datos geoespaciales por medio de la visualización? Durante los últimos 15 años, muchos fueron los esfuerzos de investigación enfocados en el desarrollo de herramientas para la representación de información geográfica, especialmente en la construcción de herramientas con el objetivo de entender fenómenos complejos, generación de conocimiento, y la exploración de estructuras y relaciones complejas de los datos geoespaciales.

MacEachren y Kraak [65] describen la geovisualización como un proceso de generación de conocimiento que provee las herramientas necesarias

para la adquisición de los datos, su transformación en información útil a través de la interpretación de esos datos, y finalmente la síntesis de ese conocimiento. El conocimiento es, entonces, adquirido a través de inferencias realizadas sobre la información.

Más recientemente, en el año 2009, las comisiones y grupos de trabajo de la ICA, prepararon una agenda de investigación en cartografía y ciencias geoinformáticas [129] donde establecieron nuevos desafíos para la geovisualización. Esta agenda enfatiza la aplicación de avances en el área de juegos serios para su utilización en geovisualización. En palabras propias de la agenda de investigación generada por el ICA: “Los nuevos desarrollos en el campo de los juegos y simuladores pueden ser provechosamente examinados con el fin de adoptar novedosas y efectivas herramientas y métodos para aplicar en geovisualización”. La figura 2 muestra un subconjunto de los principales temas y subtemas de investigación. En esta figura se pueden observar resaltados los temas más relevantes para la presente tesis. En los próximos párrafos se desarrollan algunos de estos desafíos.

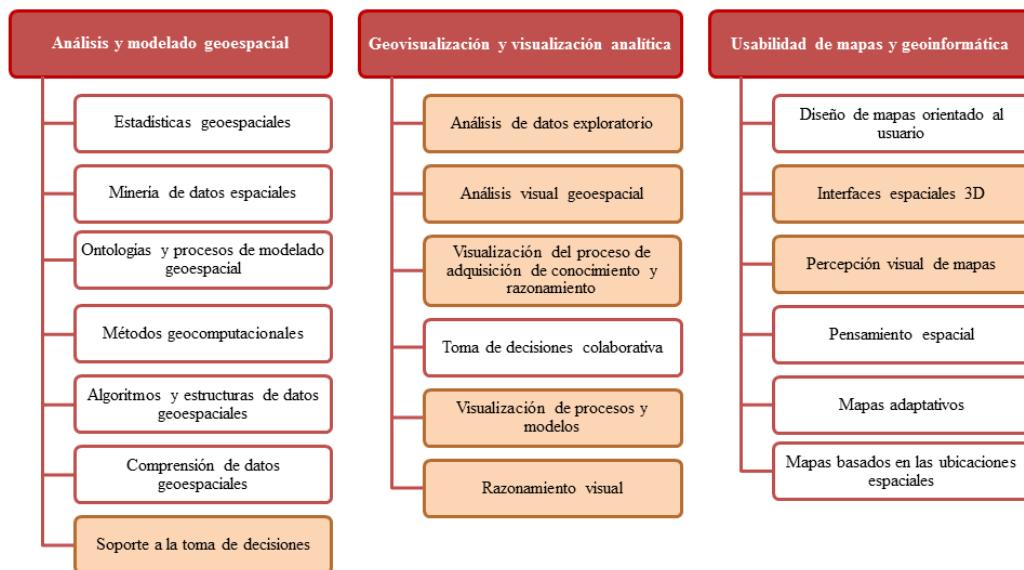


Figura 2: Subconjunto de temas y subtemas presentados en la agenda de investigación de la ICA, en 2009 [129].

Geovisualización y visualización analítica

La agenda de investigación presentó diferentes desafíos, no sólo relacionados a la representación de los datos, sino también a la interacción del usuario con estas visualizaciones. Dentro de los desafíos se encuentran: cómo integrar técnicas interactivas como “linking and brushing” con técnicas de “análisis exploratorio” y algoritmos de “minería de datos” con el fin de colaborar y asistir al usuario en el proceso cognitivo. Otro desafío comprende cómo diseñar soluciones para la toma de decisiones por medio de técnicas de visualización y razonamiento analítico. Hemos enfrentado estos desafíos en la parte II de la presente tesis proponiendo casos de diseño eficientes de soluciones basadas en visualización analítica. Nuestras soluciones asisten a los usuarios en los procesos de toma de decisiones y análisis de datos necesarios para las tareas de pronóstico operativo del estado del tiempo.

Análisis y modelado geoespacial

El propósito de este tema consiste en desarrollar técnicas de modelado y análisis que soporten la creación de conocimiento espacial y también soporten la toma de decisiones. Este tema se aborda en la parte II de la presente tesis.

Usabilidad de mapas y geoinformática

Los objetivos principales de este tópico conciernen a la identificación de tipos de usuarios a los cuales la aplicación va dirigida y a la construcción de diseños centrados en las necesidades del usuario. Algunas tendencias se refieren a métodos para mejorar los mecanismos de interacción y potenciar el análisis de los datos utilizando las capacidades de las bases de datos geoespaciales. Hemos analizado algunos temas relacionados con este tópico en la parte III de nuestra tesis.

1.1.3 Planteo del problema

La previsibilidad del tiempo afecta diversos aspectos de la vida humana y constituye un gran desafío [32]. Las ciencias de la atmósfera generan un gran número de simulaciones y observaciones a diario. Estos conjuntos de datos tienen asociado una incertidumbre en tiempo y espacio. La naturaleza caótica de la atmósfera amplifica aún más estos errores del pronóstico del estado del tiempo imponiendo un límite a la previsibilidad [50, 91]. Existe una gran necesidad de incorporar visualizaciones y mecanismos de interacción a las herramientas de meteorología con el fin de facilitar el análisis de tendencias, fenómenos asociados, incertidumbre, y errores del modelo. Las siguientes secciones describen los desafíos y problemas abiertos abordados en cada uno de los enfoques desarrollados en esta tesis.

Visualización Analítica Geoespacial para pronóstico del estado del tiempo operativo

La visualización analítica provee de herramientas para procesar, representar, explorar datos, obtener nuevos conocimientos de los mismos, y descubrir patrones de información ocultos, en un diálogo fluido entre el usuario y la aplicación. Los procesos de visualización analítica son incrementales, iterativos, y extendidos en el tiempo.

Andrienko et al. [5] resaltó la importante conexión entre la visualización analítica y la geovisualización. Ellos manifestaron la necesidad de una nueva generación de herramientas donde se combinen métodos computacionales con la experiencia y el conocimiento humano.

En nuestro caso particular, los meteorólogos especializados en pronóstico operativo, o abreviadamente pronosticadores, necesitan realizar análisis ágiles y tomar decisiones rápidas a diario.

El diseño efectivo de herramientas de visualización analítica requiere un entendimiento claro del flujo de trabajo de tareas y las necesidades específicas de los expertos del dominio. Las simulaciones del pronóstico del estado del tiempo presentan diferentes niveles de complejidad y resolución, tanto en las dimensiones de tiempo y espacio, y por ende, demandan dife-

rentes niveles de abstracción para facilitar su análisis. Los pronosticadores se enfocan en la identificación de posibles tendencias y en la detección de fenómenos del estado del tiempo, la incertidumbre asociada a la predicción de los mismos, y el análisis de errores en los modelos numéricos. Por esta razón, requieren de un acceso inmediato a la información y una comunicación ágil de los resultados de su análisis. Esto es crítico para la prevención de posibles desastres, para salvar vidas, y también evitar pérdidas económicas. La agenda de investigación presentada por Andrienko et al. [6] identificó los desafíos que involucra el diseño de soluciones basadas en visualización analítica para la toma de decisiones con soporte para datos espaciales. En su agenda, ellos manifiestan que para algunos dominios de aplicación específicos donde, “*La evaluación debe realizarse en unos breves períodos de tiempo, se requiere de herramientas analíticas y visualizaciones altamente eficientes*”.

Sacha et al. [103] exploraron el proceso de visualización analítica y expusieron los mecanismos cíclicos que comprende el discurso visual y analítico que conlleva a la construcción visual del conocimiento. La figura 3 muestra una versión adaptada del ciclo de visualización analítica propuesto por Sacha et al. En nuestro trabajo incorporamos el concepto del modelo de generación de conocimiento para visualización analítica con varios ciclos de procesamiento y análisis con el fin de construir soluciones eficientes que satisfagan las necesidades específicas al pronóstico operativo del estado del tiempo.

Aplicación de tecnologías de juegos serios a la geovisualización

Las nuevas tendencias en geovisualización se están enfocando en la representación realística de ambientes topográficos, en el modelado de procesos y fenómenos espacio-temporales (ambientes topográficos, imágenes satelitales, modelados de rutas, redes de afluentes y ríos, ambientes urbanos, estadísticas demográficas, datos de clima, entre otros). Pero más aún en nuevos métodos y puntos de vista que puedan proveer la utilización de las características de los videojuegos, y su forma de interacción. Un ejemplo de éstas características que puede ser utilizada y que está actualmente implementada en distintos videojuegos es el modelado multi-resolución (multi-

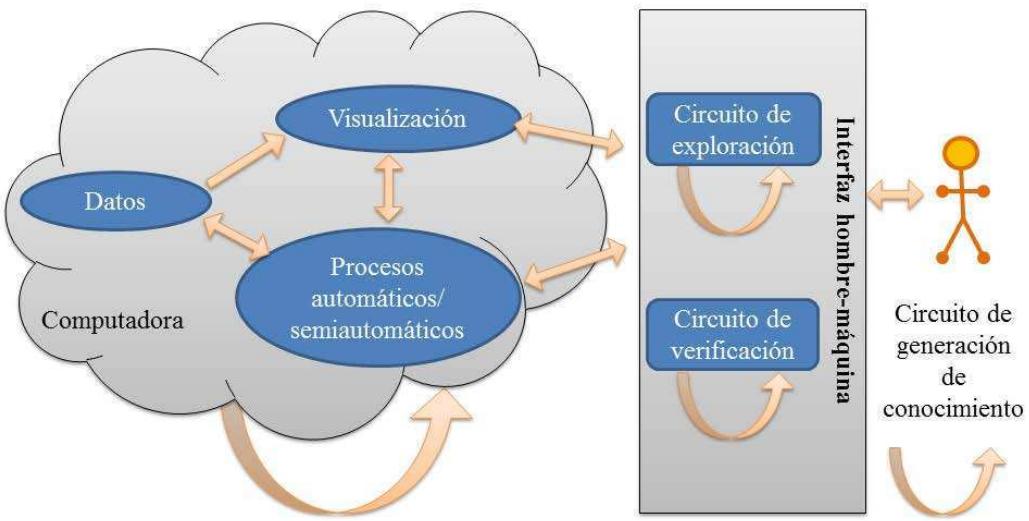


Figura 3: Modelo de generación de conocimiento para visualización analítica (reproducido de Sacha et al. [103]).

resolution modeling): esta técnica concierne la administración de niveles de detalle (LoD: Level of Detail) en aplicaciones de tiempo real.

Un ejemplo de técnicas de multi-resolución es el método de Clipmaps desarrollado por Losasso y Hoppe [64]. El desarrollo de estas técnicas de manejo de nivel de detalle es fundamental para proveer una experiencia de usuario interactiva, con un número de cuadros por segundo razonable (frame-rates).

Otra característica muy útil es la técnica de multi-textura (Multi-texturing). Esta técnica se refiere al mapeo de datos multidimensionales en el espacio. Por ejemplo, Han y Hoppe [40] desarrollaron una técnica para generar transiciones suaves entre imágenes satelitales de diferente resolución.

1.1.4 Metas

La presente tesis aborda la visualización geoespacial aplicada a la meteorología desde dos perspectivas diferentes: desde la visualización analítica haciendo foco en la toma de decisiones para el caso particular de pronóstico operativo del estado del tiempo, y desde un punto de vista tecnológico combinando motores de visualización con modelos de datos y bases de datos geoespaciales. En este trabajo nos enfocamos en el diseño de solu-

ciones de geovisualización centradas en las necesidades de los usuarios, utilizando una metodología de diseño incremental e iterativa. Hemos trabajado en estrecha colaboración con nuestros expertos de dominio, haciéndolos partícipes en cada etapa del proceso de diseño, con el fin de establecer la combinación más eficiente de técnicas de visualización y mecanismos de interacción que mejor se ajusten a sus necesidades específicas.

En la parte II abordamos el “diseño de soluciones basadas en visualización analítica para pronóstico operativo del estado del tiempo”. El propósito principal es el estudio de diferentes diseños basados en visualización analítica para asistir a los pronosticadores en el análisis de datos y toma de decisiones. Existe una oportunidad de mejora para cubrir la brecha explotando no sólo las técnicas de visualización interactiva, sino también, combinando las capacidades de cómputo y análisis automático con la experiencia y conocimiento de los usuarios, con el fin de diseñar herramientas efectivas y eficientes.

En la parte III planteamos la aplicación de tecnologías de juegos serios a la geovisualización. El propósito de este estudio es explorar el uso de diferentes simuladores de vuelo y motores de rendering para la construcción de aplicaciones de visualización que permitan al usuario navegar, y a su vez analizar la información desde un punto de vista en primera y tercera persona.

1.1.5 Contribuciones

Esta sección describe las principales contribuciones de cada uno de los enfoques desarrollados durante esta tesis.

Diseño basado en visualización analítica para pronóstico operativo del estado del tiempo

En esta parte proponemos dos estudios de diseño de soluciones basadas en visualización analítica para meteorología, específicas para pronóstico operativo donde los pronosticadores necesitan asistencia y análisis de datos inmediato para una ágil toma de decisiones. Hemos integrado nuestro en-

foque en una solución llamada Visual Interactive Dashboard (VIDa). La solución está diseñada utilizando una arquitectura cliente-servidor con una interfaz visual web, tomando ventaja del poder de cómputo de las Graphics Processing Unit (GPU) del lado del server, y de la flexibilidad que proveen las interfaces de usuario web, del lado del cliente. La misma fue desarrollada bajo un proceso iterativo e incremental tanto de diseño como evaluación, con la colaboración y validación constante de parte de los expertos de dominio, que son meteorólogos especializados en pronóstico del tiempo. Las interfaces visuales de nuestra solución guían al usuario desde resúmenes visuales simples hasta vistas con niveles de abstracción visual más avanzados y mecanismos de interacción más complejos, con el fin de analizar tendencias del pronóstico, la incertidumbre asociada a los mismos, y los errores del modelo utilizado. Para el primer estudio de diseño, desarrollamos el componente “Semillón” para pronóstico numérico a corto plazo, y luego para el segundo caso de estudio, desarrollamos el componente “Albero”, implementado como un add-on de la solución VIDa.

Para la evaluación de nuestros enfoques, proveemos de diversos casos de estudio que permiten que los expertos de dominio evalúen la efectividad y el potencial de nuestro enfoque.

Aplicación de tecnologías de juegos serios a la geovisualización

En esta parte de la tesis, realizamos un estudio de distintas librerías, simuladores de vuelo y motores de rendering con potencial para ser utilizados en geovisualización. En este estudio se resalta la posibilidad de reutilización de distintas técnicas de visualización, actualmente implementadas y evaluadas en distintos motores de juegos. También, proponemos la integración de mecanismos de interacción, visualización científica, y bases de datos geoespaciales en un prototipo de visualización 3D de topografías y datos meteorológicos. Nuestra solución incorpora fuentes de datos provenientes de satélites, datos de terreno, y de observaciones del estado del tiempo.

1.1.6 Publicaciones

Los resultados principales de esta disertación fueron publicados en:

- P1: 2016 L. Pelorosso, A. Diehl, K. Matković, C. Delrieux, J. Ruiz, M. E. Gröller, S. Bruckner. “Visualizing Uncertainty for Probabilistic Weather Forecasting based on Reforecast Analogs”. Poster aceptado para presentar en la EGU General Assembly 2016. Sesión: “Information in earth sciences: visualization techniques and communication of uncertainty”.
- P2: 2015 A. Diehl, L. Pelorosso, K. Matković, C. Delrieux, J. Ruiz, M. E. Gröller, S. Bruckner. “Albero: A Visual Analytics Tool for Probabilistic Weather Forecasting”. Poster presentado en el workshop “Big Data and Environment: Seizing the Data Deluge in Environmental Sciences”. Buenos Aires, Argentina, Noviembre 2015.
- P3: 2015 A. Diehl, L. Pelorosso, C. Delrieux, C. Saulo, J. Ruiz, M. E. Gröller, S. Bruckner. “Visual Analysis of Spatio-Temporal Data: Applications in Weather Forecasting”. Artículo presentado en la conferencia EuroVis 2015 y publicado en el journal “Computer Graphics Forum” 34 (3), 381-390.
- P4: 2013 A. Diehl, S. Bruckner, M. E. Gröller, C. Delrieux, C. Saulo. “Visual Trend Analysis in Weather Forecast”. Poster presentado en la conferencia IEEE VIS 2013.
- P5: 2012 A. Diehl, C. Delrieux. “Applications of Serious Games in Geovisualization”. Capítulo en el libro de M. Cruz-Cunha (Ed.), Handbook of Research on Serious Games as Educational, Business and Research Tools (pp. 25-46). Hershey, PA: Information Science Reference. doi:10.4018/978-1-4666-0149-9.ch002
- P6: 2009 A. Diehl, H. Abbate, C. Delrieux, J. Gambini. “Integration of Flight Simulators and Geographic Databases”. CLEI Electron. J. 12, 3 (2009), 1–8.

- P7: 2009 A. Diehl, C. Delrieux, H. Abbate, M. Mejail and M. Sanchez. “3D Geographical Environments and Geospatial Data Exploration Using Flight Simulators and Geodatabases”. Microsoft eScience Workshop 2009.
- P8: 2008 A. Diehl, H. Abbate, C. Delrieux and J. Gambini. “Integración de Simuladores de Vuelo y Bases de Datos Geográficos”. Conferencia Latinoamericana de Informática, CLEI. Septiembre 2008, Santa Fe, Argentina.
- P9: 2008 A. Diehl, H. Abbate and C. Delrieux. “Integración de Modelos de Entornos Topográficos Aplicada al Desarrollo de Simuladores de Vuelo”. Workshop de Investigadores en Ciencias de la Computación (WICC 2008). Gral. Pico, La Pampa, Argentina. Publicado en los anales de la WICC 2008.
- P10: 2008 A. Diehl, H. Abbate, C. Delrieux, J. Gambini. “Integración de Visualización de Ambientes Topográficos y Bases de Datos Geográficos Aplicado a Simuladores de Vuelo”. Primera Escuela de Ciencias de las Imágenes (ECImag 2008), Argentina.

Como es usual en la comunidad de visualización, las contribuciones al trabajo de investigación presentado en esta tesis es el resultado de una estrecha colaboración entre expertos de múltiples disciplinas. La autora de esta tesis es la principal autor de las publicaciones P2 a P10. En todos los casos, la autora realizó el análisis del software, fue responsable de la escritura de los papers, y contribuyó sustancialmente en la implementación y generación de los resultados. El principal autor de P1 es Leandro Pelorosso, quien está actualmente realizando su tesis de grado bajo la dirección del Dr. Juan Ruiz y de la Ing. Alexandra Diehl (quien es la autora de la presente tesis).

El Prof. Meister Eduard Gröller y el Prof. Stefan Bruckner contribuyeron desde los comienzos hasta la evaluación de la solución VIDa y sus componentes (Semillón, Albero, Hornero), apoyando y colaborando fuertemente con ideas y supervisión. Ambos son co-autores de las publicaciones P1 a P4. El Prof. Kresimir Matkovic contribuyó con ideas y sugerencias para el

proyecto Albero. La Prof. Celeste Saulo contribuyó desde la incepción con ideas, especificación de los requerimientos, y la evaluación de VIDa y Semillón. Dr. Juan Ruiz contribuyó con ideas, especificación de requerimientos, y la evaluación de todos los componentes de VIDa. Las tareas de programación, implementación, y desarrollo para VIDa y Semillón, fueron realizadas por la autora de esta tesis con contribuciones de Leandro Pelorosso para la implementación del servidor multi-hilos. Leandro Pelorosso es el responsable de la programación de Albero, mientras que Rodrigo Pelorosso es el responsable de la programación de Hornero. La evaluación de VIDa y Semillón fue realizada por el Dr. Juan Ruiz y la Prof. Celeste Saulo. La evaluación de Albero fue realizada por expertos del dominio pertenecientes al Centro de Investigación del Mar y de la Atmósfera (CIMA) y al Servicio Meteorológico Nacional (SMN): Dr. Juan Ruiz, Dr. Yanina García Skabar, Laura Aldeco y Cynthia Matsudo. Las publicaciones P5 a P10 corresponden al proyecto de aplicación de tecnologías de juegos serios a geovisualización. El Ing. Horacio Abbate y la Dra. Juliana Gambini contribuyeron con ideas y supervisión, siendo el Ing. Horacio Abbate director del proyecto.

Los directores de la presente tesis, el Prof. Claudio Delrieux y la Prof. Marta Mejail contribuyeron con ideas e inspiración, además de la supervisión de la presente tesis.

1.1.7 Estructura de la tesis

La presente tesis está dividida en cuatro partes. La presente introducción, parte I, resalta aspectos claves de la geovisualización y posiciona los problemas abiertos y desafíos, presenta el planteo del problema, las metas y contribuciones. La parte II presenta la investigación realizada en el área de visualización analítica para el diseño de soluciones de geovisualización. En particular para soluciones basadas en visualización analítica enfocadas en el pronóstico operativo del tiempo. Esta parte de la tesis describe dos casos de diseño de soluciones que integran visualización analítica y geovisualización para el análisis visual del pronóstico numérico operativo a corto plazo, y el análisis visual del pronóstico probabilístico. La parte III explora

los beneficios que la aplicación de tecnologías de juegos serios puede traer a la geovisualización. La parte IV presenta una discusión sobre el trabajo realizado y las perspectivas de trabajo futuro.

Cada parte contiene capítulos que cubren el estado del arte, el diseño de las soluciones, las conclusiones y limitaciones.

1.2 Resumen parte II

En esta parte de la tesis se analizan los distintos pasos del diseño participativo e incremental de soluciones basadas en visualización analítica. En particular se presentan casos de diseño para la toma de decisiones rápida para el pronóstico operativo del tiempo, para la detección de errores en el modelo, y análisis de la incertidumbre. El capítulo 1 presenta los antecedentes de trabajos previos relacionados a la visualización de pronóstico del tiempo y clima. El capítulo 2 describe el diseño de soluciones basadas en visualización analítica, sus componentes, las visualizaciones elegidas para el diseño de las soluciones, y los detalles de implementación. El capítulo 3 ilustra el proceso de evaluación a través de una serie de casos de estudio que integran los distintos aspectos requeridos para el análisis del pronóstico operativo, tanto numérico como probabilístico, las lecciones aprendidas y las limitaciones de nuestros enfoques.

1.3 Resumen parte III

La visualización tridimensional (3D) de ambientes topográficos e información geoespacial asociada permite la exploración de una diversidad de fenómenos geográficos (cambio climático, prevención de desastres ambientales, estudios de ecología, teledetección en tiempo real, entre otros).

Adicionalmente, el poder computacional de las placas gráficas hace posible el cómputo y visualización no solo en supercomputadoras, sino también en computadoras de escritorio y en dispositivos móviles. Desarrollar herramientas de visualización y exploración de ambientes topográficos interactivos utilizando esta tecnología representa un gran desafío, más aún cuando

se los combina con información proveniente de diversas fuentes en tiempo real que proveen datos meteorológicos.

Alcanzar niveles de interactividad razonables en este tipo de aplicaciones, como por ejemplo sobrevolar una topografía con técnicas de rendering realístico sobre un área geográfica dada, mientras se analiza información de variables atmosféricas, implica, potencialmente, el desarrollo de mecanismos interactivos avanzados. Las tecnologías de juegos serios han incorporado gradualmente técnicas de rendering realístico junto con otras funcionalidades interactivas de gran utilidad. Por lo tanto, estas tecnologías representan una gran oportunidad para desarrollar nuevas formas de interacción y exploración de datos geoespaciales [26, 72].

El capítulo 1 presenta antecedentes de trabajos en el área de simuladores de vuelo y motores de juegos, y sus aplicaciones al área de geovisualización. El capítulo 2 examina la factibilidad de nuestro enfoque por medio de pruebas de concepto de soluciones de geovisualización, reutilizando ciertas características de un motor de rendering, OpenSceneGraph (OSG) popularmente usado para el desarrollo de simuladores de vuelo.

1.4 Conclusiones

1.4.1 Visión general

En esta tesis estudiamos diferentes diseños de soluciones eficientes de visualización analítica para el área de geovisualización. Específicamente, herramientas para asistir a los pronosticadores en el análisis y toma de decisiones asociadas al pronóstico operativo del estado del tiempo. También, hemos analizado el uso de tecnologías existentes de juegos serios para geovisualización. Los antecedentes en este área postulan las siguientes preguntas y problemas abiertos:

- ¿Cómo influye el campo de aplicación en el diseño de las visualizaciones?
- ¿Cuál es la importancia de participar al usuario en el diseño?

- ¿Cuánto es el costo asociado a la adaptación de herramientas genéricas o librerías para un dominio específico en términos de curva de aprendizaje, tiempos de desarrollo y esfuerzo de adaptación e implementación?

En esta tesis nos preguntamos cómo debe ser un diseño eficiente de aplicaciones de visualización analítica, cuando uno se enfrenta a problemáticas específicas de un determinado dominio de aplicación.

Los estudios de diseño realizados durante esta tesis giran alrededor de las preguntas anteriores y tratan de dar respuestas a las mismas, y también a las siguientes preguntas:

- ¿Cómo integramos la dinámica de la operatoria de trabajo a los sistemas visuales existentes o a nuevos sistemas?
- ¿Cómo integramos las visualizaciones interactivas a los flujos de trabajo de tareas de los usuarios?
- ¿Cómo pueden las visualizaciones interactivas ser utilizadas para que los flujos de trabajo sean más eficientes?

Hemos dado respuestas parciales a estas preguntas por medio de un proceso de diseño participativo e integrando herramientas de visualización analítica a los flujos de trabajo de tareas de nuestros expertos de dominio. Hemos consolidado las lecciones aprendidas en una lista que sirve como guía para el proceso de diseño de soluciones basadas en visualización analítica para geovisualización:

1. Realizar un análisis exhaustivo de los flujos de trabajo de tareas involucrados en los procesos de análisis de los expertos de dominio.
2. Identificar los problemas existentes, funcionalidades deseadas, y oportunidades de mejora para los flujos de trabajo actuales y herramientas que estén utilizando los expertos.
3. Analizar las necesidades específicas de los usuarios, sacando provecho del conocimiento previo y la experiencia de los usuarios, con el fin de mejorar la productividad cognitiva.

4. Participar a los usuarios en todas las etapas del proceso de diseño, organizando el mismo en iteraciones cortas que consisten en identificar y priorizar las oportunidades de mejora, sugerir visualizaciones y posibles soluciones, discutir con los usuarios cuáles son los mejores enfoques, producir prototipos de software incorporando esos enfoques, diseñar casos de uso, y repetir el ciclo de diseño incremental.

Nuestros enfoques condujeron también a nuevas preguntas: ¿Hasta qué punto el diseño de las visualizaciones pueden recaer en la experiencia previa del usuario? Más aún en aplicaciones científicas donde se trata de utilizar métodos objetivos de validación de los algoritmos y técnicas propuestas. En nuestro caso particular, en el que tratamos de visualizar la incertidumbre para la mejora de la toma de decisiones: ¿Hasta qué punto se puede incorporar el conocimiento previo del usuario para reducir la incertidumbre, o medirla con mayor precisión, siendo que el conocimiento en sí mismo presenta una incertidumbre asociada, difícil de cuantificar?

1.4.2 Principales hallazgos

Las soluciones basadas en visualización analítica para la meteorología logran asistir a los usuarios en al menos cuatro tareas: (1) identificación y clasificación de fenómenos meteorológicos, (2) análisis de errores y sesgos sistemáticos de los modelos NWP, (3) análisis de la incertidumbre, y (4) toma de decisiones para pronóstico del tiempo operativo.

En primer lugar, encontramos una oportunidad de mejora en las soluciones actuales basadas en visualización analítica para pronóstico operativo del estado del tiempo, dado que no cubrían las necesidades de los pronosticadores con el fin de realizar análisis rápidos e interactivos. Las herramientas actuales requieren adaptación para las necesidades específicas, programación, y muchas de ellas configuraciones complejas.

Nuestra solución aborda esta problemática utilizando la metodología de visualización analítica o “mantra de visualización analítica”, proponiendo el uso de VIDa y su componente Semillón, e integrando nuestra solución a su flujo de trabajo de tareas de nuestros expertos de dominio. Ellos eva-

luaron nuestro enfoque a través de varios casos de estudio. Durante estas pruebas se completaron exitosamente tareas de análisis del pronóstico del estado del tiempo, y los expertos dieron comentarios positivos acerca de la usabilidad de nuestra solución.

En segundo lugar, hemos encontrado una mejora potencial del análisis del pronóstico probabilístico, descomponiendo el método automático, en pasos intermedios y visualizándolos. De esta forma el experto de dominio tiene acceso a información interna del método, visualizada a través de vistas vinculadas, que conectan distintas piezas de información. Hemos introducido un nuevo flujo de trabajo de tareas interactivo, también siguiendo el mantra de visualización analítica, para el análisis del pronóstico probabilístico del tiempo, agregando diferentes niveles de análisis en el componente Albero. Esta aplicación descubre los pasos internos del método y revela nueva información utilizada para generar los pronósticos probabilísticos utilizando la técnica RAR. Nuestra solución utiliza el concepto de Multiple Coordinated Views (MCV) con visualizaciones interactivas, vistas vinculadas, y métodos de análisis semi-automáticos. Este enfoque fue diseñado y evaluado utilizando una metodología de diseño incremental y participativa.

Finalmente, hemos realizado un estudio de algunos de los motores de rendering y juegos serios más populares. A partir de este estudio hemos propuesto un modelo de datos que se inspira en los modelos de datos de los simuladores de vuelo más populares, y un prototipo de aplicación visual 3D que le permite al usuario sobrevolar ambientes topográficos en primera y tercera persona, y analizar los datos meteorológicos en un entorno simulado realista. Hemos encontrado que diseñar y desarrollar soluciones para geovisualización utilizando motores de rendering o de videojuegos puede acortar los tiempos de prototipado, haciendo uso de funcionalidades originalmente desarrolladas para juegos pero adaptadas para propósitos de análisis visual de datos geoespaciales.

1.4.3 Limitaciones

Una de las limitaciones proviene de la escasez de datos y calidad de los mismos. La exactitud de los datos podría mejorarse utilizando data-sets locales provenientes de radares y estaciones meteorológicas de la región. En el caso de los datos de radar, todavía en el país se necesitan definir estándares de formatos e interfaces de datos.

También reconocemos la necesidad de mayor evaluación del enfoque de uso de juegos serios en aplicaciones de geovisualización. Actualmente, estamos considerando la construcción de un sistema de visualización 3D y seguimiento de tormentas utilizando motores de juegos.

1.4.4 Trabajo futuro y problemas abiertos

Como próximo paso se diseñará una nueva serie de sesiones de evaluación de nuestros enfoques por parte de los expertos de dominio para la solución VIDa y sus componentes. Estas sesiones propondrán la aplicación de medidas cuantitativas de la usabilidad, eficiencia, y experiencia del usuario.

Asimismo, cuando se estabilice el código se publicará como código abierto. Esto no solo servirá para que otros expertos evalúen la herramienta y contribuyan con más mejoras, sino también para que otros países y organismos con problemáticas similares puedan usar este trabajo como punto de partida.

Otro desafío futuro consiste en aplicar visualización analítica para el seguimiento de tormentas y el análisis del pronóstico de fenómenos relevantes como precipitación intensa, granizo, y tormentas severas.

1.4.5 Perspectivas y visión a futuro

A lo largo de esta tesis hemos presentado varios enfoques y casos de diseño de aplicaciones de visualización analítica para el pronóstico operativo del estado de tiempo. Hemos demostrado la aplicación de un proceso de diseño participativo e iterativo que utiliza el mantra de visualización analítica, y que está centrado en las necesidades del usuario. Hemos demostrado como

las herramientas de visualización analítica pueden asistir al pronosticador en la toma de decisiones, de manera eficiente. La guía propuesta en esta tesis para el diseño eficiente de soluciones basadas en visualización analítica puede generalizarse y reutilizarse en otros dominios de aplicación. Nuestro trabajo provee evidencias de utilidad en el diseño de herramientas basadas en visualización analítica aplicadas a la meteorología.

Chapter 2

Introduction

2.1 Overview

During the last decade, visualization as a discipline has experienced a gradual shift from the focus on visualization techniques to the focus on human centered visualization [61]. This shift partially comes from the intrinsic and growing connection between visualization and visual perception areas, and partially from the understanding and the lessons learned during the interaction with domain experts and users [79].

Early in 2004, Tory and Möller [120] highlighted the importance of considering human factors in the design of visualization algorithms. The authors stressed how user perception and visualizations can influence their understanding of the data. They also proposed in [121] a taxonomy that characterized visualization algorithms taking into account how human factors affect design models. The authors conceptualized a user model as a visualization idea that the user has in mind. This idea is translated to a visualization design model by an appropriate selection of visualization techniques and representations selected by a designer. Then, a developer creates a solution based on the design model. The validation process is performed by the user in the frame of an iterative and incremental methodology. Figure 4 shows the relationship between a design model and a user model.

Ten years later, Tory [119] emphasized how user studies became a standard practice in visualization research.

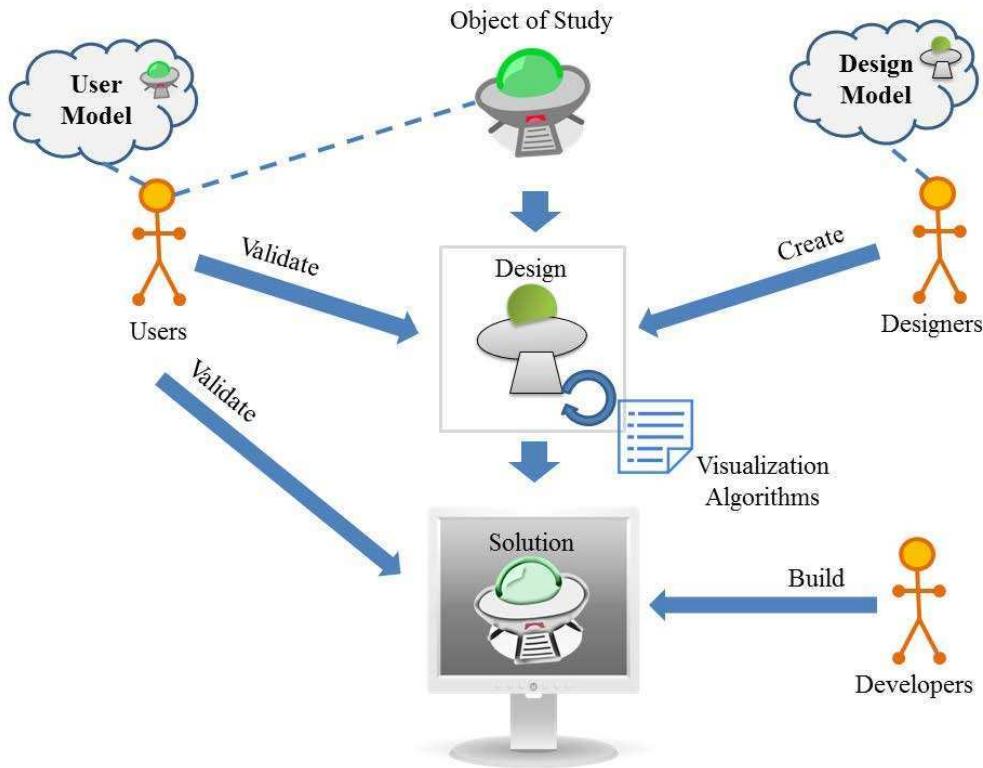


Figure 4: Relationship between user model, design model and participants of the design and development process of a visualization solution (based on Tory and Möller [121]).

Sedlmair et al. [107] defined a design study as “... *a project in which visualization researchers analyze a specific real-world problem faced by domain experts, design a visualization system that supports solving this problem, validate the design, and reflect about lessons learned in order to refine visualization design guidelines*”. They also proposed a design study methodology that we followed during all the stages of the present thesis.

Laramee and Kosara [61] identified challenges and open problems associated with human centered visualization. The authors have categorized three different classes of challenges: human centered, technical and financial. Table 2 shows a synthetic summary of the three categories of challenges. These challenges are aligned with the areas of “geovisualization” (geographic visualization) [27], “visual analytics” [54], and in particular with the recent field of “geovisual analytics” (geospatial visual analytics) [5].

Categories	Challenges
Human centered	<p>Improve communication and knowledge transfer between domain experts and computer scientists.</p> <p>Evaluate usability: interfaces, visual encoding, metaphors and abstractions.</p> <p>Find effective visual metaphors.</p> <p>Choose optimal level of abstraction from users' viewpoint.</p> <p>Promote collaborative visualization.</p> <p>Promote effective interaction: developing intuitive interactive techniques, especially in the field of virtual reality.</p> <p>Represent data uncertainty.</p>
Technical	<p>Enforce scalability and large data management.</p> <p>Tackle high data dimensionality (multidimensional and multivariate data).</p> <p>Create effective data filtering and aggregation algorithms.</p>
Financial	<p>Decide which technologies to use, among the large amount of possibilities.</p> <p>Introduce standards.</p> <p>Transform research into practices and contributions to society.</p> <p>Create platform independent visualizations.</p>

Table 2: Challenges and unsolved problems adapted from Laramee and Kosara [61].

2.1.1 Challenges

Geovisualization emerges as a branch of geotechnologies with the purpose of providing methods and tools for the visual exploration, analysis, synthesis and representation of geographic information [27, 99]. Geovisualization as a research activity was formally encouraged by the International Cartographic Association (ICA) commissions [45] and the National Science Foundation (NSF) [83]. In 2001, the commission on visualization and virtual environments of the ICA established a research agenda on geovisualization with focus on interaction and dynamics. The agenda stated that “...still, *visualization is not being taken advantage of to exploit the full potential of geospatial data*” [65]. This statement maybe rephrased as: “How to exploit the full potential of geospatial data by means of visualization?”. At that time, geovisualization set the focus on four main points:

- Representation of geospatial information.
- Integration of computational and geovisualization methods.
- Creation of effective interfaces for computer-aided graphical tools.
- Usability of visual environments.

More recently, in 2009, the ICA commissions and working groups prepared a Research Agenda on Cartography and Geographic Information (GI) science, where they have classified research topics and established new research challenges [129]. Figure 5 shows some of the subdivisions of the main topics and subtopics of research given by the 2009 agenda. Some of the topics are related to different aspects of our work.

Geovisualization and visual analytics

The agenda presented several challenges, not only related to the visual representations of data, but also to the user interaction. Among those challenges are how to integrate interactive techniques such as “linking and brushing” with “exploratory analysis” and “data-mining algorithms” to collaborate with the cognition process. Another challenge comprises how to design

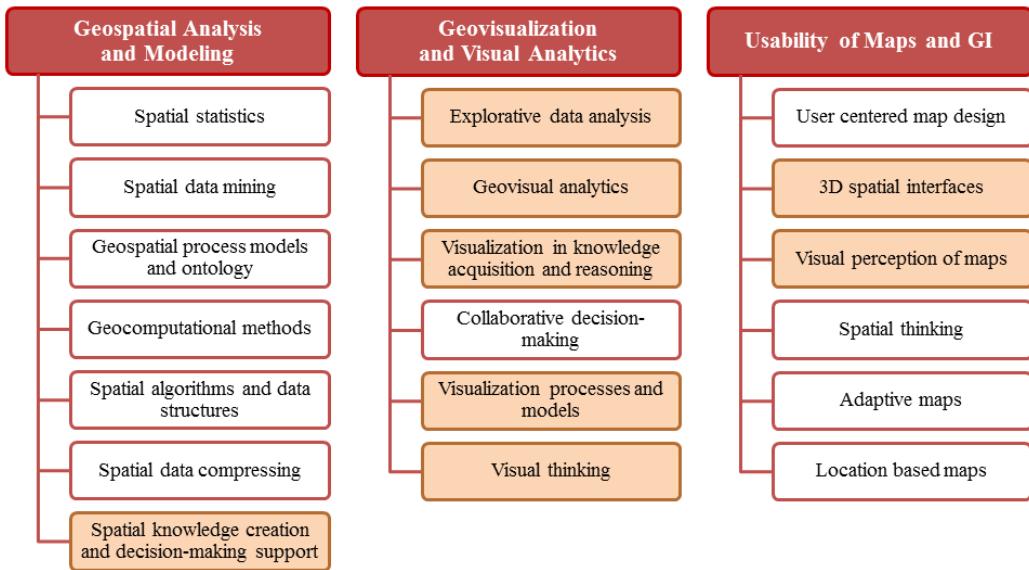


Figure 5: Some of the main topics and their subtopics of the 2009 ICA research agenda [129]. Highlighted are the subtopics that are more relevant to the present thesis.

decision-making solutions by means of visualizations and analytical reasoning. We faced these challenges in Part II of the present thesis by proposing efficient designs of visual analytics solutions for assisting users in their decision-making process of operational weather forecasting.

Geospatial analysis and modeling

The purpose of this topic is to develop techniques for creating spatial knowledge and to support spatial decision-making, this topic is addressed in Part II. In addition, new developments use realistic processes and spatial models to improve the representations of the real world, which is analyzed in Part III of this thesis.

Usability of maps and GI

The main challenges concern the identification of target users to whom the application is orientated to, and the construction of a user-centered design. The agenda encouraged the development of “*New geovisualizations using*

games and simulators technologies" [129]. Some trends are related to methods to enhance interaction and to empower data analysis by means of geodatabases. We elaborated this topic in Part III by integrating standard features from game frameworks with scientific visualization and geodatabases.

2.2 Problem statement

The predictability of weather influences many aspects of human life and constitutes a grand challenge [32]. Atmospheric sciences generate a large number of simulations and observations on a daily basis. These data-sets have associated uncertainty in time and space. The chaotic nature of the atmosphere amplifies weather forecast errors imposing a limit on predictability [50, 91]. There is a strong need to incorporate visualization and interaction capabilities to facilitate the analysis of trends, associated phenomena, uncertainty and model errors. Section 2.2.1 and section 2.2.2 describe the challenges and open problems faced on each path of the research done in this thesis.

2.2.1 Visual analytics for operational weather forecasting

Visual analytics provides tools for processing, representing and exploring data to get new insights and discover hidden information patterns, in a seamless "dialog" between the user and the application [131]. Visual analytics processes are incremental, iterative and extended in time. Therefore, communication, collaboration and adaptability are key components of these processes and require special attention.

Andrienko et al. [5] emphasized the important link between visual analytics and geovisualization. They also stated that a new generation of tools is needed combining computational methods with human experience and knowledge. In their words: "...*The illness of current tools is unrelated to computational power or the advancement in technologies. The lack of effectiveness is due to disregarding the domain experts knowledge and experience*" [5].

Meteorologists specialized in operational weather forecasting, shortly forecasters, have to perform agile analysis and make quick decisions on a daily basis, they and other stakeholders are the main decision-makers. The design of effective visual analytics tools requires the understanding of how forecasters work and their specific needs. Weather forecast simulations present different complexity and resolution in time and space, demanding different levels of abstraction to facilitate their analysis. Forecasters focus on the identification of possible trends and detection of weather phenomena, the associated uncertainty of predictions and the errors in the numerical models. They need immediate access to information to communicate them on time. This is critical to prevent possible disasters, save lives and avoid economic loss.

Andrienko et al. [6] identified the challenges involved in the design of geovisual analytics solutions for spatial decision-support. Their agenda stated that, for some application domains, “*The evaluation taking space in the context of time frames requires high and efficient analytics tools and visualizations*”. The authors raised challenging questions such as: “*...Do different types of problems require different kinds of geovisual support? If so, what are the relevant differences that influence the choice of approaches and methods? Is it productive to develop separate methodologies for different phases of problem-solving including intelligence, design and choice?*”. Colin Ware described in his book [131] key ideas for effective design of visualization solutions and also asked a critical question: “*...How best to transform the data into something that people can understand for optimal decision-making?*”. Quinan and Meyer [98], in their work about visual comparisons of weather features in weather forecasts, discussed: “*...What is the proper way to design, what will assist a set of experts, who have similar goals but individualized processes and domains?*”.

This thesis proposes different approaches and design studies for short-term weather forecasting, probabilistic forecasting and weather analysis. Our research is supported by a strong collaboration with domain experts and a careful selection of well-established visualization techniques. We present two different design studies with appropriate choices of different visual encoding, visualization techniques and interactive mechanisms, com-

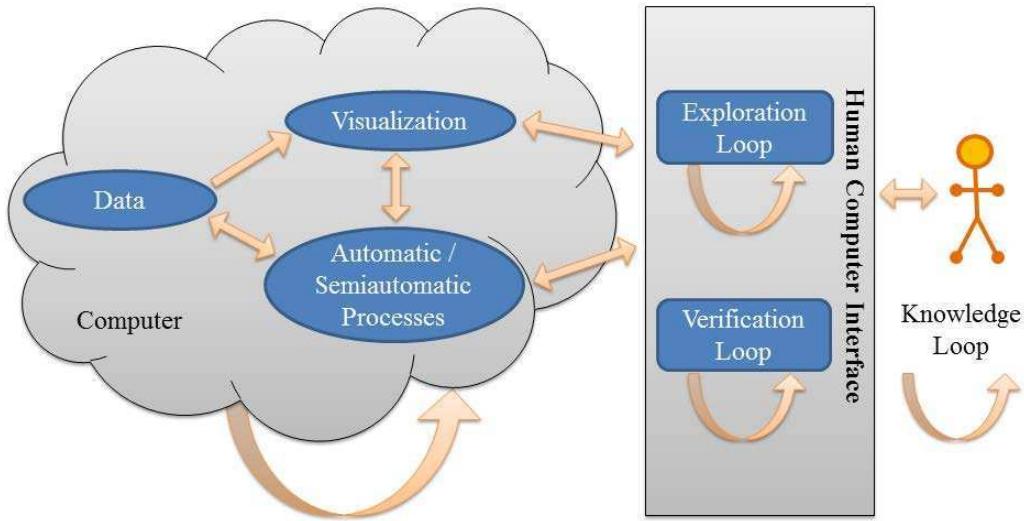


Figure 6: Knowledge generation model for visual analytics (reproduced after Sacha et al. [103]).

plemented with automatic and semi-automatic processes. These components are essential for an effective and efficient design in solutions for decision support. For this purpose, it is important to take into account the user's task workflow. The visualization solution should be integrated seamlessly to the task workflow, to maximize usability in order to facilitate the analysis of data. The design should empower the communication between the user and the application. Each visual interface should lead the user to a deeper understanding of the object of study.

Sacha et al. [103] explored the visual analytics process unveiling the cyclical mechanisms of the analytics discourse to convey visual knowledge. Figure 6 shows an adapted version of their “visual analytics loop”. We embrace the visual analytics loop concept and provide forecasters with efficient visual analytics solutions that address their specific requirements.

2.2.2 Serious games applied to geovisualization

New trends in geovisualization are setting the focus on the use of games and simulation technologies. In particular, the ICA Research Agenda emphasizes the application of game research to geovisualization: “...*Further new visualization developments in the field of games and simulators can be profitably*

examined in order to adopt novel and effective tools and methods for geovisualization” [129].

The use of game research advances and developments stands as a major venue to explore and enrich the capabilities of geospatial visualization tools [26, 51, 122]. Previous work on serious games explored interactive visualization of geospatial data. Among them, Hildebrandt [42] presents an application built by adapting Microsoft Flight Simulator 2000 with the goal of visualizing terrain and geospatial information. The author motivates his work stating that the adaptation of flight simulator technology has the potential to deliver very low cost 3D geospatial visualization capabilities compared with building an ad-hoc system or using other platforms and frameworks. McDonald et al. [72] presented a system for geographic visualization of landscape patterns, ecosystems and biodiversity, using flight simulator technologies to explore geographic regions. Also, Jones and Cornford [48] described a data-driven landscape visualization model developed to be used as a supporting model for real-time near-photorealistic visualizations. GEARViewer [67] is a geospatial rendering framework developed at VRVis in cooperation with the Austrian road and railway infrastructure. The system provides the users with alternatives for a given infrastructure project with multiple linked-views and real-time visual interaction. It also supports a wide variety of geospatial data formats. Although it does not use any specific game engine, its rendering framework includes a large number of features that are common for games, such as illumination, viewpoint controls, camera paths, scenario handling, etc.

Additionally, there are several applications such as Earth Systems and Virtual Globes that have developed features to enhance interactivity that were typical of games. One of them is Google Earth® [37] that has incorporated a flight simulator since its version 4.1. Another example is Microsoft World Wide Telescope®(WWT) [130] that allows for the creation of tours. WWT tours can navigate different environments such as the Earth, Mars, other planets, stars and galaxies.

2.3 Goals

This thesis tackles geospatial visualization for meteorology from two different perspectives: one, the visual analytics standpoint digging in the area of decision-making for the particular case of operational weather forecasting, and two, the technology standpoint by combining serious games frameworks and geodatabases with scientific visualization. It focuses on the design of geovisualization solutions centered in users' needs, using a participatory, iterative and incremental methodology of design. We have closely collaborated with the domain experts, keeping them always in the loop, in order to deliver a combination of visualization techniques and interactive mechanisms that fits the best for their specific needs.

In part [II](#), we address "Visual analytics design for operational weather forecasting". The main goal is to study visual analytics designs to support forecasters' analysis and decision-making. There is a gap between existing visualization tools for weather forecasting and forecasters' needs. This thesis aims to reduce this gap by using visual analytics approaches that take advantage of computational power and user experience in an effective and efficient design.

In part [III](#), we address 'Serious games applied to geovisualization'. The main goal is to explore the reuse of features from flight simulator and game frameworks to build 3D interactive visual applications that allow the users to navigate and analyze data from the first-person and third-person standpoint.

2.4 Contributions

This section reviews the main contributions of each of the approaches developed in this thesis.

Visual analytics design for operational weather forecasting

We propose two design studies, one for short-term weather forecasting and the other for probabilistic forecasting. Both designs are specific for oper-

ational weather forecasting, where meteorologists need assistance for immediate analysis. We integrated them by using a visual analytics solution, named Visual Interactive Dashboard (VIDa). Our solution is structured as a client-server architecture with a web visual interface that takes advantage of the power of Graphics Processing Unit (GPU) computing at the server side and the flexibility of using a web user interface at the client side. It was developed in an iterative process of design and evaluation, with constant feedback from our collaborators, who are meteorologists specialized in weather forecasting. The interactive visualization interface guides users from simple visual overviews to more advanced level of abstractions in order to identify and analyze weather forecasts. For the first design study, we developed Semillón for numerical weather forecasting and afterwards, for the second design study, we developed “Albero”, built on top of VIDa for probabilistic weather forecasting. Both of them were built to be used in the scope of operational weather forecasting. We provide several evaluation scenarios where the domain experts validated the effectiveness and potentiality of our solution. We also analyze the lessons learned, the domain experts’ feedback and the conclusions.

Serious games applied to geovisualization

We present a survey of serious games frameworks that can be used with geovisualization purposes. We emphasize reusability of visualization techniques and interactive mechanisms already implemented in serious games technologies. We propose a prototype that integrates interactive and intuitive visualizations with serious game features and geodatabases. Our solution incorporates relevant data and information sources, including terrain data, satellite imagery available from space agencies and weather forecast data.

2.5 Publications

The main results of this dissertation appeared in the following publications:

- P1: 2016 L. Pelorosso, A. Diehl, K. Matković, C. Delrieux, J. Ruiz, M. E. Gröller, S. Bruckner. “Visualizing Uncertainty for Probabilistic Weather Forecasting based on Reforecast Analogs”. Poster accepted for the EGU General Assembly 2016. Session: “Information in earth sciences: visualization techniques and communication of uncertainty”.
- P2: 2015 A. Diehl, L. Pelorosso, K. Matković, C. Delrieux, J. Ruiz, M. E. Gröller, S. Bruckner. “Albero: A Visual Analytics Tool for Probabilistic Weather Forecasting”. Poster presented at the workshop “Big Data and Environment: Seizing the Data Deluge in Environmental Sciences”. Buenos Aires, Argentina, November 2015.
- P3: 2015 A. Diehl, L. Pelorosso, C. Delrieux, C. Saulo, J. Ruiz, M. E. Gröller, S. Bruckner. “Visual Analysis of Spatio-Temporal Data: Applications in Weather Forecasting”. Paper presented at EuroVis 2015 conference and published in the journal “Computer Graphics Forum” 34 (3), 381-390.
- P4: 2013 A. Diehl, S. Bruckner, M. E. Gröller, C. Delrieux, C. Saulo. “Visual Trend Analysis in Weather Forecast”. Poster presented at IEEE VIS 2013 Poster session.
- P5: 2012 A. Diehl, C. Delrieux. “Applications of Serious Games in Geovisualization”. In M. Cruz-Cunha (Ed.), Handbook of Research on Serious Games as Educational, Business and Research Tools (pp. 25-46). Hershey, PA: Information Science Reference. doi:10.4018/978-1-4666-0149-9.ch002
- P6: 2009 A. Diehl, H. Abbate, C. Delrieux, J. Gambini. “Integration of Flight Simulators and Geographic Databases”. CLEI Electron. J. 12, 3 (2009), 1–8.
- P7: 2009 A. Diehl, C. Delrieux, H. Abbate, M. Mejail and M. Sanchez. “3D Geographical Environments and Geospatial Data Exploration Using Flight Simulators and Geodatabases”. Microsoft eScience Workshop 2009.

- P8: 2008 A. Diehl, H. Abbate, C. Delrieux and J. Gambini. "Integración de Simuladores de Vuelo y Bases de Datos Geográficos". Conferencia Latinoamericana de Informática, CLEI. September 2008, Santa Fe, Argentina.
- P9: 2008 A. Diehl, H. Abbate and C. Delrieux. "Integración de Modelos de Entornos Topográficos Aplicada al Desarrollo de Simuladores de Vuelo". Workshop de Investigadores en Ciencias de la Computación (WICC 2008). Gral. Pico, La Pampa, Argentina. Published in the proceedings of WICC 2008.
- P10: 2008 A. Diehl, H. Abbate, C. Delrieux, J. Gambini. "Integración de Visualización de Ambientes Topográficos y Bases de Datos Geográficos Aplicado a Simuladores de Vuelo". Primera Escuela de Ciencias de las Imágenes (ECImag 2008), Argentina.

As it is common in the field of visualization, the research contributions presented in this thesis are the result of a close collaboration between experts in multiple disciplines. The thesis author is the main author of the publications P2 to P10. In all cases, she performed the software analysis, was responsible for writing the paper, contributed substantial parts of the implementation and generated the results. The main author of P1 is Leandro Pelorosso, who is currently doing his undergrad thesis under the supervision of Dr. Juan Ruiz and Eng. Alexandra Diehl (the author of this thesis).

Prof. Meister Gröller and Prof. Stefan Bruckner contributed from the inception to the evaluation of VIDa and its components (Semillón, Albero, Hornero), strongly supporting the design of the tools with ideas and advisory. They are co-authors of publications P1 to P4. Prof. Kresimir Matkovic contributed with ideas and suggestions for Albero project. Prof. Celeste Saulo contributed in the inception, requirements specification and evaluation of VIDa and Semillón. Dr. Juan Ruiz contributed with ideas and domain expert knowledge in VIDa and all its components (Semillón, Albero, Hornero).

Programming and testing was performed by the author of this thesis for VIDa and Semillón, with contributions of Leandro Pelorosso for the imple-

mentation of the Multi-thread GPU server, Leandro Pelorosso for Albero and Rodrigo Pelorosso for Hornero.

Evaluation of VIDa and Semillón was performed by Dr. Juan Ruiz and Prof. Celeste Saulo. Evaluation of Albero was performed by domain experts from Centro de Investigación del Mar y de la Atmósfera (CIMA) and National Weather Services: Dr. Juan Ruiz, Dr. Yanina Garcia Skabar, Laura Aldeco and Cynthia Matsudo.

Publications P5 to P10 correspond to the project of application of serious games in geovisualization. Eng. Horacio Abbate contributed with ideas and guidance in the design.

Prof. Claudio Delrieux and Prof. Marta Mejail contributed with ideas and inspiration in addition to the guidance and supervision of the present thesis. Prof. Dr. Juliana Gambini collaborated with ideas and inspiration being a former director of the present thesis.

2.6 Thesis outline

This thesis is divided in four parts. Current Part I highlighted some key aspects of geovisualization, situated the open problems and challenges and presented the problem statement, goals and contributions. Part II focuses on research of visual analytics applied to the design of geovisualization solutions for operational weather forecasting using two different techniques: deterministic and probabilistic forecasting. Part III explores the benefits of applying serious game technologies in geovisualization. Part IV presents the conclusions and perspectives for future research.

Part II

Visual analytics for operational weather forecasting

Related work

1.1 Overview

Visualization tools have been widely used for the analysis of weather forecasts and climate data. Visual analysis can provide significant insights for meteorologists. In particular, visual analysis of spatio-temporal patterns assists them in understanding specific weather phenomena, atypical model behaviors and errors.

This chapter describes related work in the area of visualization and visual analytics for the application domain of weather forecasting. It classifies previous work based on different aspects that are relevant to our work and it is organized as follows. Section 1.2 describes current visualization systems applied to weather forecasting and climatology, Section 1.3 presents comparative visualization works and section 1.4 describes works in uncertainty visualization. Section 1.5 and 1.6 analyzes works in visual analysis of time-series data and spatio-temporal pattern research, that are related our work.

1.2 Visualization for Weather and Climatology

There has been extensive work done in the area of visualization systems and visual analytics applied to weather forecasting and climatology. In the scope of this work, only the most representative examples are summa-

ritized as background references to our approach. The University Corporation for Atmospheric Research (UCAR) presented a list of post-processing tools for weather visualization. Among them are: Integrated Data Viewer (IDV) [80], Visualization and Analysis Platform for Ocean, Atmosphere, and Solar Researchers (VAPOR) [21], and the Grid Analysis and Display System (GrADS) [123]. VAPOR provides advanced interactive 3D visualizations of massive data with focus on scientific needs. It allows the simulation and visual analysis of turbulence on gridded domains. It can run on desktop computers equipped with advanced GPUs. GrADS is a visualization system that uses a mix of Graphical User Interface (GUI) and command-line scripts to derive post-processed data visualizations. Command-line scripts and programming features can be seen as an advantage in terms of flexibility, but also, make the user experience and interaction difficult. Instead, visual interfaces are proven to be more effective in terms of cognitive productivity [131].

The Ultrascale Visualization Climate Data Analysis Tools (UV-CDAT) is a powerful system jointly developed by several institutions, universities, and private companies [136]. The tool set integrates data analytics, ensemble analysis, uncertainty quantification, metrics computation, and visual analysis applied to climate big data. Also, Song et al. [110] presented a visualization system that performs analysis of multi-dimensional atmospheric data-sets using physics-based atmospheric rendering, illustrative particles, and glyphs. The aforementioned visualization systems enable a full and flexible visual analysis, but they require many parameter settings, and in some cases, even programming tasks.

We are particularly interested in visualization systems available as web applications. Microsoft Research FetchClimate [77] is a climate data retrieval service that operates over the cloud. It displays an integrated webmap using Bing Maps, several linked-views and context-menus that visualize and interact with the data. This application focus on climatology rather than weather forecasting. Also, WeatherSpark Beta [133] provides multiple views such as an integrated mapview, charts and glyphs depicting weather conditions (sunny, cloudy, rainy, etc.), and historical data for a

given weather station or a given city anywhere in the world. In WeatherSpark application, the main component is the linked time-chart that plays a major role in the application. Furthermore, Google Maps [38] has incorporated geolocated information about weather conditions. It has integrated a new weather forecast feature that displays weather conditions of a city using glyphs. Other examples are Weather.com [116], Wetter.de [135], Weather Underground [134], and also the 3D weather visualization system Terra3D [112]. These cases are mainly targeted at general users rather than domain experts. In relation to the aforementioned tools, our domain experts stated: “the available web applications do not allow us to perform operations among forecasts, they do not provide us with a global overview, and do not allow us to compare multiple forecasts”. Moreover, we found only a few options in these applications for comparative visualization among different time-steps and/or different layers.

1.3 Comparative visualization

Gleicher et al. [34] presented a survey of visual comparison techniques. They created a taxonomy following three main categories: juxtaposition, superposition, and explicit encoding. They stated that all visual comparison designs can be constructed using building blocks of any of those categories, or a combination of them. There are several techniques that can be applied to compare 2D scalar-fields. Malik et al. [66] presented an approach for visual comparison using multi-image views that preserves contextual information. Schmidt et al. [106] proposed a multi-image view technique that used hierarchical clustering. It enables the scalability of the method with respect to the number of the images.

Köthur et al. [59] presented a visualization system that works with 2D scalar-field distributions of atmospheric data. They employed a visual analytics approach that enables users to extract and explore sets of 2D scalar-field distributions. A hierarchy of clusters is also used to aggregate different time-series according to their characteristics. Their solution allows the user to interactively explore and regroup clusters according to their needs. It has

a visual summary view that displays 2D spatial distributions side by side and allows examination of similarities and differences. SimEnvVis [85] provides a comprehensive set of comparative techniques such as clustering and parallel coordinates visualizations to compare multiple climate simulations.

1.4 Uncertainty visualization

Potter et al. [95] described an extensive overview of the state-of-the-art in Uncertainty Visualization. It contains a taxonomy of different approaches for visualizing uncertainty based on data dimensionality and uncertainty dimensionality. Thomson et al. [117] proposed a typology for uncertainty visualization that associates data, uncertainty and tasks, with an appropriate visual representation. Many tools are available for meteorologists to gain insight and to communicate the uncertainty of ensembles. Deitrick and Ed-sall [23] experimentally concluded that uncertainty visualizations can affect the decision-making process. Potter et al. [96] presented Ensemble-Vis, a framework that combines multiple linked views to facilitate the visual analysis of ensemble data, focused on short-term weather forecasting (short-time spans) and climate modeling (very long-time spans). They provided spatial and temporal overviews combined with detailed statistical views. These views give insight into the distribution and uncertainty of model simulations, and emphasize the probabilistic characteristics of the ensemble. Sanyal et al. [105] introduced Noodles, a framework that combines ribbon and glyph-based uncertainty visualization, spaghetti plots, iso-pressure color maps, and data transect plots to assess uncertainty in the data. Their approach has proven to be particularly helpful for finding outliers in an ensemble run.

1.5 Analysis of time-series data

Aigner et al. [1] provided a complete survey of different visualization techniques for time-dependent data and time-series analysis. Hochheiser and Shneiderman [43] presented TimeSearcher 1, a visual exploration tool that

combines time-box queries with overview displays, query-by-example facilities and support for queries over multiple time-varying attributes. TimeSearcher 2 [16] is an extension of TimeSearcher 1 for long time-series visualization using data filtering and pattern-query specification. It allows the user to filter the data and reduce the scope of the search, and also to perform a specific pattern search anywhere in between the remaining data. Fails et al. [28] presented a visualization tool, PathFinder, for searching and discovering temporal patterns within multivariate and categorical data-sets. The tool has two main panels: the pattern design panel and the results visualization panel. Users can define patterns by setting constraints on events and time-spans. The tool allows the analysis across those events and the connection of multiple events together. Results can be explored through coupled ball-and-chain and tabular visualizations.

1.6 Analysis of spatio-temporal patterns

There have been extensive studies in the area of spatio-temporal visualization and geospatial visual analytics. Andrienko and Andrienko [8] have broadly described different techniques for visual analytics of spatio-temporal data. Additionally, Andrienko et al. [7] proposed a methodology to analyze both temporal and spatial behavior based on a Self-Organizing Map (SOM). They presented an interactive visual interface consisting of a matrix display that provides “space-in-time SOM” and “time-in-space SOM” analysis and a control panel that provides parameter setup. Also, Jänicke et al. [46] proposed a visualization method for multivariate data that transforms high-dimensional attribute spaces into 2D point-clouds. In this work, historical climate data is grouped into seasons, and then transformed into point-clouds. Brushing techniques are applied for the analysis of temporal patterns and similarities.

Bruckner and Möller [15] proposed a system for the visual exploration of a simulation parameter space to assist the user in the generation of effects such as smoke and explosions. They split each simulation sequence into a number of the most representative segments. Then, they compared the

simulation-space similarities at different points within the temporal evolution of each simulation. They depicted subsets of the cluster’s members in a timeline at different temporal compression levels. Krstajic et al. [60] presented a technique for the interactive visual analysis of multiple time-series events, named CloudLines. They used distortion techniques on a timeline to accommodate a large number of data items. They employed different “Focus+Context” techniques for inspection of individual items in the highly compressed areas.

Kehrer and Hauser [52] presented an extensive survey about different techniques related to multivariate and multidimensional data visualization. Ware and Plumlee [132] highlighted specific design guidelines for weather forecasts visualization.

We employ a high level and qualitative variation analysis, which is based on the concept of families of curves [53, 58]. We apply this approach to the analysis of weather trends and atypical model behaviors. The ideas presented in Coto et al. [22], although not related to weather forecasts, provided an important background to our approach. Coto et al. described a classification based on “time-signal curve types” for early detection of breast cancer. Those curves were used to classify a possible tumor, and each curve type represented a characteristic of that tumor. For example, curves that show increasing values were indicators identifying possible benign lesions. For each data-point, they defined a distance-metric among temporal functions referred to as Time-Activity-Curves (TAC). Fang et al. [29] proposed a system to classify and segment 3D medical images based on temporal characteristics. Woodring and Shen [137] proposed a method to explore temporal trends at different resolutions using wavelets. With wavelets they characterized the data-points and transformed them into time-series curve sets. These points are classified into clusters of similar activity.

Chapter 2

Visual Interactive Dashboard (VIDa)

2.1 Overview

The state of the atmosphere can be described by its characteristic meteorological variables (i.e., temperature, pressure, humidity content, wind direction and velocity). The future evolution of these variables can be predicted to a certain degree of accuracy by Numerical Weather Prediction (NWP) models, given that a good representation of the current atmospheric state is provided. NWPs are used to predict current and future state of the atmosphere in time and space.

The motivation for this work arose from collaboration projects of more than four years of joint work with scientists from the TU Wien, University of Bergen, Universidad Nacional del Sur, the CIMA and the National Weather Services in Argentina. The main domain of application is *operational weather forecasting*. Operational weather forecasting requires immediate access to the results to perform quick analysis of the information. Meteorologists need a weather forecasting tool with quick visual mechanisms to explore multiple forecasts.

The result of our collaborative work is a visualization solution named “VIDa” (Visual Interactive Dashboard).

We designed our solution using a participatory, iterative and incremental methodology of design.

The participatory design process consisted on:

1. Analyze users' requirements.
2. Identify and prioritize problems and gaps.
3. Suggest visualizations and solutions.
4. Discuss with users the best approaches.
5. Produce a software prototype incorporating selected approaches.
6. Design case studies to validate the prototype.
7. Evaluate the case studies with domain experts.
8. Discuss and reflect on the lessons learned.
9. Repeat the incremental design cycle until the team agrees that the solution sufficiently satisfies the users' requirements.

VIDa is a Coordinated and Multiple Views System [100] specifically designed for operational weather forecasting. It is organized as a series of linked and coordinated views. These views are connected by interactive mechanisms in a fluent dialog (as defined by [131]) between the user and the application. We had iterative meeting cycles with the meteorologists to identify the users' requirements and to develop incremental prototypes. Our goals include understanding their current needs and the provision of new mechanisms to interactively visualize and analyze weather forecasts. At the same time, we aim to maintain the key functionality of their current system. The iterative and incremental process of refinement of our solution allowed us to align our work with their needs. Sometimes as computer scientists, we propose ideas that are novel from the visualization point of view but not suitable for the specific experts' needs. Instead, by using this methodology, in each iteration we can readjust our proposals to the requirements of our domain experts.

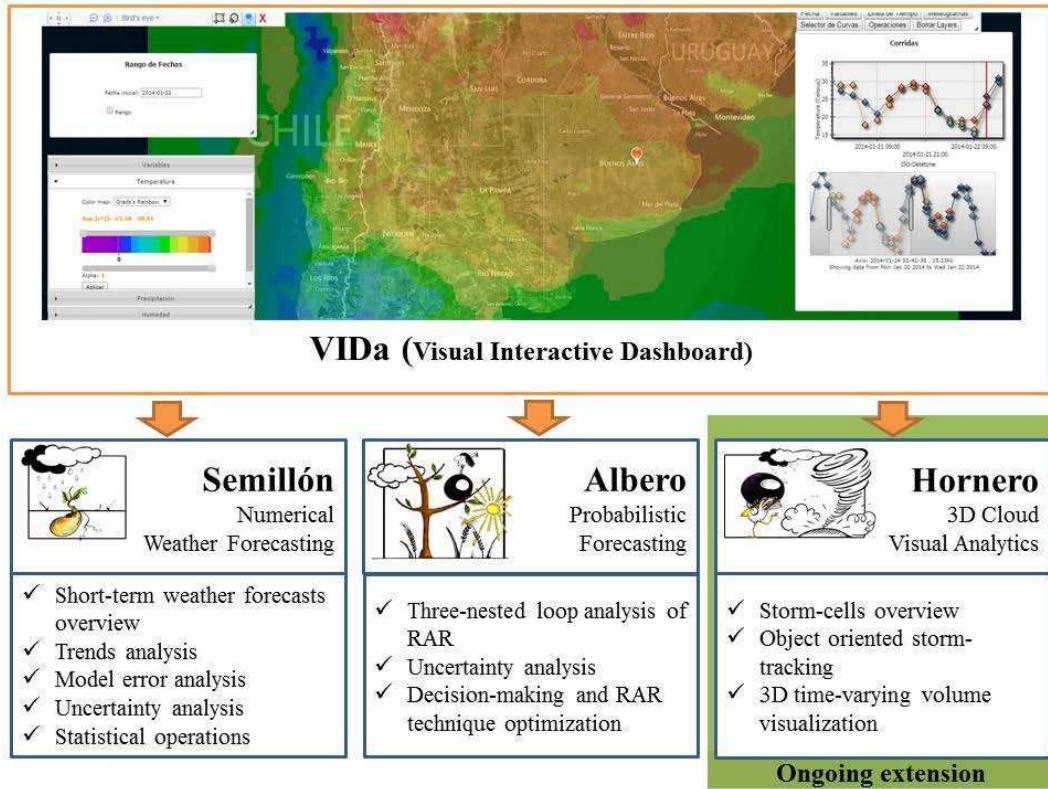


Figure 7: VIDa's component overview.

VIDa is extended to support different weather forecasting models and forecasters' needs as it is shown in Fig. 7. Currently, two models run on top of VIDa: a solution named Semillón that supports numerical weather forecasting, and another solution named Albero that works with probabilistic weather forecasting. Ongoing work will also enable the 3D visual exploration of clouds structures and radar data. This endeavor is embraced by an ongoing project named Hornero.

This chapter presents the methodology and decision-making process for the design of VIDa and its components: Semillón, Albero and Hornero. Section 2.2 introduces our solution, VIDa. It presents characteristics of the domain of application, a description of the tasks' workflow and motivations for its components: Semillón, Albero and Hornero. Section 2.4 discusses the design and implementation details of VIDa and how the components are integrated in a Multiple Coordinated Views (MCV) system.

2.2 Components

Shneiderman described the visual information seeking mantra as: “1. overview, 2. zoom and filter, 3. details-on-demand” [108]. Afterward, Keim extended it to a visual analytics mantra that emphasizes the importance of including automatic and semi-automatic analysis: “Analyze first, show the important, zoom, filter and analyze further, details on demand” [55]. The design of our solution takes advantage of the benefits of both mantras and combines them in a solution that provides overviews, zooms and filters, and also includes post-processing of data, automatic processes, and further details.

2.2.1 Semillón

The component Semillón assists users in the visual analysis of short-term weather forecasts. Our collaborators calculate short-term weather forecasts using the Weather Research and Forecasting (WRF) model [101] adapted for their geographical region. The WRF model is an open-source mesoscale NWP model built on the basis of collaborative efforts of several institutions in the United States of America, as well as contributions from researchers around the world. Its optimum configuration and performance highly depend on the specific application. It includes different parameters such as geographical area, time of the year, and local forecast errors of the regional model [102].

Forecasters need to provide early predictions about trends, specific phenomena and analysis of model errors. Also, a very important aspect for them is to make available and share their experimental short-term weather forecasts using the Internet. Forecasters need a tool that provides them with straightforward visual mechanisms to explore a set of forecasts, with an agile interface that allows them to perform different operations in time and space. These needs are not altogether covered by currently available systems. Applications such as GrADS, VAPOR, and UV-CDAT, require specific hardware, user training, and programming skills that make them unsuitable for operational needs. Our solution gives the users easy access to complete

cycles of short-term weather forecasts. It provides them with a visual overview of multiple forecast cycles at the same time. By using our application, forecasters can evaluate the complete panorama and focus on specific forecast operations and analyses, without the requirement of programming and complex settings.

Tasks workflow

In order to design a visual analytics tool that reflects domain experts' needs, we have analyzed and integrated our solution to their tasks workflow for short-term weather forecasting. The CIMA's current workflow comprises four main activities as shown in Fig. 8.

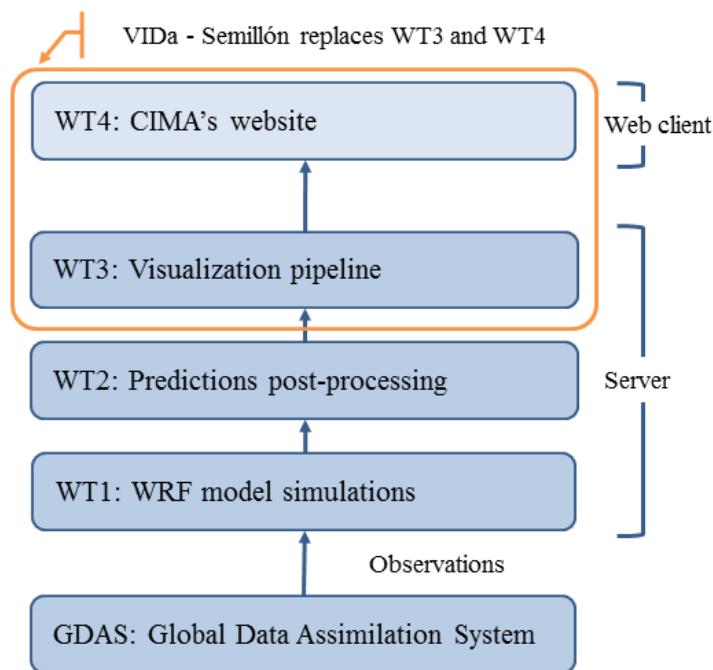


Figure 8: Semillón task workflow. The new visualization pipeline is integrated to our collaborators' workflow at tasks WT3 and WT4. GDAS provides observational data for the workflow.

First task (WT1). It consists in the generation of cycles of numerical weather forecasts twice a day, for the next 48-hours.

Second task (WT2). After the generation of forecasts, simulation outputs are post-processed to derive new atmospheric variables from their results.

Third task (WT3). In this task, the specialists create new visualizations that are generally 2D plots.

Fourth task (WT4). Finally, visualizations are exported as images and presented on a website using 2D maps and plots (see Fig. 9).

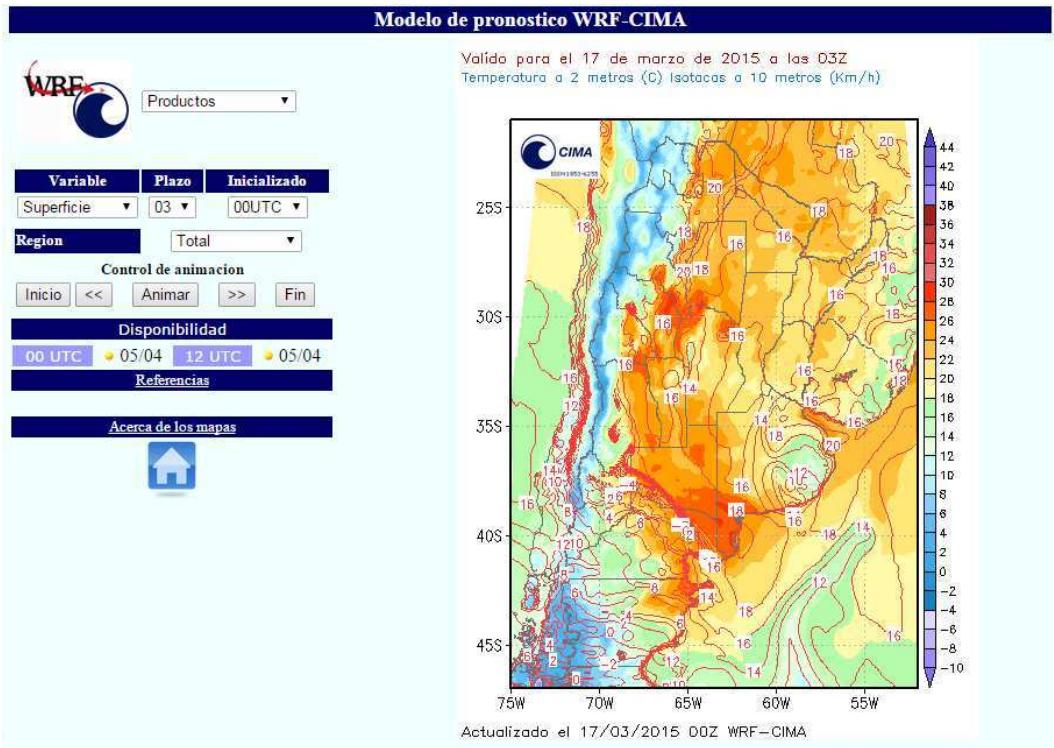


Figure 9: Screenshot of the current CIMA's website.

A vast amount of information is presented to the users in their website: http://wrf.cima.fcen.uba.ar/htm/wrf_out/index_mapas.html, however we have realized that their visualizations have certain drawbacks. For example, they visualize the information using a static interface that makes interaction difficult. Additionally, they display different views that cannot be linked or compared. We have modified their current workflow by integrating our new visualization pipeline and visual analytics solution, VIDa and its component Semillón, into tasks WT3 and WT4 (see Fig. 8).

Semillón is a component of VIDa designed specifically for operational weather forecasting. It is implemented as a client-server system with an interactive web interface that displays a full screen webmap and several linked-views. Multiple views can be visualized or hidden by using a menu, as shown in Fig. 10. These views include: date selector, parameter selector, minimap timeline, mapview, mapview toolkit with spatial filters, meteogram, curve-pattern selector and a statistical and mathematical Operations tool.

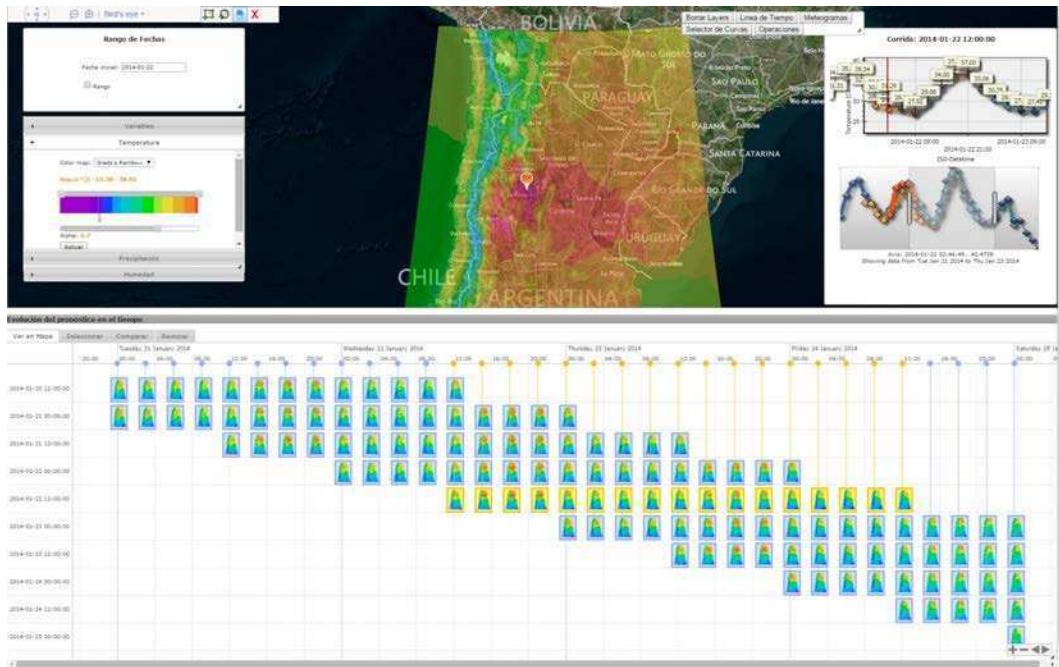


Figure 10: Screenshot of Semillón for the analysis of short-term weather forecasts.

Semillón contributions

Our solution provides visualizations and interactive mechanisms to assist users in the identification of weather trends and the visual analysis of model behaviors. The system supports 2D scalar-field grids of different meteorological variables such as temperature, pressure, and humidity. It enables the analysis and comparison of multiple NWP forecasts that describe the state of the atmosphere.

The main contributions of this work can be enumerated as follows:

- **A timeline view.** The timeline shows miniature 2D geo-referenced projections of a given meteorological variable, which are referred to as “minimaps”. This view shows an overview of the complete cycle of short-term weather forecasts. The timeline acts as a control panel of the visualization system. It shows a synthesized overview of various weather forecast cycles, each of them covering 48-hour periods. The overview comprises simultaneous views of forecasts initialized at different days/hours and targeted at different days/hours within the timeline window.
- **An Operations tool.** We present a flexible design to operate with multiple forecasts. The user can select two or more forecasts and apply different mathematical operations such as addition and subtraction between 2D scalar-field forecasts, or the computation of the mean and standard deviation over a group of forecasts. The results of the applied operations are displayed as a new layer on the mapview.
- **A novel curve-pattern selector tool.** This tool allows users to perform advanced visual analysis of multiple 2D scalar-fields, while avoiding the perceptual drawbacks of superimposing multiple colormaps. By means of this tool, users can sketch meaningful pattern behaviors, save them to be used anytime, and classify different phenomena, for example a cold front in a set of forecasts according to the defined patterns. Our tool is especially useful for quick analysis of weather forecasts in cases where the forecaster needs to identify specific data behavior.

Linked-views

The dialog between the user and the application begins when the user selects a meteorological variable and a given date and time to analyze. The system presents an overview of all computed 2D scalar-fields for the selected variable and date by means of the minimap timeline. The timeline

provides a visual overview that the user can employ to identify interesting forecasts where a possible trend or specific phenomena might be occurring. The minimap timeline is designed using the concept of “small multiples” [124]. Minimaps are connected with the mapview and the meteogram as it is shown in the Fig. 11.

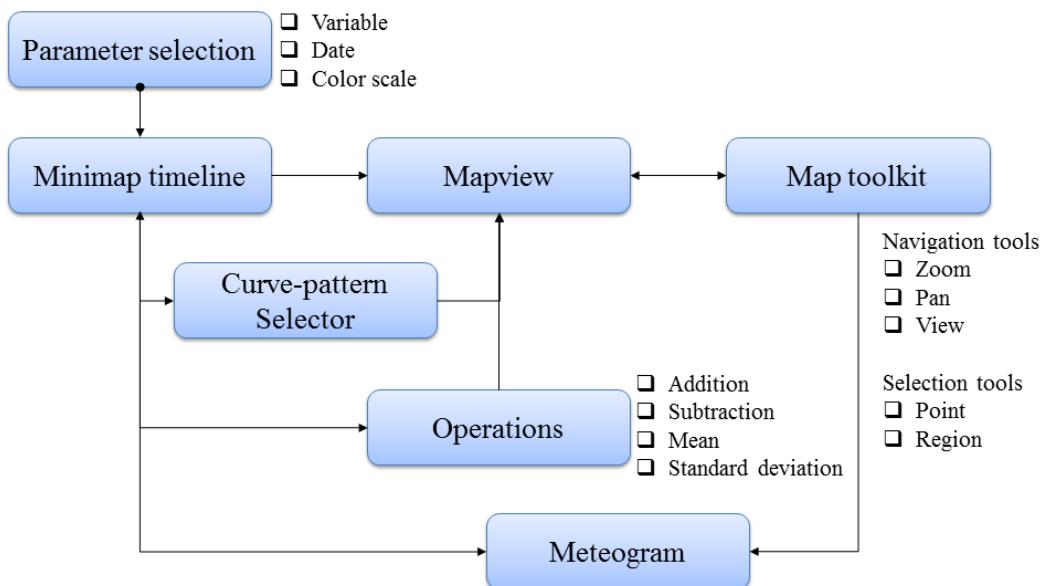


Figure 11: Semillón linked-views.

The user can select a minimap and then visualize the associated 2D scalar-field projection in an integrated mapview to analyze in detail the geographical distribution of a specific forecast. Also, the user can select two or more minimaps and perform different operations with the forecasts to aggregate the information or to emphasize some feature detected in the data. The user can visualize the results on the mapview, select different zoom levels and apply spatial and temporal filters using the mapview toolkit. This functionality is especially useful for forecasters to study a particular region or point on the mapview as part of the analysis of some specific meteorological event. The mapview controls allow the user to zoom in and out and view the forecast data in different resolutions. The user can also select a point or region on the mapview and visualize a meteogram with detailed information about the multiple forecasts initialized at different times. The

meteogram displays the temporal evolution of a variable in a specific point or region. Each run in the meteogram is linked with the same run in the minimap timeline and viceversa. It allows the user to analyze the influence of different runs in a given geolocated point or region. The minimap timeline acts as a pivot between multiple runs and forecasts. The Operations tool is linked with the minimap timeline and the mapview.

Furthermore, we introduce a novel curve-pattern selector that makes use of domain expert knowledge and semi-automatic analysis for the detection of particular local weather phenomena (see Fig. 12). It consists of an intuitive visual tool for pattern sketching and classification, a set of configuration options, and an automated technique to perform the classification of curve-patterns. Although our visual components are similar to others used in previous work, the selection and combination of them as presented in the curve-pattern selector provide a novel and flexible tool for meteorologists when utilized for weather forecasting. Using the curve-pattern selector the user can sketch meaningful patterns in a 2D coordinates graph. Once the curves corresponding to patterns are sketched, the system provides the users with auto-suggestion of a subset of trends identified using the curve-pattern selector. The suggestion of different trends allows the user to make a choice based on her own experience and knowledge. The curve-pattern selector is connected with the minimap timeline and the mapview. The user can select a set of minimaps and the selection is linked and visualized as time steps in the horizontal axis of the sketching canvas. The curve-pattern classification is done via a semi-automatic process. Results of the classification are then shown on the mapview.

Further analysis can be done in a iterative process using the multiple views and features. The analysis of combined results coming from multiple visual tasks can connect different viewpoints of a given phenomenon and provide a complete visual analysis of short-term weather forecasting.

Minimap timeline

The newly introduced minimap timeline is a key characteristic of our application given its capability to show a synthesized overview of complete 48-

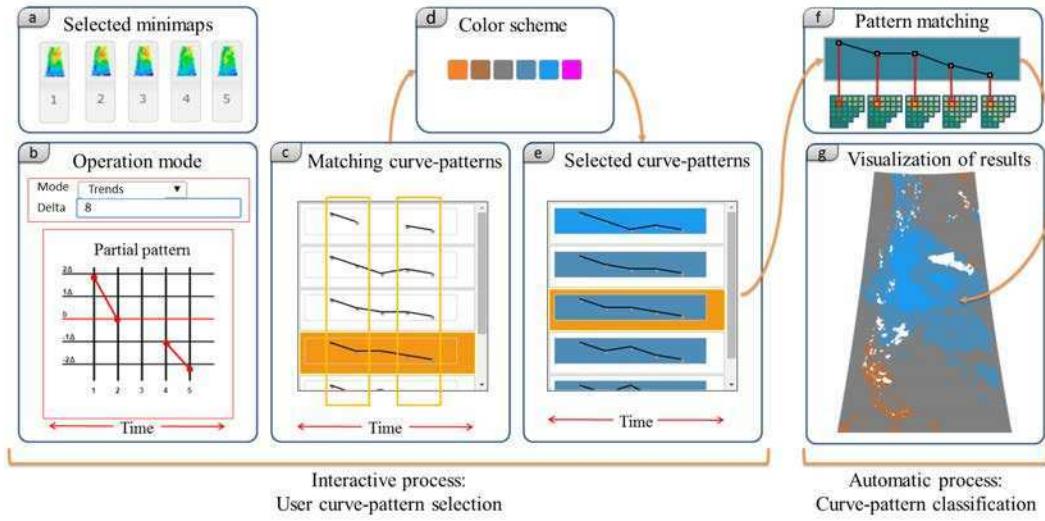


Figure 12: Curve-pattern analysis process: (a) The selected forecasts that predict 2D scalar-fields for a given date and time. (b) The curve drawing allows the user to sketch the desired patterns. (c) The selectable list of possible patterns shows the curve matches. (d) The selected color scheme. (e) The selected curves show the curve-patterns chosen by the user and its associated colors. (f) Pattern matching process for each pixel of the selected minimaps. (g) Visualization of the curve-pattern classification.

hour cycles of short-term weather-forecasts (see Fig. 13). Forecasters need to navigate the timeline easily from the current prediction through future predictions, as well as past events. The primary use of the minimap timeline is to provide an overview of multiple forecasts. Our domain experts stated that the minimap timeline is a very useful tool for selecting multiple forecast runs. Currently, salient differences in the color of the small geo-referenced projections of the 2D scalar-fields constitute an initial aid for the user in the identification of interesting events that can be later confirmed by using other components provided by Semillón.

Each minimap represents a 2D scalar-field of a forecasted meteorological variable and it is depicted as a geo-referenced 2D colormap. The colormap is a Universal Transverse Mercator (UTM) grid projection of the 2D scalar-field that is adjusted to the pixel coordinates of the webmap [75].

The timeline groups the minimaps by date and hour of each run. It shows the minimaps corresponding to all the calculated forecasts for a particular date (see Fig. 13). Using the horizontal axis, domain experts can analyze the temporal evolution of forecasts for a particular variable. Each horizontal line shows the next 48-hour predictions of a given variable in a given run. A new run takes place everyday at 12:00 AM and 12:00 PM Coordinated Universal Time (UTC). Using the vertical axis, they can analyze differences among multiple runs of predictions that correspond to a given date and time. This is very useful for analyzing the performance of the NWP model through different days. The timeline also shows the first prediction at lead 0 for a given date (lead 0 means that the forecast predicts the weather with 0 hours of anticipation). This prediction is known as the “diagnostic” and is the best representation of the state of the atmosphere at lead 0. The “diagnostic” is created with observational data and it is utilized for forecast verification.

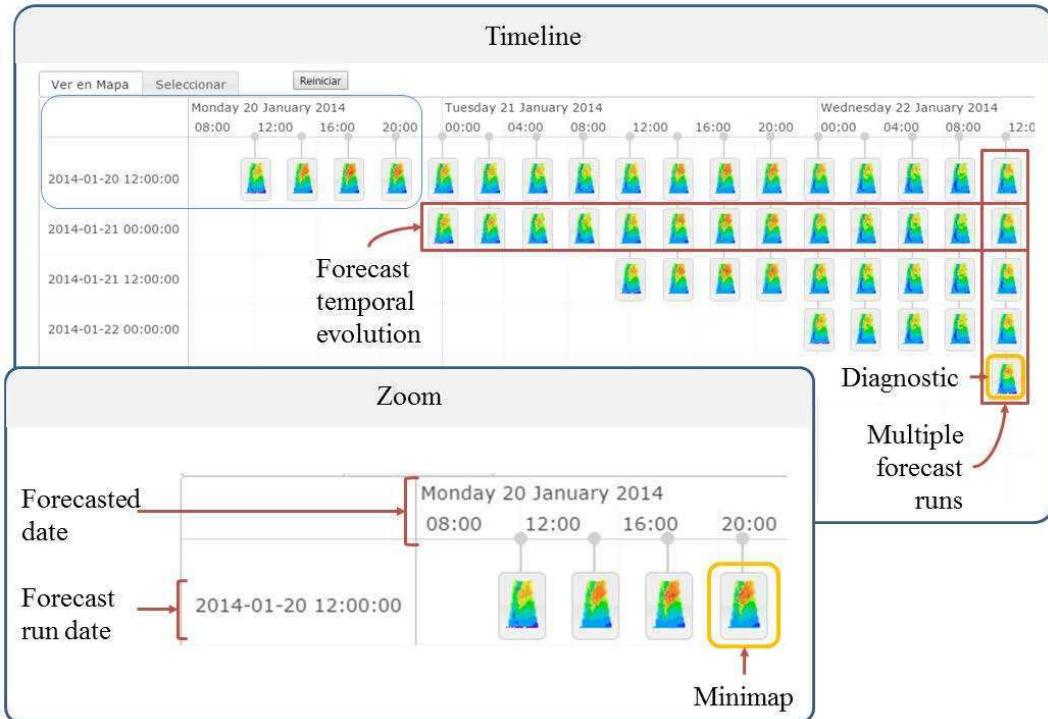


Figure 13: Minimap timeline overview. It shows multiple forecast runs and predictions, depicting the temporal behavior of a given variable.

Mapview

The mapview allows users to visualize a given forecast, or result of an operation or comparison at different levels of detail. Forecasters can focus on specific regions, apply spatial filters and analyze linked information. The mapview integrates the information of forecasted 2D scalar-fields, point-based geolocated information and spatial filters over the map. The user can visualize one or more 2D scalar-fields corresponding to a meteorological variable as overlays on the map. The mapview zoom level ranges from zero to twenty-three, depending on the geographical location. For zoom level zero, the entire world map fits into an image of 512x512 pixels, but for each increasing zoom level the pixel space expands by a factor of two in both axes.

Forecast Operations tool

The forecast Operations tool combined with the timeline allows for an extensive analysis of the short-term weather forecasts. It enables users to perform comparisons among an arbitrary number of forecasts. By using the timeline, multiple forecasts can be selected in the horizontal axis as well as in the vertical axis. The operation is applied to the selected forecasts. Operations such as subtraction, mean and standard deviation give useful information to assess forecast uncertainty. The subtraction between forecasts allows the user to detect significant changes in forecasts initialized at different times. By analyzing the subtraction results, the user can detect where and how significant the spatial differences are. Other operations such as the computation of the mean among a group of forecasts can provide a more accurate estimation of the future evolution of the atmospheric variables compared to the individual forecasts. An extensively used method for the estimation of forecast uncertainty is to compute the standard deviation among a group of forecasts predicted for the same date and time. Finally, the addition of multiple forecasts can be useful for the analysis of variables that can be accumulated, such as precipitation. In this case, the users can interactively select a consecutive set of time steps that best suits their needs.

Operation	n forecasts F_i	Formula on each point (x, y)
Addition	two or more	$A(x, y) = \sum_{i=1}^n F_i(x, y)$
Subtraction	two	$S(x, y) = F_1(x, y) - F_2(x, y)$
Mean	two or more	$\mu(x, y) = \frac{A(x, y)}{n}$
Standard Deviation	two or more	$\sigma(x, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_i(x, y) - \mu(x, y))^2}$

Table 3: Formulas of the Operations tool.

Mapview toolkit

The user can select one or more spatial filters from a toolbox located on the mapview. These filters allow the user to select a location or a region. This information is used to filter the spatial data and to retrieve information for a given geolocated point or region on the map. The mapview also provides integrated functionality that allows the user to zoom in and out and to change the view mode.

Meteogram

A linked meteogram visualizes multiple runs over time for a given meteorological variable. This view shows mean values for a given region or for a specific location selected using the map toolkit (see Fig. 14). The information is linked and highlighted across the different views.

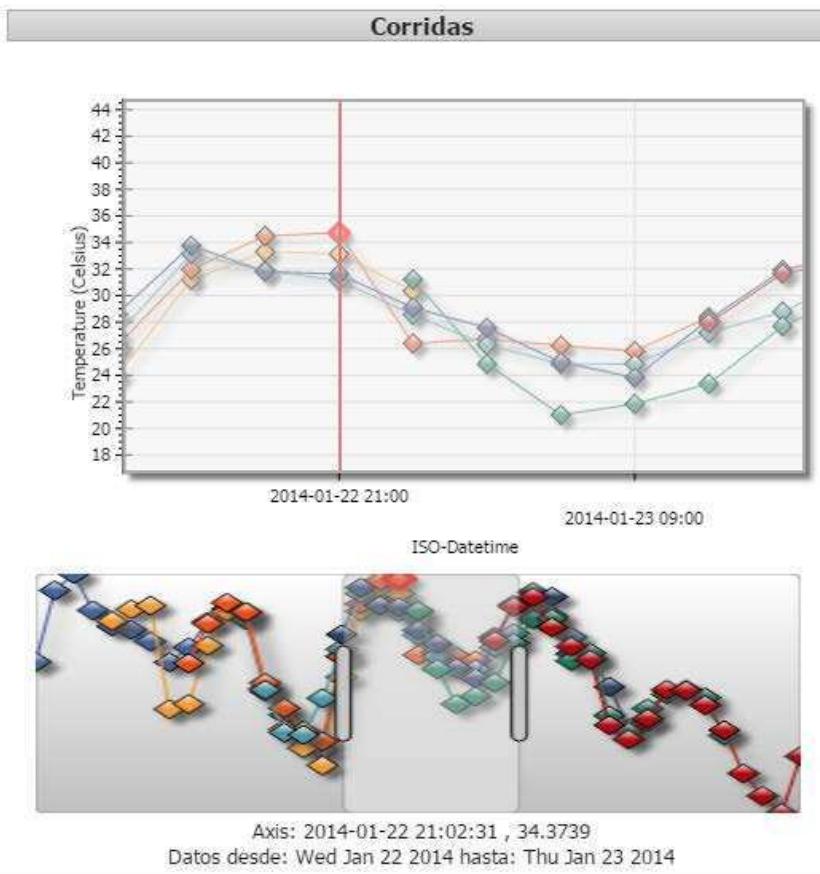


Figure 14: Meteogram displaying multiple runs over time for a specific geolocated point in the mapview.

Curve-pattern selector

Forecasters need tools for detecting and analyzing trends and anomalies, model errors and for assessing forecast uncertainty. Here, a “trend” is considered to be a general behavior that is expected to happen for a forecasted meteorological variable. For example, a continuous increase in temperature in a given region over a period of time is a possible expected behavior. On the contrary, an “anomaly” is considered as an unusual or atypical behavior of a forecasted variable. Forecasters need to quickly analyze their forecasts and estimate the effects of those phenomena in their model.

Our curve-pattern selector is especially helpful for this kind of analysis. The goal of the curve-pattern selector is the comparison and classification of multiple 2D scalar-fields in time and space simultaneously to identify trends and anomalies. A curve-pattern represents the qualitative temporal

evolution of a variable at a particular spatial location. Forecasters can specify patterns related to particular features of the data, for example, constant curves, increasing curves or decreasing curves.

By means of the curve-pattern selector, users can make interactive queries of prospective curve-patterns that could be meaningful to them, and classify and visualize the patterns on the mapview. For example, the forecaster can specify a curve-pattern that corresponds to a sudden drop on the values of a variable. In the case of temperature, it could indicate the presence of a cold front. We assume that this is possible, under the hypothesis that only a few of all possible combinations of curve-patterns are practically meaningful to the users. Those curve-patterns are subsequently used to classify the 2D scalar-fields. This classification allows the user to visualize specific characteristics such as trends or outliers. The identification of specific features in the data depends highly on the nature of the meteorological variable and on the kind of analysis the user wants to apply. For instance, a case can be applied to the temperature variable, where the user chooses increasing curve-patterns, decreasing curve-patterns and a constant curve-pattern (within a certain threshold), to describe different trends.

Curve-pattern visual query

Employing the minimap timeline, the user selects the time steps to be included in the analysis (see Fig. 12a) and span one or more segments. Then, the user sketches a pattern with the mouse. The pattern sketch can be partial in order to match different curves. Figure 15 shows examples of different curve-pattern sketches. The horizontal axis of the chart represents each of the selected time steps. The vertical axis represents the change in magnitude of the selected variable in a scale of “ Δ ”. Δ is a global parameter specified by the user. This parameter is used to measure the distance between the sketched curve-pattern and the temporal curves on each geographic location of the 2D scalar-field associated to the selected time steps.

The curve-pattern selector supports the sketching of segments and points that can start at any time step. They do not need to begin at the first time step or to sketch a single segment. The creation of a curve-pattern is straight-

forward. The user clicks on the desired time step in the drawing canvas as a starting point. She can draw a continuous line by clicking on other time steps, or break the curve with a mouse right-click. The user can create several partial line segments or points that jointly compose the final curve-pattern.

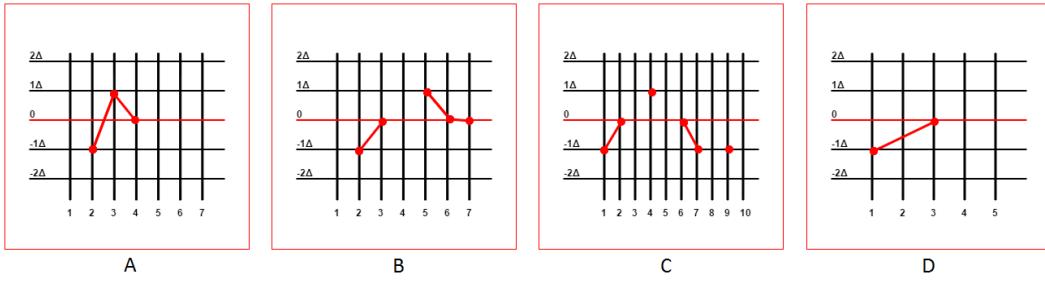


Figure 15: Click-based sketching of different connected and disjoint segments. Sketches A and B show connected and disjoint segments. Sketch C shows an example with single point segments. Sketch D shows a segment with an intermediate time step linearly interpolated.

Once the curve-pattern is sketched, the curve-pattern selector displays all possible curves that closely match the drawn curve, including the pattern itself as the first element of the list. The application establishes the maximum number of possible curves to visualize depending on performance and usability factors. From all possible combinations the user selects the curves that are meaningful for her (Fig. 12e). The tool allows the user to associate a color with each curve-pattern or a group of curve-patterns. The color is chosen from a particular color scheme (Fig. 12d). The selected curves become the curve-patterns that will be used to perform the classification (Fig. 12f). The results (Fig. 12g) are visualized as a new layer in the mapview.

The curve-pattern analysis has proven to be very useful for our domain experts to address at least two important goals; they are, weather trends analysis and forecast verification, which will be extensively described in chapter [Case studies](#).

Glatter et al. [33] technique is closely related to our approach. They developed a text search language using Flex and Bison for specifying temporal patterns. They allowed the specification of qualitative queries of temporal patterns that can be partially defined by means of regular expressions. We use some concepts from their query language and adapt them to our visual analytics tool. Their work and ours have similarities, but they differ with respect to user interactions. They present a very complete technique using string query patterns, while our work uses a visual interface for creating the queries. The curve-pattern selector functionality presented in this work are a sub-set of the pattern search language of Glatter et al., but complete enough to satisfy the needs of our domain experts. Their technique allows more powerful and flexible expressions (at the expense of increased complexity), but our approach does not require those features at the moment. Our selector allows the forecaster to draw several disjointed segments of the pattern (by using the mouse right-button to interrupt a segment). It enables any time-step that was not drawn to accept any value of the variable, thus, we include the question mark "?" symbol of Glatter et al. However, we do not include the asterisk "*" symbol (repetitive patterns).

Curve-pattern classification process

The classification based on curve-patterns can be helpful for detecting specific features in the temporal evolution of meteorological variables represented as 2D scalar-fields. Forecasters can arrange several 2D scalar-fields and analyze their evolution or changes in the temporal domain. Each 2D scalar-field of a meteorological variable is stored as a matrix with a resolution of $N * M$, longitude and latitude respectively. We define a curve for each position (*long*, *lat*) across the k matrices, where k is the number of scalar-fields. The curve represents how the meteorological variable changes in the temporal domain and it is composed of $k - 1$ line segments. Since each curve is defined in a given geolocated position, the set of all these curves describes the temporal evolution of a variable in a complete geolocated region. The curve-patterns are compared against the 2D scalar-field curves using a specific metric. The metric used for the classification is defined by the mode

of operation. The curve-pattern classification algorithm works in two different modes: temporal evolution and forecast verification.

The temporal evolution mode can be beneficial to detect significant trends or changes in the behavior of a variable (e.g., sudden drops in temperature associated to weather fronts). In order to define the metric, a global parameter Δ is introduced. The parameter Δ measures the distance between a curve-pattern segment and a curve segment as it is shown in Fig. 16a.

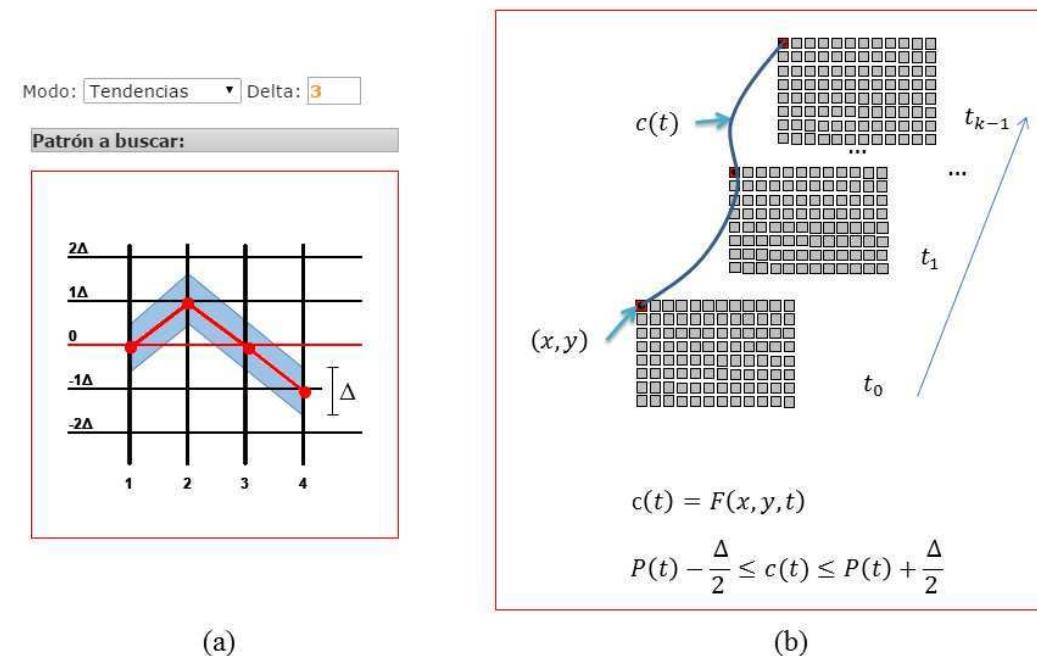


Figure 16: Curve-Pattern Δ metric: (a) a light blue highlight has been added (not visible in the application) with the values that a variable can assume in order to match the pattern. Δ value has an effect on the tolerance acceptance for the variable value in order to match the curve-pattern. (b) Function $c(t)$ is defined as a temporal function of the forecast $F(x, y, t)$ at the position (x, y) at time t and it represents a temporal curve. $P(t)$ is the temporal curve-pattern defined by the user. The algorithm selects the curves in each (x, y) that match the inequality equation.

A curve-pattern and a curve will match if each of the segments differs not more than $\Delta/2$, as is shown in Fig. 16b. Parameter Δ acts as a threshold on the Y-axis. It is also related with the slope of the line segment because

it establishes the Y-axis scale. Segments of the curve-pattern that were not specified will match with any curve.

In the forecast verification mode, forecasts are compared against the “diagnostic”, a 2D scalar-field created on the basis of available observations and calculated at lead 0. In this mode, the curve-pattern analysis can be helpful for the identification of regions with different error growth rates. It helps forecasters to gain experience on the behavior of forecast errors associated with the NWP model under different weather conditions. There are several measures to perform forecast verification [68], for example the “bias”. The bias is defined as the difference between the diagnostic and the forecast. This is a measure of the forecast error commonly used in forecast verification analysis [47]. For this mode of operation, the meteorologists use the absolute value of the bias (Eq. 2.1). The absolute value of the bias is calculated for each point i, j with $i \in [1, N]$ and $j \in [1, M]$ of a forecast grid F that has a dimension of $N * M$ and compared to the diagnostic D that has the same dimension of $N * M$.

$$|bias_{i,j}| = |F_{i,j} - D_{i,j}| \quad (2.1)$$

In both modes, the algorithm performs a classification on each point ($long, lat$) for the selected 2D scalar-field matrices and represents the results in a new matrix with the same dimension. To create a new color-code image with the results of the classification, the ($long, lat$) coordinates are transformed to pixel coordinates. In case that a curve matches one of the curve-patterns, the color of this pattern is mapped to the pixel coordinate corresponding to the ($long, lat$) of the point. Otherwise, the pixel is rendered with a transparent color. The results of this operation are visualized as a new layer on the mapview.

2.2.2 Albero

Weather forecasting based on NWP models has uncertainties arising from inaccuracies of the initial and boundary conditions and the imperfections of the model itself. To address the problem of deterministic forecasts, many

runs of the prediction model can be performed simultaneously. Both the initial conditions and settings of the physical parameterization of the model are perturbed, obtaining an ensemble of forecasts. However, the ensemble of forecasts represents only a sub-set of the possible outcomes of the NWP model, underestimating the degree of uncertainty in the prediction of the future atmospheric state [50, 91].

Hamill and Whitaker [39] introduced the “Reforecast Analog Regression” (RAR) technique to overcome the limitations of the ensemble and the deterministic forecast models. The algorithm uses a large database of weather forecasts generated by means of a specific prediction model [84] and an associated database of observations [20]. It retrieves historical forecasts that are similar to the current numerical forecast for a particular geographic region, referred to as “analogs”. Then, observations corresponding to those analogs are used to estimate a probability density function for the current forecast period. The results can be expressed as probabilistic forecasts. Aldeco et al. [2] verified experimentally the performance of the technique for Argentina. Probabilistic forecasts computed using the RAR technique allow meteorologists to have a detailed measure of uncertainty [39].

Our collaborators use retrospective forecasts (shortly reforecasts) generated by the Global Forecast System (GFS) [84] to create their probabilistic forecasts (see Fig. 17). Reforecasts are global forecasts that estimate the evolution of the atmosphere from the initial forecasted date for a maximum lead time of 16 days. In their current tasks workflow, they have to run the RAR algorithm repetitively for each threshold, lead time and additional configuration if they need to perform further analysis, as it is shown in Fig. 18).

We propose a novel interactive workflow (see Fig. 19) that enables the analysis of internal aspects of the RAR algorithm, such as the analysis of multiple probabilistic forecasts, ensemble of analogs and observations and statistical summaries. Our new workflow enables domain experts to optimize the RAR technique and forecasters to improve their decision-making. It acts in each step of the technique, visualizing intermediate results and unveiling new insights and information. Another important and novel feature is the capability to evaluate the sensitivity of results with respect to several

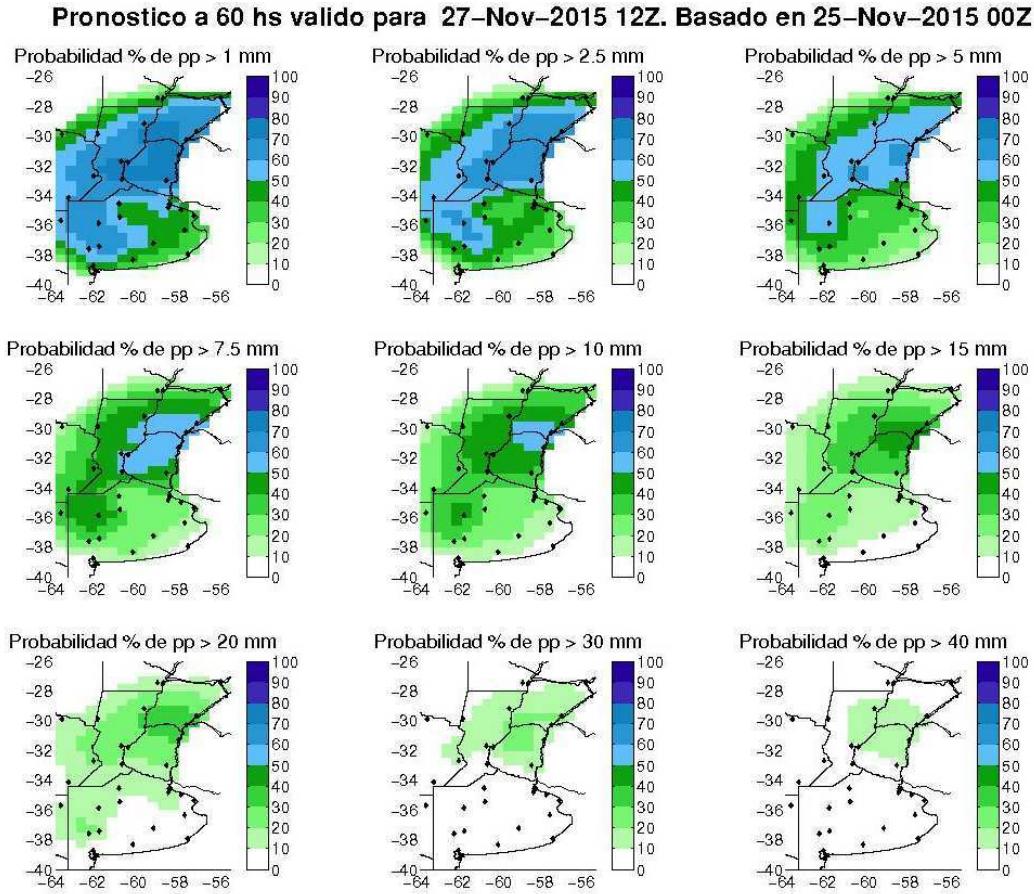


Figure 17: Visualization outputs of precipitation probabilistic forecasts at the National Weather Services, before using Albero [2].

parameters of the algorithm. Among the parameters that can be tuned are the size of the analog database, the number of selected analogs, the size of the spatial window used to select the analogs, the variables used to define the analogs and the metric used for the analog search.

To the best of our knowledge, our work is the first interactive workflow and visual analytics solution of this kind for probabilistic forecasts based on the RAR technique. We replaced the current automatic Hamill and Whitaker workflow implemented by our domain experts with a new interactive workflow (see Fig. 20). It allows forecasters to gather new insights into the probabilistic forecasts, past events and systematic errors.

We provide an efficient and interactive approach that supports forecasters to detect extreme events in advance such as flooding and severe precipi-

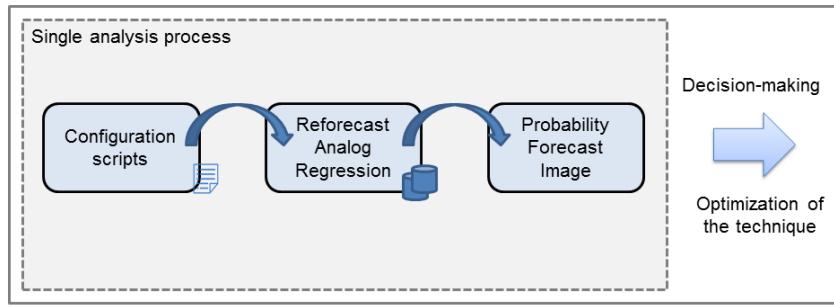


Figure 18: Probabilistic forecasts: current automated workflow.

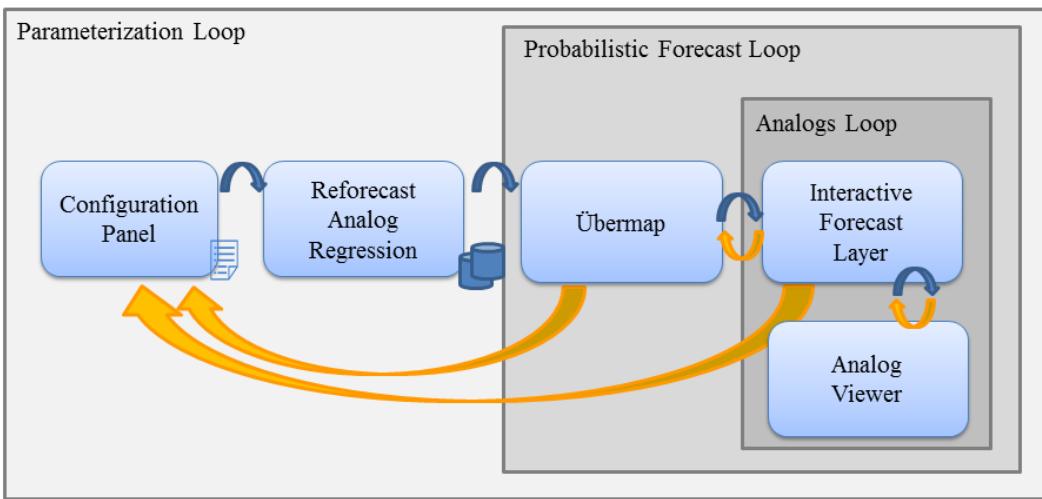


Figure 19: Albero: a new interactive workflow for probabilistic forecasting based on Analogs.

tation. In addition, it meteorologists to detect extreme values and systematic errors to improve their models.

Our solution, we named “Albero” integrates visualization techniques, such as small multiples, maps, charts and histograms, as coordinated and linked-views to facilitate the meteorologist work (see Fig. 19). We introduce a novel interactive workflow (see Fig. 19) that enables the analysis of new aspects of the RAR technique and data, such as the analysis of multiple probabilistic forecasts, analogs and observations, and statistical summaries. The “visual analytics loop” proposed by Sacha et al. [103] explored the cyclical mechanisms to convey visual knowledge. Our new interactive workflow breaks-down the automated RAR method into three principal analysis

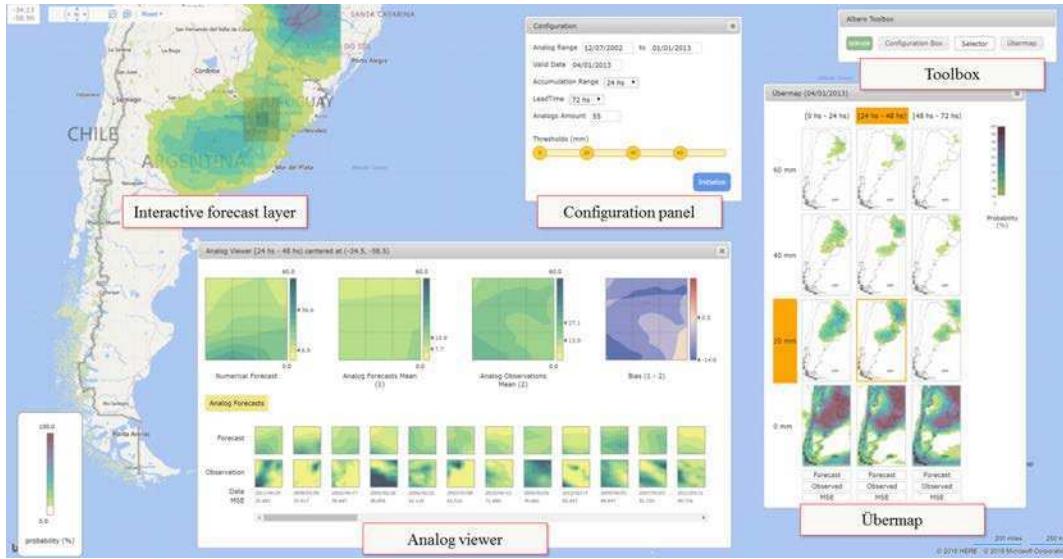


Figure 20: Screen-shot of Albero for the analysis of Probabilistic forecasts. Different linked-views are coordinated in a fluent dialog between the user and the application to assist the user in the decision making process and optimization of the RAR technique.

loops. Each loop comprises an interactive dialog between the user and different linked-views provided by Albero.

This work targets at least two different types of users: meteorologists who work in operational weather forecasting (forecasters) and meteorologists who are researchers and work on numerical weather prediction (researchers). Our approach supports forecasters to detect extreme events in advance such as flooding, severe precipitation, etc., and researchers to detect systematic errors to improve their model.

Parameterization Loop. This iterative process allows different types of users to set temporal and spatial parameters of the Hamill and Whitaker algorithm. We have identified two principal types of users: researchers and forecasters. Forecasters need to set parameters such as the lead time, accumulation range of a meteorological variable and spatial region of interest. Researchers need to optimize parameters such as historical database ranges, quantity of analogs and weight of different variables used in the algorithm.

A flexible configuration panel is provided where users can refine those parameters depending on their needs.

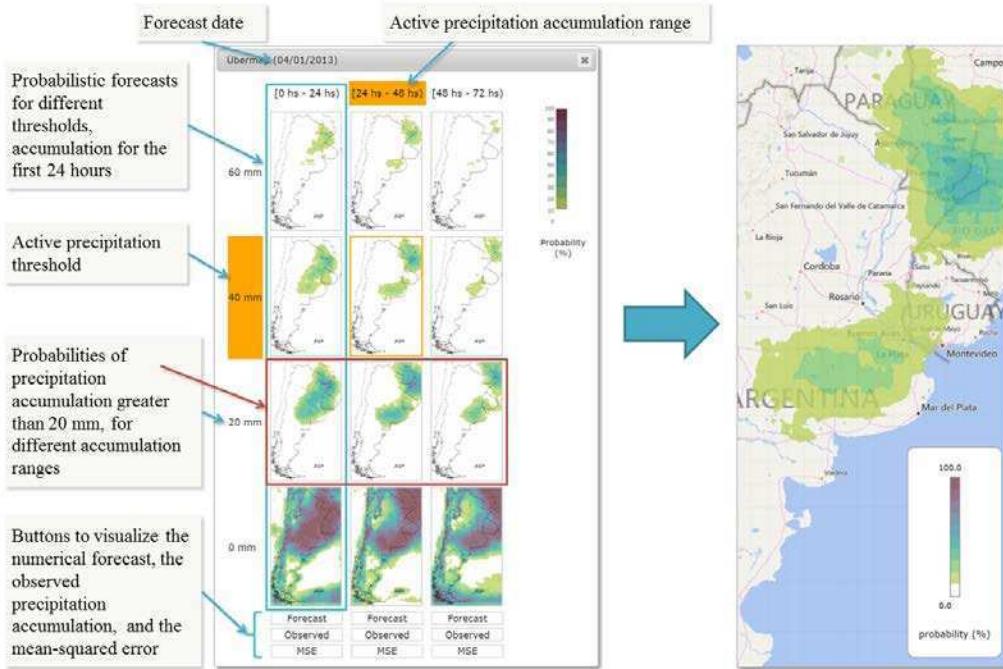


Figure 21: Übermap: overview visualization displaying probabilistic forecasts for various thresholds of precipitation accumulation and accumulation ranges. It can also display the deterministic forecast, the observations and the Mean Squared-Error (MSE), by using the corresponding buttons.

Probabilistic Forecast Loop. Users need to analyze the probabilistic forecasts with different levels of detail and compare them both spatially and temporally. In this loop users can analyze a set of probabilistic forecasts for a given meteoreological variable, summarize information about observations and evaluate their uncertainty. We defined an overview visualization named “Übermap”(see Fig. 21) and a detailed view named Interactive Forecast Layer (see Fig. 22). The Interactive Forecast Layer allows the user to visualize and explore the spatial sub-regions used by the RAR to construct the probabilistic forecasts. The tool allows a seamless exploration of the aforementioned sub-regions as well as their associated spatial errors.

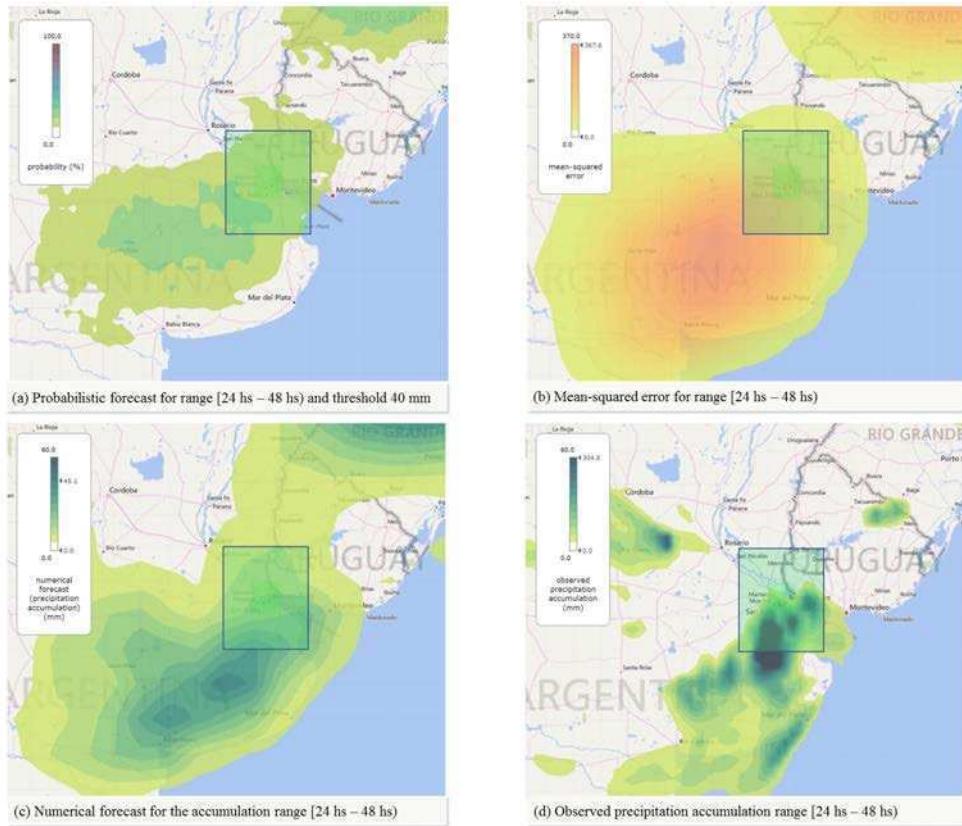


Figure 22: Interactive Forecast Layer showing a close-up of: (a) probabilistic forecast colormap showing an interactive grid divided in the sub-regions used by the RAR, (b) MSE spatial distribution of the technique, (c) numerical forecast, (d) observations. By clicking in one of the sub-regions the user has access to detailed information of the ensemble of analogs and observations used in the computation of the forecast for the given sub-region.

Analogs Loop . This loop examines intermediate steps of the Hamill and Whitaker algorithm. It has a break-down view, named the Analog Viewer (see Fig. 23), that allows users to access information about ensembles of forecasts and observations for the selected analogs' dates and algorithm configuration.

The Analog Viewer also shows different statistical summaries of the analogs and the observations. It allows users to detect extreme situations such as intense precipitation or flooding. The analysis of this detailed information provides experts with new insights into the RAR technique and help to improve their decision-making processes [118].

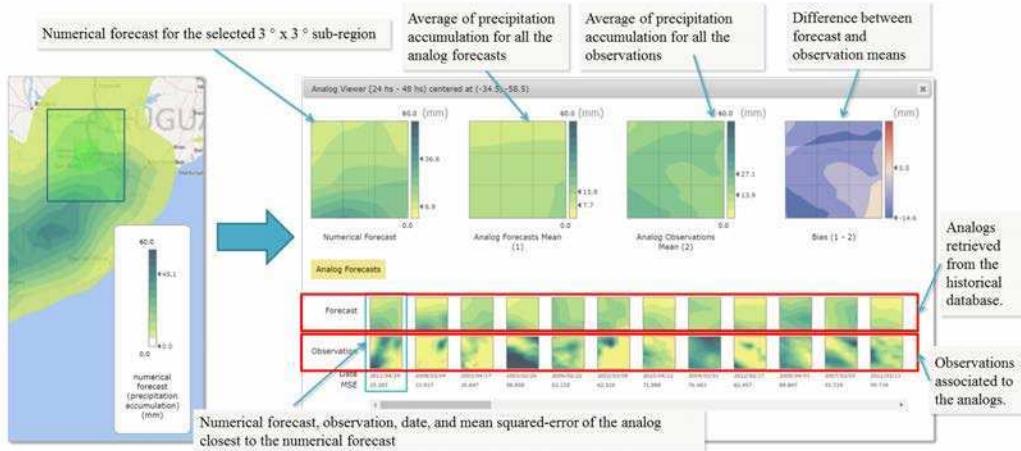


Figure 23: Analog Viewer description for the case of a specific selected region on the first accumulation range [00 hs - 24 hs).

Color scheme design

For the visualization design of the color schemes used in Albero, we have selected the HCL color model. We have chosen a single-hued HCL scale for precipitation accumulation and percentages and double-hued scales to represent the bias and Mean Squared-Error. Choosing appropriate colors palettes can strongly increase the usability of visualizations. Quantitative variables are best visualized by using single-hued color scales, using different levels of luminance to represent different levels of the variable. Monmonier [78] recommended using single-hued scales to facilitate interpretation without recurring to a legend at all the times and observed that a legend can improve the usefulness of a map but it can not make it efficient. A committee [10] was created in 1993 to study color perception and present guidelines. It offered recommendations on color usage in features commonly used in meteorology. One such recommendation was the use of green scales to indicate rain. Stauffer et.al [111] empathizes using HCL color model for meteorology, since it offers increased readability, guideline, attractiveness, and accessibility for people with visual constraints.

2.2.3 Hornero

We are working on an extension of VIDa for the 3D Web visualization of the forecasted development and growing of storm cells in time and space. We aim to answer questions such as:

- Where are the regions of intense precipitation for the following two hours?
- Where are the regions of hail formation for the following two hours?
- How is the associated uncertainty of the forecast?

2.3 Data sources

The CIMA institute works with the WRF model customized for South America [101]. The WRF model is implemented over a region centered at longitude 63°W and latitude 38°S , located in Argentina. Each simulation predicts several variables (usually more than 10) and more than 50 vertical levels. Among the meteorological variables are temperature, pressure and winds. The covered area is discretized into a regular grid, with a resolution of 149 by 299 points, in Lambert Conic Conformal projection (LCC) and an approximate distance between points of 15 km. As a result, a total of $149 \times 299 \times 50$ ($\text{long}, \text{lat}, \text{Z}$) data points per forecast and per variable are produced.

2.3.1 Semillón

For Semillón we only considered the value of the meteorological variables near the surface (i.e., as a single vertical level) even though the regional WRF model can provide results at different altitudes.

Every day, two 48-hour cycles of short-term forecasts run at time 00:00 UTC and at time 12:00 UTC in CIMA's servers. Each cycle generates 17 predictions (one every 3 hours) that are representations of the atmospheric state corresponding to different times in the future. Afterward an automatic task post-processes the data and generates new derived variables.

Then, in VIDA's server, an automatic Extract-Transform-Load (ETL) process is called. The ETL process extracts the information via File Transfer Protocol (FTP), transforms it and loads it in the system. This automatic process takes each simulation from the FTP server, extracts information about the date and time of the run and the time-stamp of each of the simulations, generates aggregated variables, loads the aggregated variables in a geospatial database and saves the new files in the file system. Afterward, another process is launched to generate the images. GPU computing is used to generate the images that will serve the Front-End.

2.3.2 Albero

For Albero we used reforecasts generated by the GFS. Reforecasts are global forecasts that estimate the evolution of the atmosphere from the initial forecasted date for a maximum lead time of 16 days. These forecasts include meteorological variables such as temperature and precipitation among others. They cover the entire globe with a horizontal resolution of 28 kilometers between grid points. The horizontal resolution drops to 70 kilometers between grid points for forecasts between one and two weeks lead time. The ensemble of forecasts consists of 11 members generated by slightly varying the initial conditions of the forecasts, giving place to multiple possible scenarios. We only use the mean of the ensemble to search for analogs.

This data-set is generated with the goal of building a retrospective database of forecasts using a specific NWP model. This allows for a detailed study of the systematic-biases of the model and the quantification of its uncertainty. The main motivation of this data-set is to develop a robust historical database to quantify the model uncertainty and in subsequent research steps to improve the quality of the model.

Another data-set required is the observations database. We use observational data from the CMORPH (Climate Prediction Center - NOAA) [20]. It produces global precipitation analyses at very high spatial and temporal resolution. They are based on satellite information using radiation emitted from the Earth and the atmosphere in the microwave frequency ranges to

estimate liquid water and ice content in clouds. The surface precipitation is inferred using this data. This information, combined with geostationary satellites data, is used to estimate the displacement of precipitation areas. Both estimates were verified to be valid for the region of interest by meteorologists from the National Weather Services [104].

2.4 Implementation

2.4.1 Architecture and technologies

Our solution is built as a multilayer system with a presentation layer that exposes a web Front-End, a Business layer that implements all the meteorological algorithms and a persistence layer based on a geospatial database that completes the server Back-End (see Fig. 24). The Front-End presents a visual interface implemented as a web application with multiple linked-views. Those views were developed using HTML5, JavaScript and jQuery technologies. TeeChart charting components (<http://www.steema.com/>) were used to implement the meteogram and vis.js (<http://visjs.org/>) was utilized for the timeline.

The Business layer uses the GPU computing engine and core libraries for processing weather forecast information. It was developed using C++ and OpenCL technologies for the General-Purpose computation on Graphics Processing Units (GPGPU). Bing Maps services were used for the mapview component. The persistance layer is based on a PostgreSQL geospatial database management system (DBMS) and automatic processes to import, transform and store the information provided by the WRF model.

2.4.2 GPU visualization pipeline

There is an intrinsic connection among large-data visualizations and High Performance Computing (HPC). Visualization of large amounts of data requires huge processing capabilities, especially when interactivity is required. In this part of the work, we focus on meteorological variables that are repre-

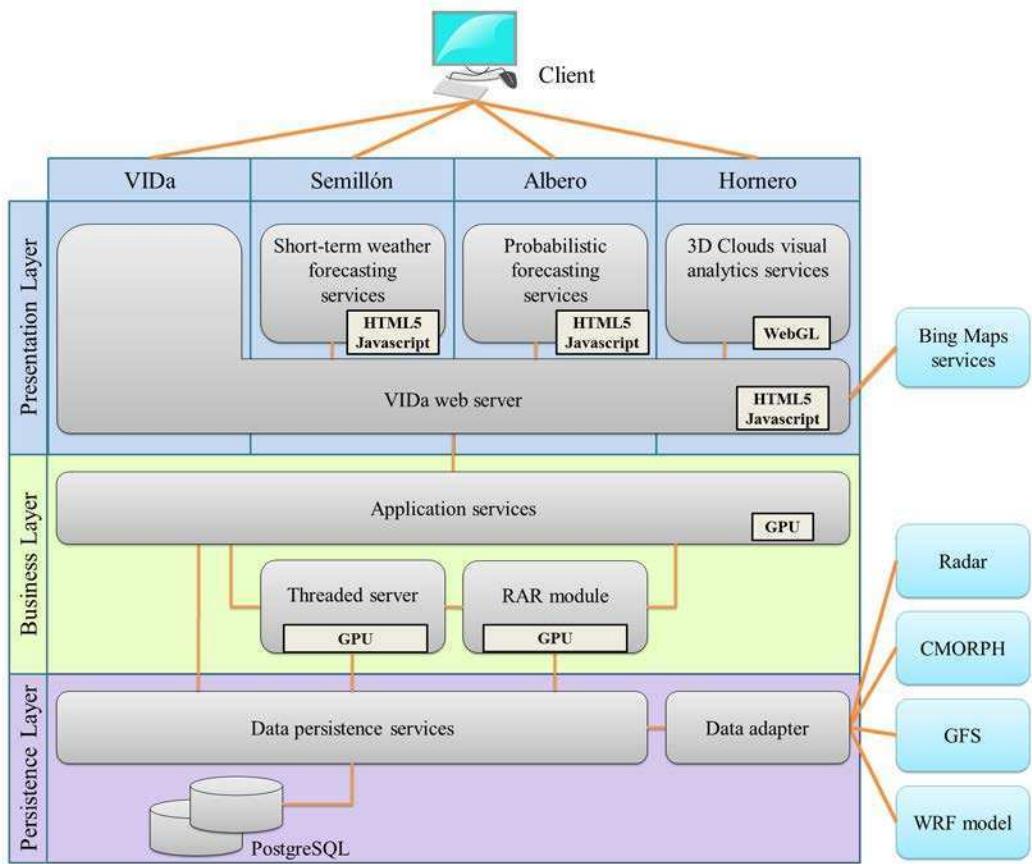


Figure 24: VIDa's architecture.

sented as 2D scalar-fields, such as the case of temperature. Our algorithms process and compare the 2D scalar-fields in time and space.

We used GPGPU programming for the development of the algorithms that run in VIDa's server. The experimental graphics server counts with a Tesla C2070 graphic processor with 448 CUDA cores and 515 gigaflops in double precision.

Different GPU kernels were programmed for the acquisition, transformation and visualization of the meteorological data coming from the CIMA, GFS and CMORPH systems. Figure 25 depicts how the processes are integrated in a GPU visualization pipeline. They all are implemented in OpenCL in the graphics server. These GPU kernels implement algorithms for the following processes:

- Acquisition and transformation of geospatial data to geographic coordinates.
- Reprojection from LCC projection to Web-Mercator, which is the coordinate projection used by the map server [74].
- 2D scalar-fields Color mapping.
- Map server tiling [75].
- Comparison of 2D scalar-fields and classification of curve-patterns .
- Operations among of 2D scalar-fields.

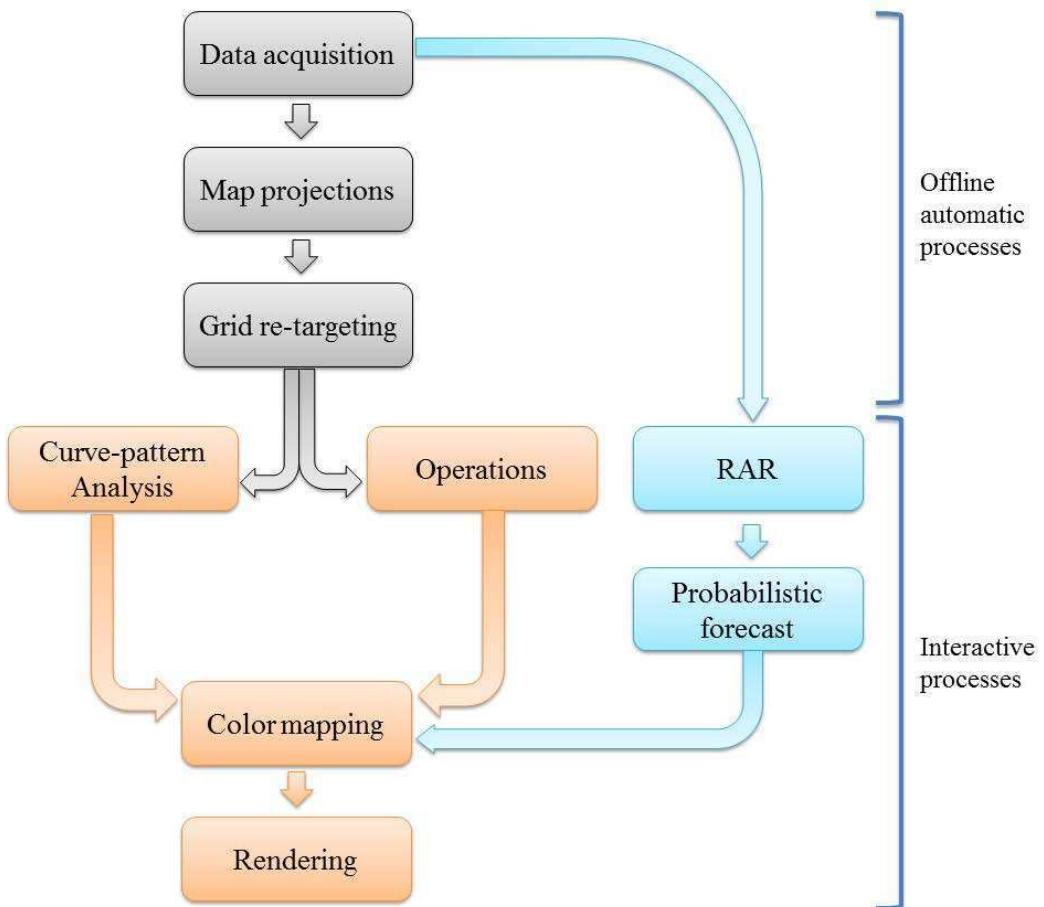


Figure 25: GPU visualization pipeline.

Data acquisition and transformation

Data coming from the CIMA's server are stored in binary format customized to be processed by GrADS [123]. Each file contains information about 11 different surface variables for each forecast date and time. This is an extract of the format information of the binary file:

```
dset ^SURFACE2014012012d01F36.dat
options byteswapped
undef 1.e30
title OUTPUT FROM WRF V3.0.1.1 MODEL
pdef 149 299 lcc -38.000 -63.000 ... 15000.000 15000.000
...
...
```

The name of the file contains information about the forecasted date and time, in this case “2014012012” indicates that the data file corresponds to the forecast 2014/01/20 12 : 00 : 00 and “d01F36” means that it was forecasted 36 hours before. We can also derive the forecast run from this information subtracting 36 hours from the forecasted date and time. The following lines inform about format, precision and other file settings.

This data and the data values of the meteorological variables are processed and stored partially in a geospatial database and partially in the file system. This information is accessed by other GPU kernels to post-process the information.

Map projection

We work with 2D scalar-fields of meteorological variables represented as 2D matrices. We developed GPU kernels to transform the 2D scalar-field coordinates from LCC projection to Web-Mercator projection. Figure 26 shows the difference between both projections. The matrices are uniformly distributed in LCC projection, so after this process, the distribution is not anymore uniform (Grid re-targeting).

The Web-Mercator projection is performed according to the webmap services selected by our solution, in our case Bing Maps. Bing Maps assumes

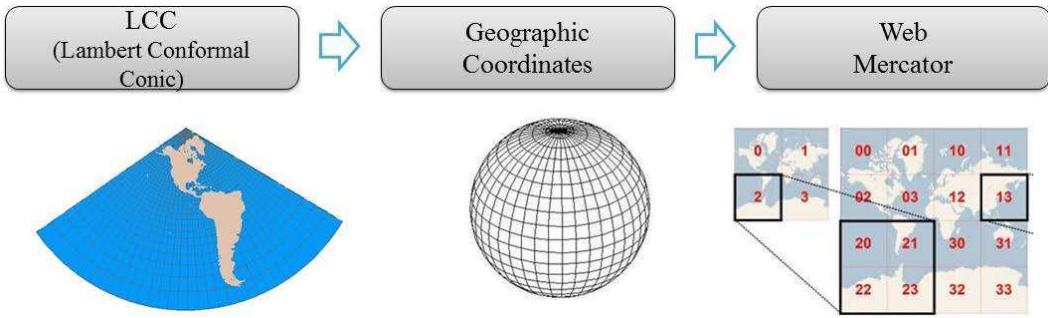


Figure 26: The transformation from LCC to Web-Mercator consists in two steps. Transformation from LCC to geographic coordinates and transformation from geographic coordinates to Web-Mercator. The webmap uses Web-Mercator coordinate system. Images adapted from [25, 74].

that the geographic coordinates use a WGS84 datum. Longitude is in the range of $(-180, 180)$ degrees and latitude is in the range of $(-85.05112878, 85.05112878)$. These ranges are established in order to avoid the singularities in the poles given by Web-Mercator projection.

Color mapping

The color mapping kernel takes the Web-Mercator matrices and transforms them into pixel coordinates. Pixel coordinates are calculated according to the webmap services in use by our solution. Afterward, the kernel uses the transformed matrices in pixel coordinates, colorcodes the data values according to the colormap selected and interpolates pixel values using a bilinear interpolation. The result of this process is a new image encoding the 2D scalar-field.

Each 2D scalar-field is displayed as a new layer on Bing Maps. Bing Maps Services allows us to provide different levels of detail, by means of different zoom levels. Each zoom level requires a new image or layer with a different resolution. The amount of squared meters per pixel depends of the zoom and the latitude. The size of the image depends on the latitude and longitude values to visualize and the zoom level. For zoom level zero, the complete world map is visualized in 512×512 pixels [75]. The height

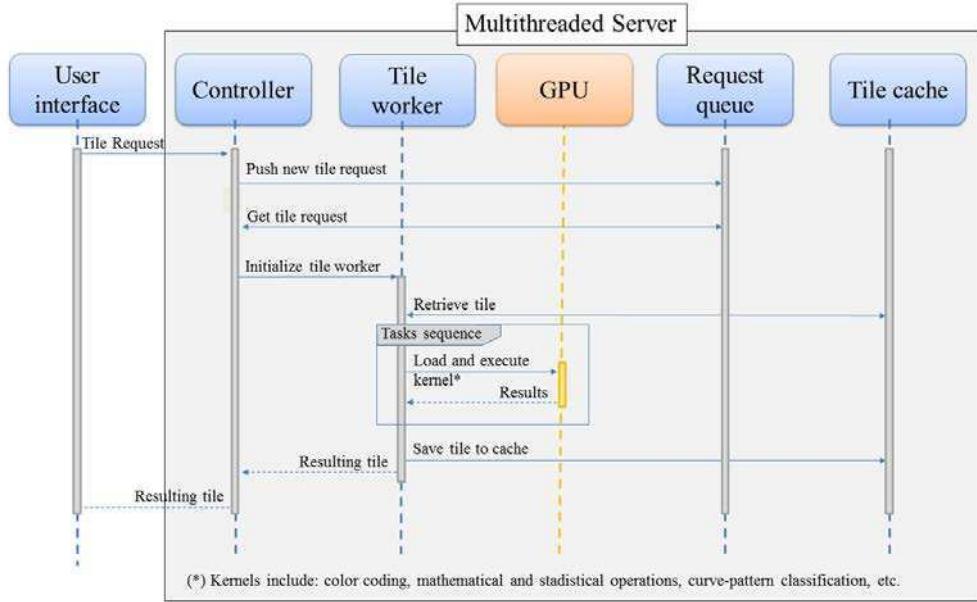


Figure 27: Sequence diagram of the multithreaded server.

and width doubles in size when the zoom increases, scaling the images by a factor of two.

Tiling

A multithreaded server was created to manage requests of larger images as it is shown in the architecture of the system (see Fig. 24). The multithreaded server offers an interface using sockets. The client calls the interface specifying the task to perform (see Fig. 27). For example, if the task is an operation, it specifies the selected operation, the time steps to use (geographic location, meteorological variable, forecasted date, forecasted run date) and the zoom level. This request is queued and control is returned immediately to the client.

The multithreaded server has a controller thread that is continuously monitoring the input queue. When the controller finds a new request, it pops the request, it creates a new processing thread and assigns the new work. The controller also monitors the completion of the processing threads. When a request has been processed, the controller returns a response to the client by using the same socket and then terminates the processing thread

and closes the socket. Finally the controller optimizes the performance by using a caching mechanism. When it receives a new request it checks whether exactly the same request has been previously processed and, if that is the case, it returns to the client the same tiles that were previously generated and conveniently saved in a caching repository.

Curve-pattern classification for Semillón

This algorithm is implemented using GPU kernels for the temporal evolution mode and the forecast verification mode. The algorithm is described in detail in Section 2.2.1. The kernels receive 2D matrices representing the selected 2D scalar-fields, a list of curve-patterns, a list of its corresponding color codes and the parameter Δ . The kernels are called by the tile worker, described in Section 2.4.2, previously verifying that the request was not already processed and stored in the cache. If it is a new client request, the kernel is executed for each required tile, given its zoom level and region of the map. The results of the classification are returned to the client. Fig. 27 shows how the kernel is invoked from the client and integrated in the multithreaded server.

2D scalar-field operations

The algorithms corresponding to the operations of the Operations tool are implemented using different GPU kernels. Each kernel receives 2D matrices representing the selected 2D scalar-fields, color scale range and meteorological variable range. The operation is applied to the set of matrices and the result of the operations are returned to the client. Fig. 27 shows how the kernel is invoked from the client and integrated in the multithreaded server.

RAR algorithm for Albero

We adapted and implemented the “rank analog technique” with smoothing of Hamill and Whitaker [39] into the GPU. The algorithm reads the reforecast data and stores only a range of 90 days for each reforecasted year in memory. A 90 days-range is selected around the forecasted date, as ex-

plained in Hamill and Whitaker's work. The algorithm only stores data corresponding to the geographic coordinates of the region, used as a bounding box. The bounding box corresponds to a Bing Maps's tile. It stores reforecast data in memory and transfers it directly to the GPU. The algorithm processes the input data for each tile. Its output is a list of analogs and their associated MSE. Finally, the analogs are processed to create the probabilistic forecasts.

Chapter 3

Case studies

3.1 Overview

We show the potential of our solution through a series of case studies. We worked with several domain experts that are specialists in weather forecasting research. Since the inception of VIDa, we have worked with two senior researchers with more than 30 years of combined experience: Prof. Celeste Saulo, who is the director of the National Weather Services in Argentina and Dr. Juan Ruiz, who is a senior researcher at CIMA. They have participated in the design and evaluation of VIDa and its component Semillón. In a second stage of the project more specialists joined us: Dr. Yanina Garcia Skabar, who is the head of the Research and Development department at National Weather Services and two members of her team, Laura Aldeco and Cynthia Matsudo. They contributed in the design and evaluation of Albero.

While working with their usual tools and procedures, our domain experts perform most of the data analysis manually. They use tools such as GrADS to generate the 2D plots and later they visualize and compare them side by side. They use MATLAB to perform operations over the forecasts, visualize the results and compare them using GrADS. For the RAR analysis the National Weather Services also use its own visualization system [2]. This system only shows the final results of the technique, 2D probabilistic forecast plots.

3.2 Semillón

The case studies for Semillón cover an analysis of the temporal trends, an analysis of model errors, and an analysis of forecast uncertainty among multiple runs. For all of them, the meteorologists chose the temperature variable at an altitude of two meters above the land surface. They have chosen temperature because it is one of the most influential variables in activity planning, decision making and productivity.

To perform the case studies we conducted unguided sessions where the application was presented to the users. While the users explored the application, we took notes and learned about their feedback. These sessions were done several times, each one followed by a period of time where feedback was incorporated into the application. These iterations concluded when their feedback was entirely positive. The case studies shown in this chapter were some of the results from the last sessions.

3.2.1 Temporal trend analysis

In this case study the users were looking for salient features or trends. Examples of weather trends are significant changes in temperature magnitudes that could possibly indicate the passage of cold or warm fronts. A key point in this analysis was to differentiate cases that contained unexpected information from those that contained expected information. The meteorological variable temperature expressed two well-known behaviors corresponding to a diurnal curve (increasing values) and a nocturnal curve (decreasing values) of temperature development (see Fig. 28).

The meteorologists focused on interesting patterns that might occur at the same time of the day but on different days, i.e., one, two, and three consecutive days. In this case, the users selected a set of forecasts from the same run, predicted at the following time steps: 0-hour, 24-hour, and 48-hour. Then, they selected the operation mode as temporal evolution, indicating a $\Delta = 8^\circ$ Celsius as differences between forecasts. Finally, they selected a qualitative list of different curve-patterns represented in terms of a large positive variation, a small positive variation, a large negative vari-

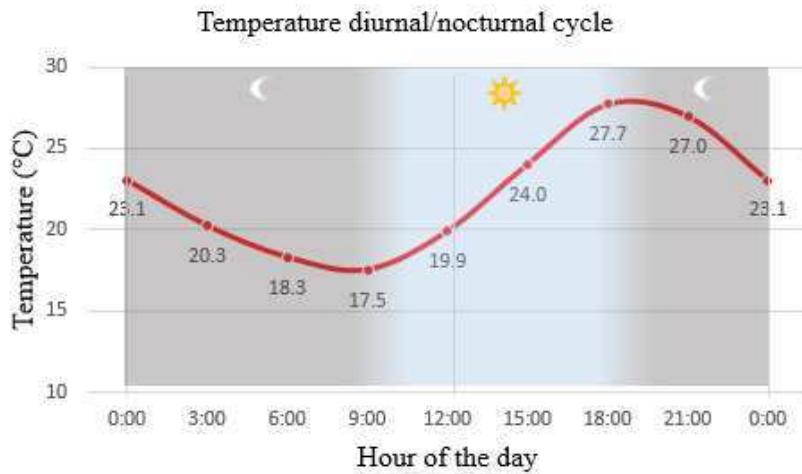


Figure 28: Diurnal and nocturnal temperature development.

ation, a small negative variation and no variation and their corresponding colors, using the "I" component of the YIQ (Luma In-phase Quadrature) color space.

Figure 29 shows a visualization of the results of the curve-pattern classification using two different Δ values. The classification is applied to the three aforementioned time-steps. In the resulting images, the spatial areas with decreasing temperature are visualized in cyan tones, and increasing temperature areas are visualized in orange tones. A large spatial region is covered with cyan tones in the center region of the map. This region corresponds to a cold front event that was moving from south to north near the center of the domain producing a significant temperature drop. The figure shows two results associated with Δ values of 6° and 8° Celsius. Larger Δ values create more relaxed conditions of similarity against the patterns, therefore larger spatial areas can be associated with each pattern. This can be noticed by comparing the cyan spatial areas in the results. White areas correspond to regions of the map that do not match any pattern. This case study shows the effects in the temperature caused by the presence of a cold front that was identified and analyzed. Now, forecasters are able to detect this kind of patterns in a straightforward way by using our curve-pattern selector.

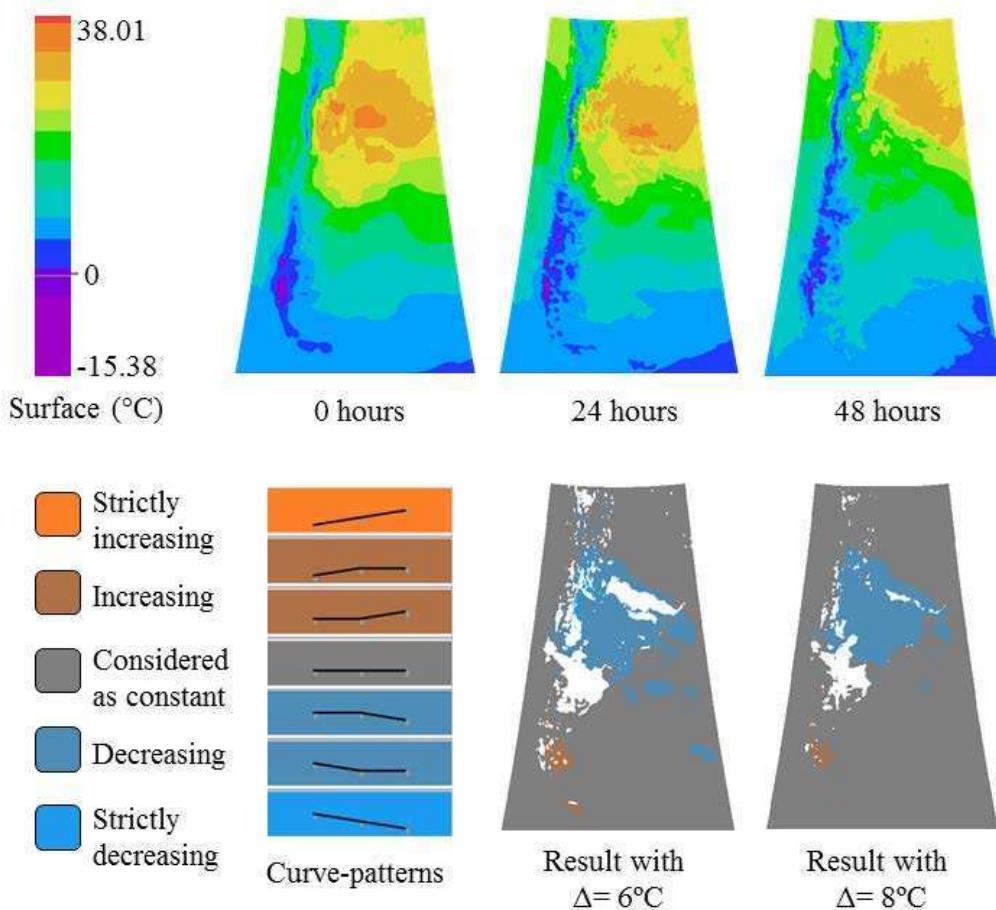


Figure 29: Trend analysis: the largest temperature drops, just after the passage of the cold front, are indicated in cyan tones. Orange tones in the south of the map indicate increasing temperature. Results are presented using two different Δ values: 6° and 8° Celsius.

3.2.2 Forecast verification

The main goal of forecast verification is to improve the quality of weather forecast models. Improvement of a weather forecast model requires a robust error analysis. The forecast error can be obtained by comparing the numerical forecast with the observations. Another possibility is to compare the forecast against atmospheric state diagnostics that are routinely generated to be used as initial conditions for the NWP models. The diagnostics are generated as an optimal integration of numerical simulations and observations coming from different sources (e.g., satellites, weather stations and radiosondes). In this work we use the diagnostic that comes from the Global

Data Assimilation System (GDAS) [35] as a reference in the computation of the forecast error. In this case study we present an analysis of model errors among multiple runs. Due to the chaotic nature of the atmospheric flow, a forecast error usually increases with the forecast lead time. However, the rate of growth strongly depends on the weather phenomena occurring at the particular time and region. In this analysis, the users want to identify spatial areas where the error increased and analyze if there are some interesting phenomena affecting the model.

Every day, two short-term 48-hour forecast cycles were run at 12:00 AM and 12:00 PM UTC. Therefore, we had 4 forecast runs and a diagnostic in a complete 48-hour cycle. From the timeline, the user selected a targeted date and time for the forecast runs. The vertical axis aligned all the runs from the selected date and time (see Fig. 13).

In this case, the users chose the operation mode for forecast verification and a suitable $\Delta = 8^\circ$ Celsius. They also selected a subset of curve-patterns, representing increasing errors and associated colors. They applied the YIQ color space as before, but only with orange tones to represent positive changes.

The spatial areas where the error presented a major increase and significant rate of change are depicted with orange tones (see Fig. 30). These areas correspond to the position of the cold front discussed in the previous case study. As cold fronts are characterized as areas where the temperature gradient is strong, small errors in the location of these atmospheric boundaries produce large errors in the forecasted temperature as can be seen in the figure. This case study shows how the tool can facilitate the analysis of the performance of the numerical model under different atmospheric conditions. In this case, the performance is affected by the cold front introducing large errors in the forecasts.

Also, in the same figure, systematic errors are observed under the Andes mountains range. This is caused by the low resolution of the model there, where the topography is very irregular. The visualization of topography and visual comparison of model and observations is a very interesting topic that will plan to address in future work.

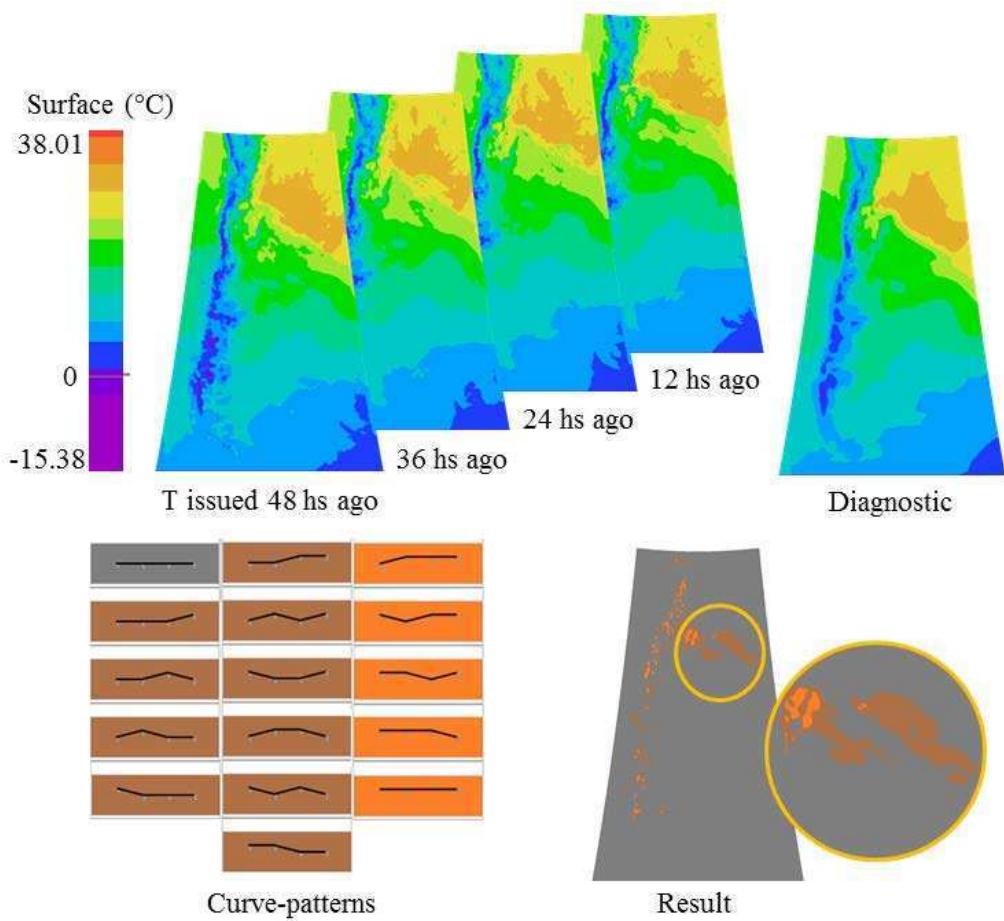


Figure 30: Forecast verification using multiple runs. It shows the passage of a cold front. This is observed as errors in the forecasted temperature shown in orange tones in the “Result” image.

3.2.3 Forecast uncertainty analysis

In this case study we present an analysis of forecast uncertainty among multiple runs. Forecast uncertainty can be addressed by analyzing the results of different operations such as standard deviation, subtraction and mean value. The standard deviation operation is applied on the multiple runs to visualize its dispersion, which is a measure of forecast uncertainty. The users selected the same four multiple forecasts as in the previous case study. They found more dispersion in the results over north central Argentina (see Fig. 31). This area of large forecast uncertainty is associated with a displacement of a cold front. Fronts are associated with strong gradients in the temperature field, and even small changes in the forecasted positions

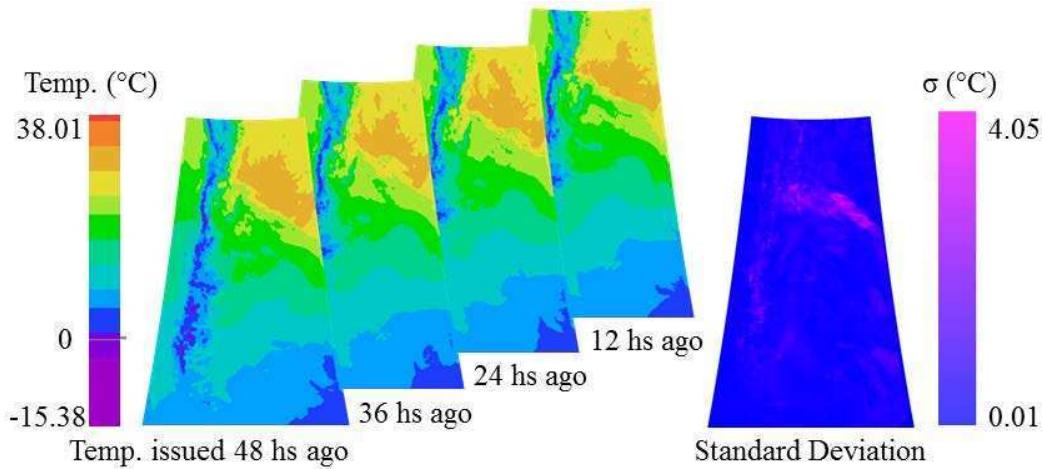


Figure 31: Forecast uncertainty analysis. The standard deviation shows the dispersion of the forecasts in different regions of the map.

of these systems can produce large changes in the values of the forecasted temperature, resulting in an increased uncertainty in the area.

The subtraction operation allows the user to detect significant changes between forecasts such as a temperature drop in the central part of the country (see Fig. 32).

3.2.4 Lessons learned

An important outcome of this iterative participatory design approach was that we quickly realized the importance of having a visual overview of the entire set of short-term weather forecasts. The minimap timeline turned out to be a key component of the whole system. This is due to its capacity to show multiple forecasts in a single view, with the additional benefit that it could be extended to visualize information related to multiple ensemble forecasts. Furthermore, the constant feedback by experts also showed the importance of keeping a simple visualization interface.

The meteorologists' feedback was highly positive. The applications they currently use do not allow them an easy comparison between forecasts initialized at different times. Instead, our solution provides them with easy-to-use functionality, specifically designed for this task. One of the prominent aspects is the flexibility with which the user can select a group of forecasts

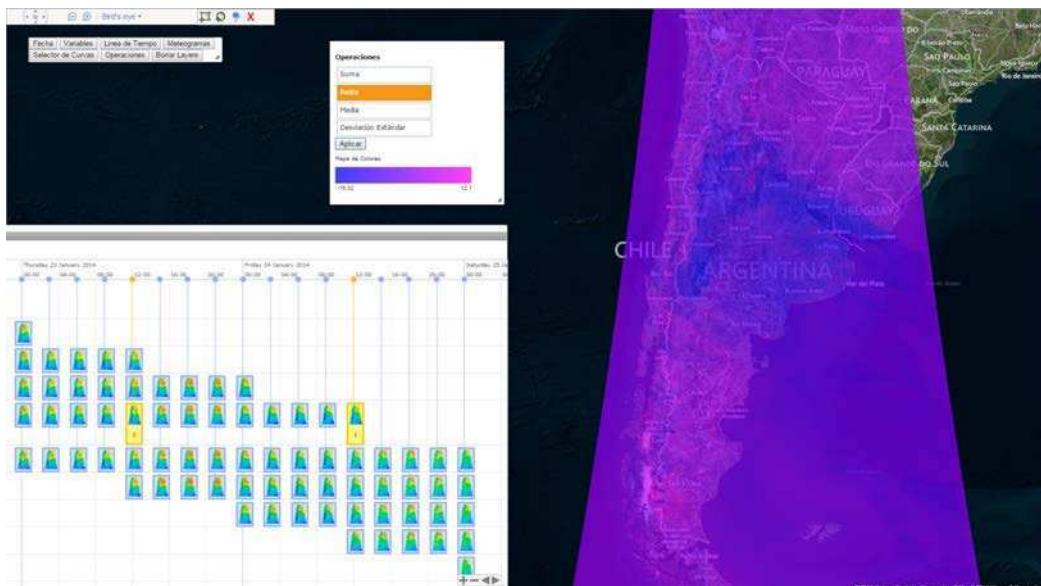


Figure 32: Forecast subtraction analysis. The subtraction among two forecasts shows distinguishable changes in the central part of the country.

and compute operations. Another valuable feature is the simplicity of the user interface. In particular, the domain experts emphasized the straightforward use of the timeline and the curve-pattern selector. One of our domain experts stated: “Now, with these mechanisms, forecasters who are interested in different areas of the country, can focus on what is happening there, what the tendencies would be, and the significance of the model errors”.

The experts said that one of the major impacts of our system is the capability to display all the cycles of short-term weather forecasts in a single view. They mentioned that with our solution, it is now possible for them to synthesize the temporal evolution of weather forecasts along the timeline. Moreover, they can spot groups of forecasts where salient features and significant variances are occurring at first glance.

Further feedback was related to the use of our tool to perform forecast operations. The capability of Semillón to perform simple operations between multiple forecasts with few clicks allows users to measure forecast uncertainty in different ways. For a forecaster, it is very useful to compare differences among the latest forecast and the previous forecasts.

The domain experts also stated that the curve-pattern selector and the curve-pattern classification are key components of Semillón. The flexibility

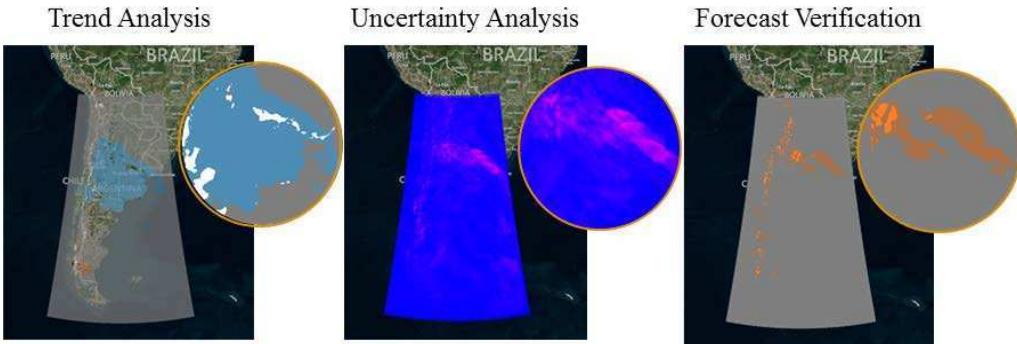


Figure 33: Summary of the case studies that demonstrate the usage of the tool, showing the curve-pattern behavior, dispersion, and errors associated with a cold front displacement.

of the curve-pattern selector opens new scenarios of analysis. These new capabilities allow them to quickly define a set of curve-patterns based on their experience and detect them in the forecast evolution of different variables. This is particularly useful in operational weather forecasting where they need a rapid analysis of forecast trends. One of our domain experts stated: “In operational meteorology, it is helpful to have a tool with the functionality that offers the curve-pattern selector. In a few steps forecasters can define a set of meaningful patterns and see which regions are associated with them”.

Also, the curve-pattern selector allows the users to define, save and reuse a given set of curve-patterns. They found this functionality very useful for some scenarios that cover large number of possible curve-pattern behaviors making the selection of an appropriate set, a challenge. By saving and reusing the curve-pattern sets, they can continue refining them repeatedly until they find an appropriated set.

Semillón allows them to analyze the evolution of the forecasts and identify specific phenomena in the data, such as the analyzed weather fronts. Figure 33 shows how the three different visual analysis described before assisted the users in the detection of the cold front phenomenon. This visual analytics process can be used for diverse kinds of situations that forecasters need to evaluate everyday.

Limitations

Our curve-pattern selector is limited by the number of different color categories the user can associate to the curve-patterns. Usually these categories are less than 12 different colors due to limitations of visual perception.

Another consideration relates to temporal scalability. Forecasters work with short-term forecast cycles where the visual analysis can be done using a reduced number of time-steps. However, the curve-pattern selector might need further extensions to overcome possible scalability issues when used for long-term forecast cycles. Examples are climatology studies where the temporal scale is significantly larger. Also, the minimap timeline could get overcrowded when larger forecast leads are visualized, for example in a 15-day forecast cycle. In this case other techniques such as visual lens, clustering and compression techniques could be evaluated.

3.3 Albero

For Albero we also conducted unguided sessions where the application was presented to the domain experts. We took notes and learned about their feedback. The evaluation case study focused on the analysis of the precipitation accumulation variable. We chose the precipitation accumulation because it is a key variable to analyze cases of extreme precipitation as well as flooding, which severely affect our region. For the evaluation, we have done two guided interviews with seven domain experts, with different

We exemplify the capabilities of Albero by means of two scenarios that cover operational weather forecasting and retrospective weather analysis, for an event of extreme precipitation in a short period of time. This event took place from the 1st to the 3rd of April in 2013, causing flooding in different areas of the north-eastern part of our country. All domain experts evaluated these scenarios.

3.3.1 Extreme precipitation analysis

The first scenario is a case study that covers the analysis of extreme precipitation in a short-period of time. It was evaluated by our collaborators, who analyze probabilistic forecasts on a daily basis. In this case study, the meteorologists knew that the selected event was under-estimated by the numerical model, since it was a rare and very extreme situation and, in general, precipitation is one of the most difficult meteorological variables to forecast.

The domain experts wanted to analyze the uncertainty associated to the probabilistic forecast and RAR technique. To analyze this particular scenario, a forecaster selected a date and a region of interest. She wanted to detect the areas with the largest probability of precipitation accumulation. The user configured the forecasted date to the 1st of April 2013, the lead time to 72 hs, the range interval to 24 hs and the thresholds (above: 0 mm, 20 mm, 40 mm, 60 mm and 80 mm). Then, she visualized the results in the Übermap (see Fig. 34a). She looked at the second time interval ([24-48 hs]) for the larger thresholds such as above 60 mm to detect extreme events of precipitation accumulation.

The forecaster selected a probabilistic forecast from the Übermap where the probabilities were above 60 mm. She focused her attention on a specific sub-region near Buenos Aires city and zoomed on the mapview (see Fig. 34b). She noticed that during the second time interval, corresponding to the second day, it is forecasted the heaviest rain. Figure 36 shows a close-up of the probabilistic forecasts for the second time interval ([24-48 hs]), for the thresholds: 20 mm, 40 mm, 60 mm and 80 mm.

Using the Analog Viewer she analyzed the numerical forecast, the analogs and the observations. Figure 35 shows the first 8th analogs and its corresponding observations. The difference between the analogs and the observations brings information about the systematic errors in the model associated with situations that are similar to the one anticipated in the current forecast.

By using Albero, meteorologists can study systematic model-bias and get further insight into their model errors. The bias is also displayed in

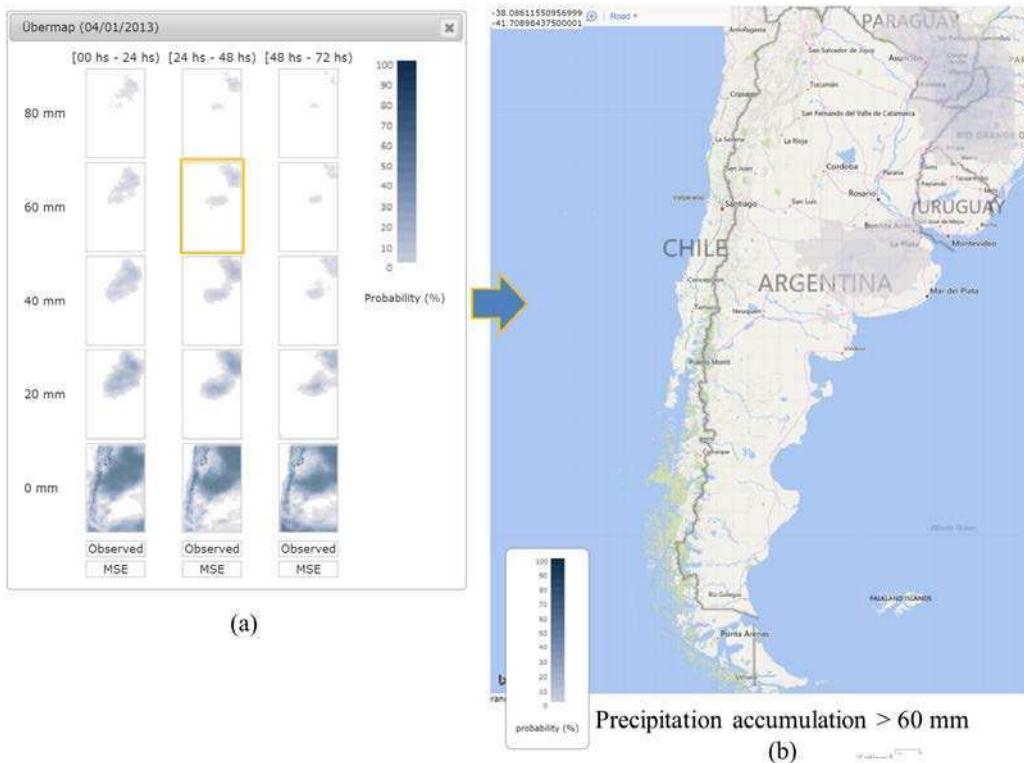


Figure 34: Albero showing (a) the Übermap for the 1st of April, 2013 for the next 72hs and (b) a detailed view of the probabilistic forecast above 60 mm.

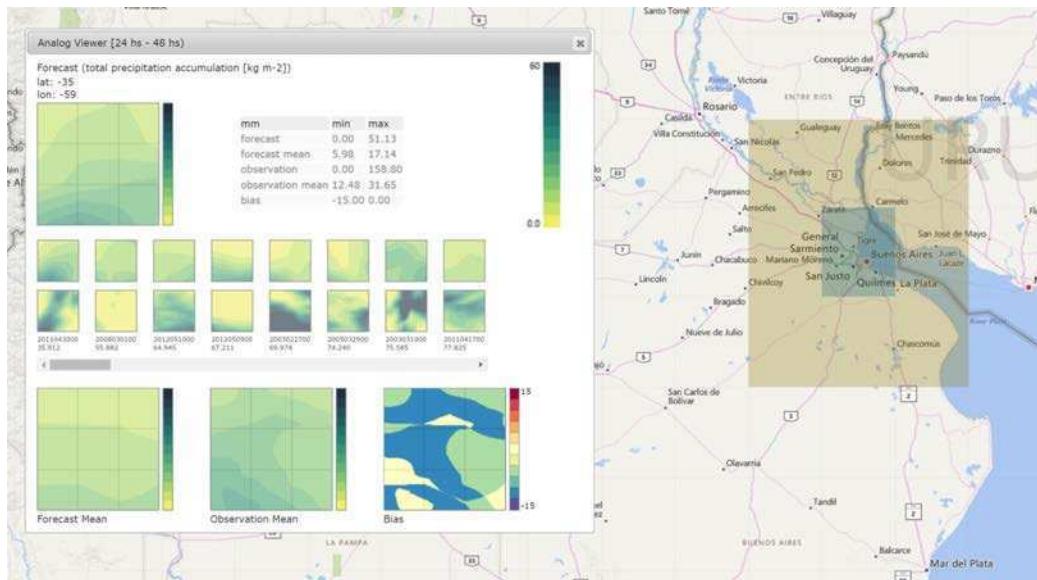


Figure 35: Albero Analog Viewer showing the RAR technique results and intermediate information over a region centered at longitude 59°W and latitude 35°S, surrounding the city of Buenos Aires.

Fig. 35. It presents large blue and green shades covering the selected sub-region, indicating areas of higher precipitation when compared with the analogs forecast mean. The user finds that the numerical forecast systematically under-estimated the amount of precipitation accumulation since the observation mean is higher than the analogs forecast mean.

Using the Übermap, the forecaster can select the Mean Squared-Error spatial distribution for a given accumulation range. This is a measure of how accurate the technique is for the particular situation the user is evaluating. Figure 38 shows a visualization of the Mean Squared-Error spatial distribution. The highest errors are located in the central part of the country. They are shown in red shades in the figure.

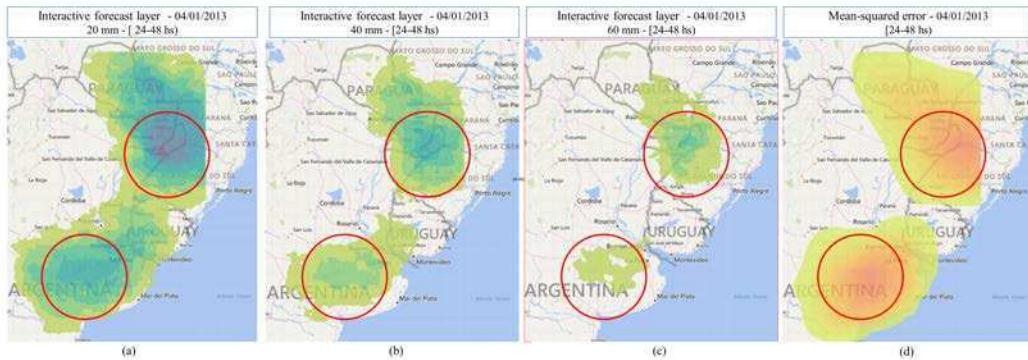


Figure 36: Interactive Forecast Layer showing different information from forecasts, observations and errors for the second day. The interval of precipitation accumulation is [24-48 hs), valid for April, 1st, 2013. The user finds that the probabilistic forecasts shows high probabilities of precipitation over the Pampa and Mesopotamia regions.

3.3.2 Albero for Technique Optimization

In this scenario, the meteorologist analyzes previous events. She knew that the selected event was under-estimated by the numerical model, since this was a rare and very extreme situation, and in general, precipitation is one of the most difficult meteorological variable to forecast.

Using the Analog Viewer she analyzed the numerical forecast, the analogs and observations. The difference between the analogs and the observations,

the bias, brings information about the systematic errors in the model associated with situations that are similar to the one anticipated in the current deterministic forecast. In that way, she compared whether the numerical forecast under-estimated or over-estimated the precipitation accumulation across the three days.

She analyzed the bias between the analogs mean and the observations mean, detecting large variations between them. Figure 37 shows that the numerical forecast systematically under-estimated the amount of precipitation accumulation for a selected sub-region in the given accumulation range, since the observation mean is higher than the analogs mean. This can also be observed in the individual analogs and their associated observations, and can be considered a systematic model-bias.

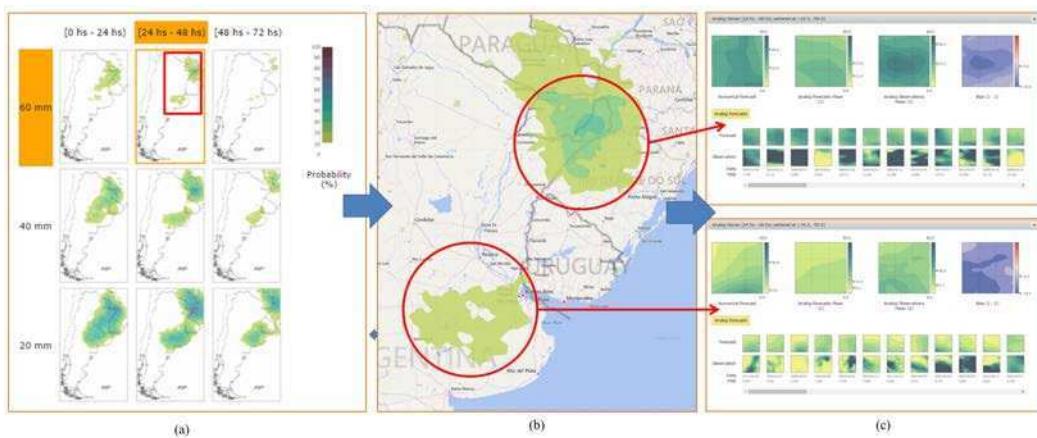


Figure 37: Albero screen-shot showing the extreme precipitation which happened during April, 2nd, 2013. We showed the second day because it was the date with the heaviest rain. (a) The user identifies a probability forecast with a large area of precipitation accumulation above 60 mm. (b) The user zooms-in and highlights two different areas of interest. (c) Afterwards, the user analyze the numerical forecast for that region and its associated analogs and observations. She finds out that the numerical forecast systematically under-estimated the amount of precipitation accumulation since the observation mean is higher than the analogs mean and also that analog forecasts for both regions presented noticeable differences.

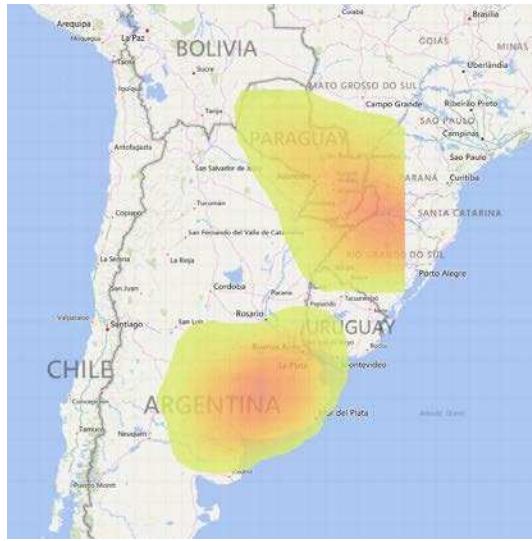


Figure 38: Distribution of the Mean Squared-Error for a given forecasted date and accumulation range. A higher Mean Squared-Error in the central north of the map indicates that the forecast had increased differences when compared to the observations.

3.3.3 Lessons learned

The general feedback was positive and encouraging. Meteorologists stressed out the potential of Albero. One of them expressed: “The capacity to connect and compare different pieces of information from probabilistic forecasts and internal views of the RAR technique brings us new available information. Previously we could only visualize the probabilistic forecasts. Albero enables us to do different kinds of analyses interactively, which were impossible before.”. Another expert said: “Now, with Albero we have the possibility to select different lead times and accumulation times, which were unavailable with our previous tools”. But they also alerted us: “Although satellite observations have the advantage of a vast spatial and temporal range, the data have lower precision at southern latitudes than in the central latitudes. It would be helpful to incorporate data from land weather stations as part of future work, the observations will be more accurate”.

3.4 Conclusions

VIDa demonstrated to be a powerful tool for operational weather forecasting, supporting a complete analysis of short-term weather forecasts by means of Semillón and an exploratory analysis of probabilistic forecasts by means of Albero. Our solution assists forecasters in their daily work where they have to process large amounts of data and extract from them trends, anomalies, analyze uncertainty and the errors in their model.

Moreover, VIDa helps forecasters to perform a step further in the analysis and identify what are the phenomena behind a given trend, how is the uncertainty associated to these events, and how the errors are affecting the model. This analysis is done in an integrated, quick and efficient process. This is key for operational weather forecasting, where forecasters need to perform a quick analysis of the information and communicate the results immediately, so stakeholders can be ahead of possible events and make decisions in advance.

A meteorologist stated: “One of the most important and challenging issues related to weather forecasting is the assessment of forecast uncertainty. VIDa provides us with different ways to assess it. Uncertainty changes from one forecast to another, and with VIDa we can estimate it using the extensions that it provides (Semillón and Albero)”. The meteorologist added: “These capabilities are extremely useful to detect changes in the atmospheric variables and the level of forecast uncertainty for a particular situation”. Analyzing ensemble uncertainties and keeping a historical record of the output may serve as a basis for further investigation of model errors, error growth and regional error behaviors among many other applications.

While our software is still a prototype, our collaborators plan to deploy it as a staging tool available for operational weather forecasting in parallel with current tools. Our solution implements a web Front-End that runs on the Internet, which facilitates the broadcasting of the information. Semillón and Albero extensions are intended to be publicly available. This action will allow us to perform a future exhaustive test of our solution.

Part III

Serious games applied to geovisualization

Related work

1.1 Overview

Roughly defined, a computer game is a contest with rules where the user tries to accomplish specific goals. Most computer games have entertainment as motivation, however serious games have a different purpose that is to impart certain knowledge or expertise to the users [139].

The proliferation of general game engines [3, 14], as well as the development of specific engines and editors for particular games [19], contributed to establish the use of games for other purposes. Among the most influenced areas, there are applications of serious games in training, education, health, CAD, marketing, simulation and social sciences. Education and edutainment, especially in training and research, are two of the most active application areas of serious games. Zielke et al. [138] for instance, proposed the use of serious games for cultural training. In their work, an application is designed using two game engines (RAGE and Natural Motion). The user learns by means of storytelling the complexities and nuances of a culture and the whole game is used as a training platform. Another example is the work of the Case Western Reserve University's Freedman Center [17] described in Bendis [12], in which a Virtual-Reality (VR) gallery market is developed using several game engines. This tool enables educators and students to create virtual exhibitions.

Computer games exert a strong influence in several other research areas, as tools or frameworks that provide features to construct virtual environments and also virtual objects. In VR, Trenholme and Smith [122] proposed the use of an engine or a combination of several engines to take advantage of the already existent computer game functionality. This approach presents the advantage of using functions that are fully tested in a product. This ensures robustness, usability and better performance. Another example in the area of computer networks, is Kienzle et al. [57]. They used multiplayer games to simulate network situations and to detect and analyze them. In health research, Immune Attack [56] is a game designed to teach immunology to students of different backgrounds, sponsored by the National Science Foundation (NSF). Figure 39 shows several areas of application of serious games such as geovisualization.

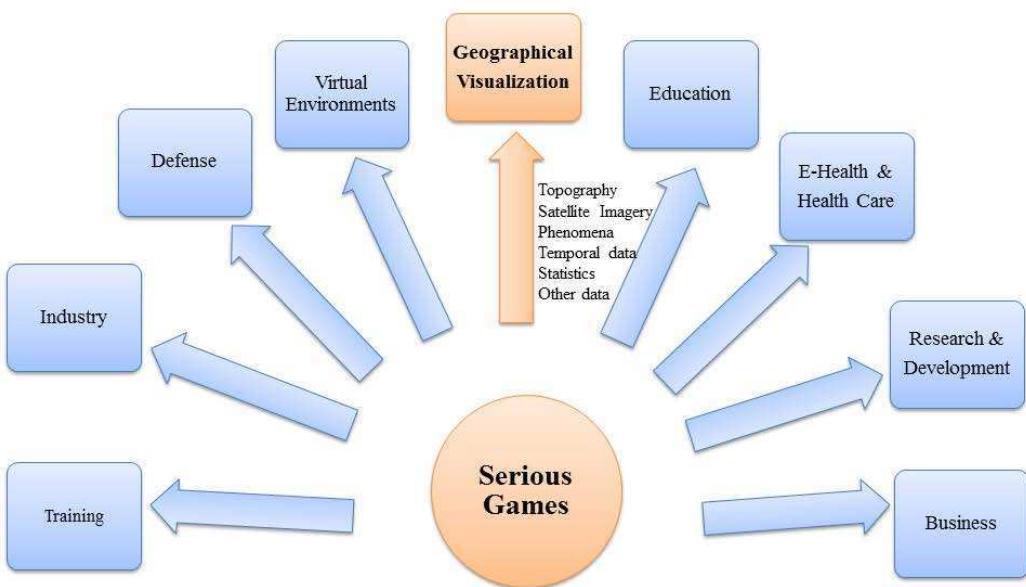


Figure 39: Typical areas of application of serious games.

This chapter presents in section 1.2 an overview of some available frameworks for the implementation of serious games. Section 1.3 explores available flight simulators technology. Section 1.4 presents a description and discussion of OpenSceneGraph (OSG). This framework was used in this thesis for the development of the case study presented in the chapter “3D Geo-

visualizer”. The case study illustrates how serious games can be used for geovisualization.

1.2 Game engines

Game engines are the result of efforts done by game developers to reuse common features of games. The game structure can be generally divided into the story, the game logic and the media resources [69]. Each game develops its own story and game logic, but a game engine should be able to separate the specific content from the common features facilitating reuse across several games [14]. Although it is possible to define a generic engine with common features that work in any game, there are game engines specialized in different kind of games. Some popular categories are: First-Person-Shooter (FPS) games, sports games, adventure games and simulation games (Sims-type) [122].

Early concepts of game engines emerged with general-purpose graphics package such as Criterion Software’s RenderWare, Argonaut Technologies’ BRender and Rendermorphics’ RealityLab. Bishop et al. [14] described the design of a particular game engine, setting the focus on the scene management features and the modularization and encapsulation of the reusable common features of games. Quake III Arena [44], and Unreal Tournament [127], both FPS games, were released with their respective engines. Id Tech 3 engine (Quake III Arena engine) was designed as a multi-player game engine, and became a de-facto standard for commercial games. During the last two decades, many other game engines have appeared, mainly FPS. An example is Source Engine (Half Life 2 engine) [128] which was used in many other areas such as health, education and training. Also, Unreal Engine [126] (used to develop Unreal Tournament 2003 and 2004) has had many other application areas, for example: virtual museums [62], interactive storytelling [18] and landscape visualization [90]. In the area of simulation-based game engines, FlightGear flight simulator [31] has also gained great popularity. This popularity raised interest in other frameworks such as JSBsim [49] for flight dynamics modeling (FDM) and OpenSceneGraph (OSG) [71] for scenes cre-

ation and 3D graphics support. The appearance of advanced 3D graphic capabilities in consoles and PCs was fueled by the emergence of dedicated graphic hardware (GPUs). Among these 3D engines, there are well-known commercial engines: C4 Engine, Torque Game Engine and 3DGameStudio, just to name a few. There are also several well-known open-source engines such as: OGRE, Irrlicht, Panda3D, Crystal Space, jME, Blender Game Engine, Reality Factory, The Nebula Device 2, RealmForge, etc. Figure 40 shows a basic and generic game application model. This model describes how the “3D engines” act as a middleware between the specific “Game source-code” and the other low level services, for instance, “GPU drivers”.

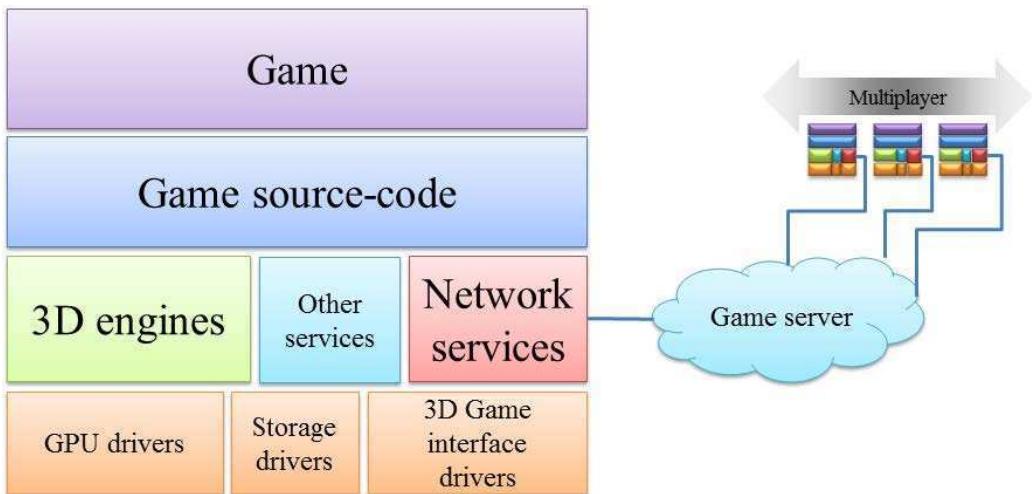


Figure 40: Generic game application model.

There is a wide spectrum of commercial and open-source game engines that can be used for geovisualization purposes. A specific choice must take into account some aspects of the final application [63,122,139], but in general the following features are desirable.

Black box. A game engine module may be considered as a black box that can be included in the development of an application, even when the source code is available.

Extensibility. Specific features of an application domain may not be included in most common 3D engines. For example, if the application requires

a specific model of atmospheric and climate data, then the representation should rely on the engine's extension capabilities (extensibility) to store ancillary attributes and also to adapt or create new algorithms.

Complexity. The use of a 3D engine has an associated learning curve, but when combined with other APIs or engines, the learning curve's slope might increase, along with the development complexity (that depends on the modularity, cohesion, extensibility and generality of each engine).

Support. Open-source engines may constitute a good choice when developing a project, but sometimes their assistance to developers relies only on community support, which may become risky when compared with official support from copyrighted products.

Game engines and frameworks can be used as a basis for the construction of interactive applications for geovisualization. This approach takes advantage of the use of functionality that is fully-tested in a product, increasing robustness, usability and performance. Engines designed for flight simulators are well-suited for the construction of virtual environments and the representation of several phenomena. However, building geovisualization applications requires additional scientific concepts. For example, they might require the analysis of geographical information and spatial statistics [129]. Most of the geovisualization tools adopt a third-person or high-scale standpoint to explore through different layers of information. The interactive and dynamic first-person standpoint used in serious games represents a profitable opportunity for the creation of geovisualization tools.

1.3 Flight simulators

Flight simulators emerged as a specific technological product by themselves. They have experienced a remarkable evolution during the last years in performance, realism and interactive features. This is mostly due to the leverage of GPUs power to render increasingly complex 3D models in real-time.

The quest for photo-realism, advanced interactive features and dynamic aspects, provides a huge thrust to commercial and scientific research of this kind of applications. Some typical flight simulator features have a great potential for their usage in the development of geovisualization applications, for example terrain rendering [4,9,86], modeling and visualization of clouds and atmospheric phenomena [41] and other natural phenomena [30,92], just to mention a few.

Some of the most popular flight simulators are FlightGear 2.0 [31], X-Plane 9 [73] and Microsoft Flight Simulator X [76]. They have several common features. Their topographical model divides the Earth in a uniform grid with tiles limited by the Earth's latitudinal and longitudinal lines. Each cell of the grid contains information about the terrain elevation data and objects positioned over the terrain. These objects can be point-based positioned over the terrain (i.e., light beacons), or distributed in an area (i.e., rivers, roads, etc).

1.3.1 FlightGear

FlightGear simulator is a product of an open-source, multi-platform and cooperative development project [11]. It is copyrighted under the terms of GNU General Public License (GNU GPL). FlightGear has its own internal coordinate system. Each tile is by itself a scene. The scene's structure can be separated in three main parts: terrain elevation data, airports geometry and information about other objects on the topography. As a part of its features, it provides several tools that can be used for data acquisition and manipulation of shape files and DEM (Data Elevation Model) data. For example, "TerraGear" [113] is a complete open-source sub-project of FlightGear. It supports the creation of internal files representing the elevation and texture of the Earth, including airports, cities, fields, forests, rivers and roads. FlightGear supports several different FDMs that are chosen at runtime, for example JSB flight model developed by Jon Berndt (which is part of a stand-alone project called JSBSim) [13].

1.3.2 Microsoft Flight Simulator X (FSX)

The Microsoft Flight Simulator X [76] is a flight simulator commercialized by Microsoft. It has its own coordinate system as an oblate spheroid, an ellipse rotated about its minor axis. Information is saved using their own internal coordinate system. Each tile contains information about the geographical region, population density, land and water classification and season. FSX offers a diverse set of Software Development Kits (SDK) for the modification and extension of its features. An SDK can be used to create or modify add-on components such as aircraft, missions, scenery, terrains, airport ground vehicles, airport runways and buildings, special effects, etc.

It also provides an FDM, SimEngine, that simulates various aircraft systems. Some of these systems are: electrical, fuel, oil and cooling systems, among others [140].

1.3.3 X-Plane

X-Plane is a flight simulator built by Laminar Research [73]. X-Plane is fully oriented to train professional pilots. It is used by defense contractors, air forces, aircraft manufacturers and space agencies for flight training. Its topographical model contains three different types of information: terrain data, airport data and navigational aids (i.e., light beacons and windsocks). X-Plane is built using OpenGL APIs and provides extensibility by means of plug-ins that developers can create. Some of the graphic features are: fog, lighting, alpha testing, alpha blending, depth testing and depth writing. X-Plane also allows for the construction of scenery with new objects, aircraft and navigation aids. It has its own FDM that includes subsonic and supersonic flight dynamics. It also contemplates changing weather in simulations such as rain, snow and storms and maintains information about thermals and real weather conditions taken from internet web services.

1.3.4 OpenFlight

There exists an open standard available for topographical models, named OpenFlight [97]. It is a generic standard without specific semantics to entities. Several component types are used to model the Earth and its geospatial objects (see Fig. 41). The Earth is represented by grid structures (modeled as a Grid structure node) composed of one or more grids (Grid node). Each Grid structure node has an associated grid resolution (Level node) and a terrain extension (Coverage node). OpenFlight describes the scene by a hierarchical tree structure that models entities and relationships among them. An entity is represented by a container (Container node) composed of geometries (Geometry node). Containers have associated features such as Level Of Detail (LOD node) and Degree Of Freedom (DOF node). Geometries are composed of vertices (Vertex node). Different Geometry node types represent text (Text node), light features (Light point node), meshes (Mesh node) and textures (Face node).

1.4 OpenSceneGraph (OSG)

OpenSceneGraph (OSG) is an open-source, cross-platform, high performance 3D graphics toolkit, used by application developers in fields such as visual simulation, games, VR and scientific visualization [89]. Several products were built on top of OSG, including FlightGear.

Its rendering framework provides a scene graph API [71]. The core of the framework supports the creation of a scene graph of nodes, geometries, textures, transformations and state management. It also contains algorithms for node finding, callbacks, user input functions, multi-camera control, camera layouts, custom windows and support for GLSL shader programming language (OpenGL 2.0) [70]. There are a lot of implementation examples of 3D graphical features, animations, interactions and special effects such as the usage of billboards, particle systems, fog effects, smoke, tessellation, terrain rendering, texture compression, flight simulation, gliders, etc. Many of them apply well-known computer graphics techniques

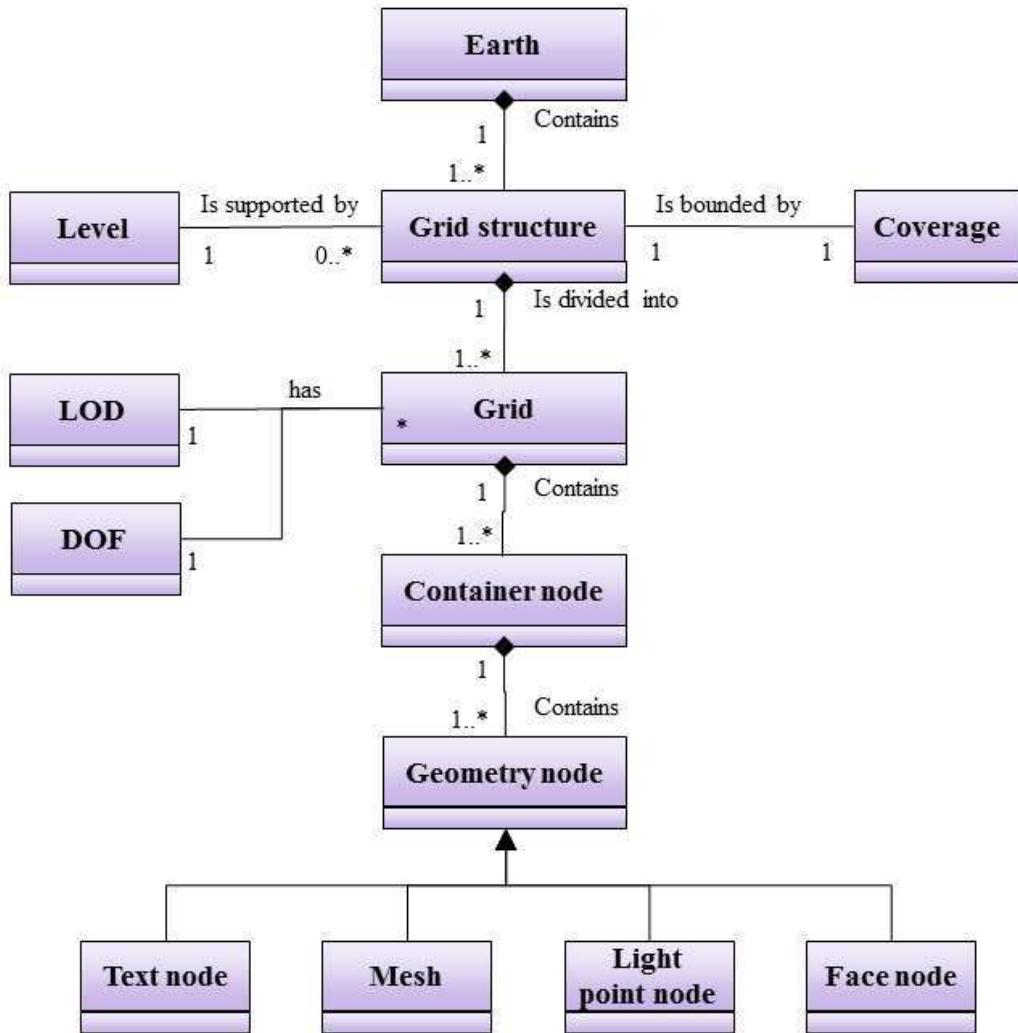


Figure 41: Simplified class diagram of an OpenFlight specification.

such as Clipmaps [64] for terrain rendering, Delaunay triangulation, particle systems, cloud rendering and marching cubes, just to mention a few. In this thesis, we use OSG features to build a 3D visualization tool that allows the user to fly over a topographical environment and visualize different weather forecast variables.

Chapter 2

3D Geovisualizer

2.1 Overview

Serious games technologies can be a suitable option for building weather analysis solutions. Reusing some features from engines and frameworks that were already created for game development can be profitable and can reduce implementation time and improve user experience. This involves an appropriated design of several components of the solution. We present our design choices for data interfaces, a topographic data model inspired on data models used in current flight simulators and a 3D visualization prototype that allows the exploration of topographic environment and weather data.

This chapter is organized as follows: section 2.2 presents our design choices for the topographic and object data model. Section 2.3 describes the implementation of the visualization prototype. Section 2.4 reviews some of the results from our 3D visualization prototype. Section 2.5 discusses limitations of our approach and future work.

2.2 Design

We present a 3D visualization application to explore the geographical information from different viewpoints. Our solution allows the user to fly over a

topographical area and interact with the data using first- and third-person views.

2.2.1 Data sources

The Open Geospatial Consortium (OGC) [87] proposes a set of technical documents and specifications grouped in a standard called OpenGIS [88] to encourage the integration of geographic information. The main goal of the standard is to generate open and public interfaces and encoding to store and exchange spatial data all over the world. OpenGIS implements Geography Markup Language (GML) [93] as a standard language to store and exchange geospatial information. GML is an XML grammar written in XML Schema for the description of application schemata as well as the transportation and storage of geographic information.

Also, Google has developed its own markup language called Keyhole Markup Language (KML) [36] to exchange and transport geographic data, for its product Google Earth [37]. KML is a language focused on data visualization including maps and annotations using the Internet. It has a wide and varied community of users, for example, casual users upload their information to the Internet using KML spatial references to their homes, to describe trips, or adventures. Also, scientists share information of models, climate phenomena, volcanic eruptions, weather patterns, earthquake activities, etc. Organizations such as National Geographic [82], UNESCO [125], and the Smithsonian Institute [109] use KML to display their rich sets of global data as services. KML 2.2 is now a part of GML. It uses certain elements derived from GML 2.1.2 [87].

We have analyzed OGC standards to represent geographical information and decided to use GML (Geographic Markup Language) [93]. GML provides the following options: GML schema and GML profiles. GML schema has a template and a closed syntax while GML profile allows for the adjustment of the semantic and syntax to the application that is using it. For our solution, we have adopted GML profiles.

2.2.2 Data model

Our approach provides an adequate modeling of global information and detailed information. It ranges from a large scale representation of the entire Earth surface to a small scale representation of geolocated point-based objects. The data model is divided in a “Topographic Environment Model (TEM)” that represents the Earth and an “Object Data Model (ODM)” that represents geolocated objects on the Earth. The TEM represents environments independently of their extension by using a large scale model of the overall terrain surface. It utilizes some of the concepts presented in the OpenFlight [97] de-facto standard and other widespread flight simulator SDKs.

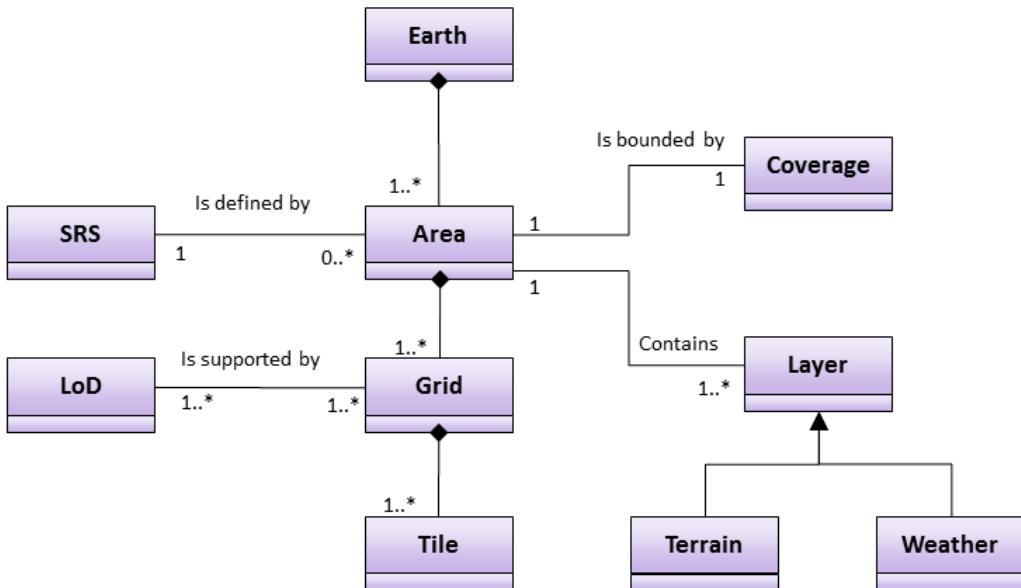


Figure 42: High level class diagram of the topographic environment data model of our 3D Geovisualizer.

Figure 42 shows a class diagram representing a subset of the TEM. It divides the Earth in a uniform grid with parallel lines to the latitude and longitude lines. The main idea is to factorize the Earth (Earth class) in a set of areas (Area class). Each area represents a tile (Tile class) in a grid (Grid class) with boundaries that are parallel to the latitude and longitude lines. An area has a covered surface extension (Coverage class) and a Spa-

tional Reference System (SRS class). The SRS class contains information about the coordinate system type, datum and projections. Each area contains information of different layers. Raster information about latitude, longitude and height is represented in a Terrain layer (Terrain class). The weather layer (Weather class) contains georeferenced data about weather (temperature, humidity, wind speed and direction). The model supports different Levels of Detail (LoD class), dividing each area in one or more reticulated grids. Each grid tile keeps information related to the initial position using the original geographic coordinates.

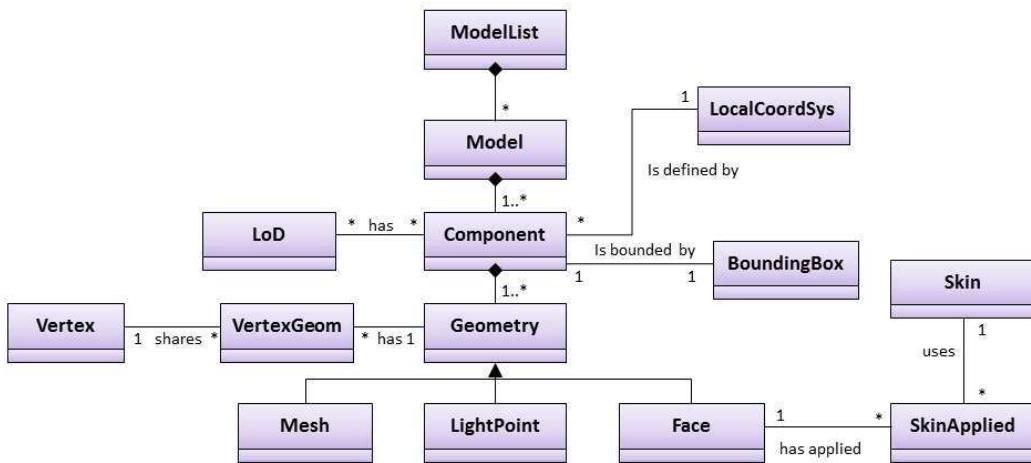


Figure 43: Objects in the scene represented in a class diagram.

Figure 43 shows a class diagram representing the scene's objects of the object data model. Each area has an associated model list (ModelList class) representing the objects that compose the scene contained on a given parcel of terrain. For this work, we have only modeled point-based geolocated objects. A model (Model class) is a high level object that can be composed of one or more components (Component class). Each component is associated with a local coordinate system (LocalCoordSys class) and a box or a minimal volume that includes it (BoundingBox class). The components are built from geometric structures (Geometry class) that include faces (Face class), meshes (Mesh class), or lighting points (LightPoint class). Each of the geometries is formed by one or more vertices (Vertex class) that store the position of the point in local coordinates. The Vertex class stores the vertices of each component. These vertices are shared among each of the geometries using

the Geometry Vertex class (VertexGeom class). The model incorporates the concept of level of detail by using the Skin class, the SkinApplied class and the local Level of Detail (LocalLoD) class. Each component is associated to one or more skins. Skins store the textures of each local Level of Detail (localLoD) that will later be applied on each face, depending on the distance to the observer viewpoint.

2.3 Implementation

There is a wide variety of commercial and open-source database management systems (DBMS) options that allow managing massive geographical data. These software applications include special functions, geospatial indexing and new data types for geospatial data management. Among these available options, we have chosen PostgreSQL [115] with PostGIS [94] extension. PostgreSQL is a well-known open-source software, and its PostGIS extension that is used to manage geospatial data is compliant with the OGC standards. The implementation of the visual application is based on OpenGL and Visual C++ technologies. It allows the representation of terrain elevation data, point-based objects and meteorological information.

2.4 Results

We have incorporated data from several sources such as topographic data that has been taken from the NASA's Shuttle Radar Topography Mission (SRTM) [81] (3 arc-sec of latitude by 3 arc-sec of longitude). In addition, a LANDSAT-7 7 [114] image (30m per pixel) was mapped over the topography. The application allows the user to overview the environment and recognize punctual glyphs and continuous layers representing geographical and ancillary information. One of them is a weather layer in which there are representations of weather measurement points positioned over the area. The local temperature is visualized over the topographic environment, where the local values are linearly interpolated using a Delaunay triangulation, as it is shown in Fig. 44 and Fig. 45.

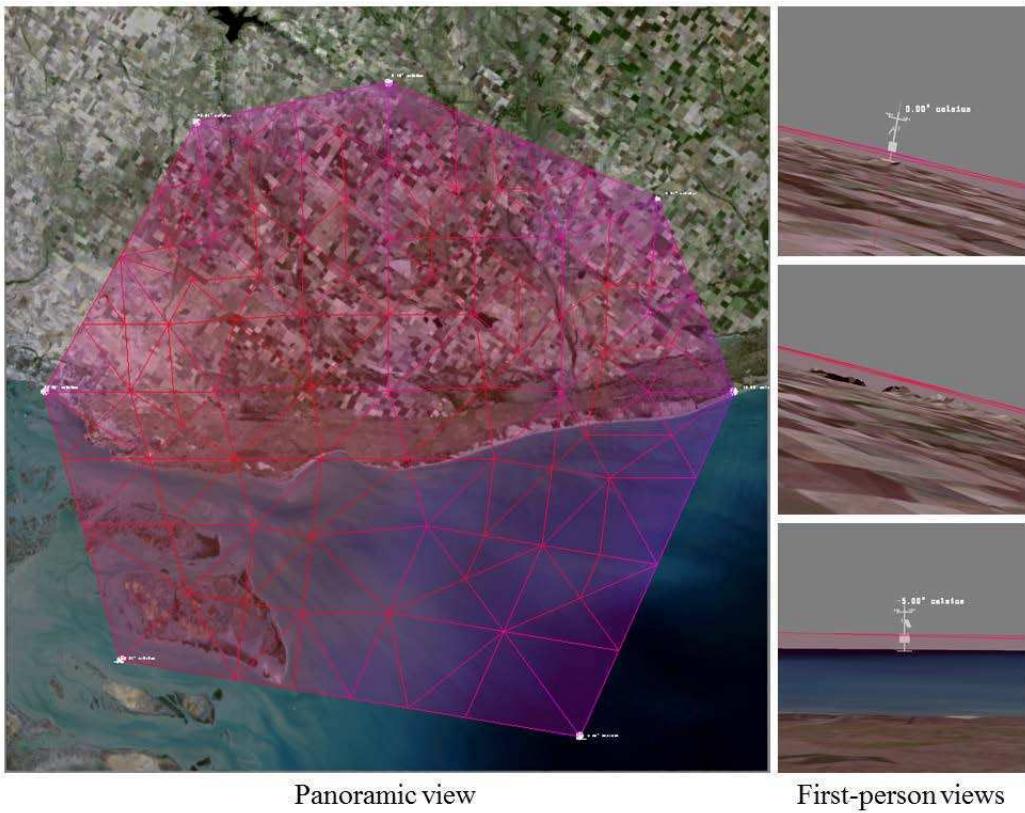


Figure 44: 3D Geovisualizer first- and third-person views. Visualization of the area of Bahía Blanca, Argentina.

The model described above presents several advantages. It is generic, since it facilitates the modeling of both simple and complex objects. It is also flexible, since it allows the user to specify the global reference system used in their database (global coordinate system, datum, projection) and a local coordinate system for each modeled object in the scene. Finally, it is extensible, since it allows adding new geometries that are inherited from the geometry entity without modifying existing entities and their relationships.

The use of PostGIS (GIS extension for PostgreSQL) to support geographical information allows us to validate and ensure compliance with the OGC standards. The performance of the database operations can be improved by using specific functions for geographical data, coordinate system conversions and spatial indexing.

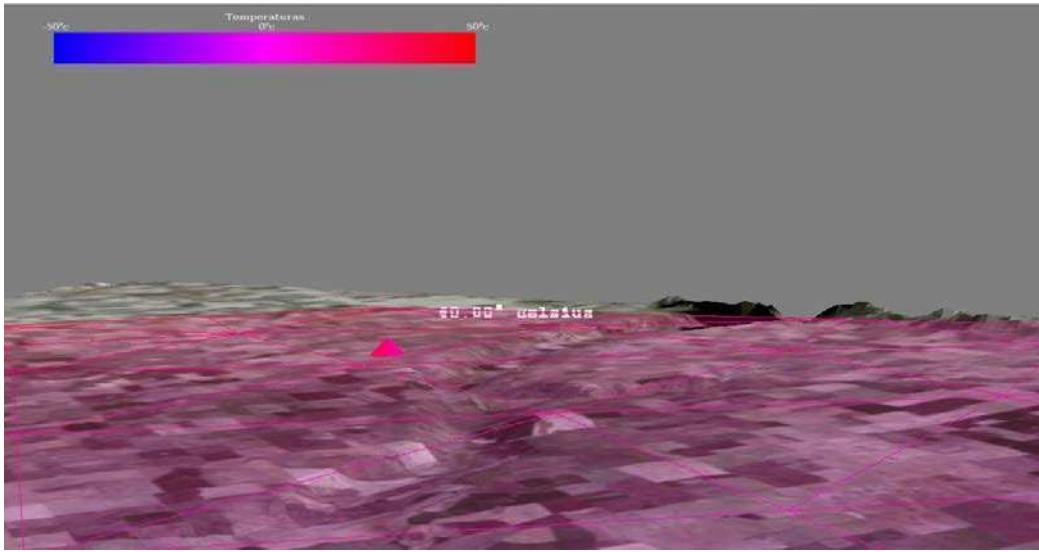


Figure 45: First-person overview of the area of Bahía Blanca, centered in (38.90S, 61.57W), Argentina.

2.5 Limitations and future work

Although we have generalized the data model and data interfaces, we still have opportunities to generalize the use of serious game engines. For a better evaluation of usability of serious games, we need a generic “adapter” between the engine and the geovisualization tool. This adapter would simplify the prototyping and construction of visualizations and reinforce usability. Our current prototype is using LANDSAT-7 images provided by the CONAE which have a low resolution. Although, high resolution images are not mandatory for weather analysis. We plan to include imagery with higher resolutions to provide a better visual appeal. Also, we are currently evaluating Delta3D [24], a game engine built on top of OSG to reuse further features provided by game engines.

Part IV

Conclusions

Reflections on our work

1.1 Overview

In this thesis we studied different designs of efficient visual analytics solutions for the area of geovisualization to support forecasters' analysis and decision-making on operational weather forecasting. We have also analyzed the use of existing game technologies for geovisualization purposes.

Previous literature on this subject posed the following questions and open problems:

- How does the domain of application influence the visualization design choices?
- How important is participating the users in that design choices?
- What is the associated cost of adapting generic tools or frameworks for a specific domain, in terms of learning curve, developing time and customization efforts?

The design studies performed during this thesis were related to these questions and tried to give answers to them and also to the following ones:

- How can we integrate the operational tasks workflow of forecasters with existing or new visual analytics solutions?
- How can we translate their operational mechanisms into visual analytics mechanisms?

- How can interactive visualizations be useful to make their tasks workflows more efficient?

We gave partial answers to these questions by applying a participatory design process of visual analytics tools to be incorporated in the users' tasks workflow. We embodied the lessons learned into a suggested guideline list to structure the process of visual analytics design for geovisualization:

1. Perform a thorough analysis of the domain experts tasks workflow.
2. Identify existing problems, gaps and improvement opportunities in their current workflow and tools.
3. Solve the specific user needs, taking advantage of their previous knowledge and experience in order to improve cognitive productivity.
4. Participate the users in all the stages of the process of design, by working in short iterations.

Also, this thesis raised challenging questions: Up to which point can results and visualizations rely on the users' experience? This concern gets increased importance in a scientific domain, in which scientists apply objective techniques and algorithms to measure their object of study and to validate their methods. In particular for the case of uncertainty visualization, how can we incorporate the user knowledge and experience to the visualization technique, given that the user experience itself is difficult to quantify and has an associated uncertainty?.

1.2 Main findings

The proposed visual analytics solutions assist the users in at least four main tasks: (1) identification and classification of meteorological phenomena, (2) analysis of systematic biases and NWP model errors, (3) analysis of uncertainty and (4) decision-making for operational weather forecasting.

First, we found a gap between existing visual analytics solutions and forecasters needs. Generic tools for operational weather forecasting require

complex settings, are time-consuming, and some of them require programming skills. We filled this gap by applying the visual analytics mantra, proposing VIDa and its component Semillón and intervening their current tasks workflow. The domain experts evaluated our approach by means of several case studies. They could perform successfully their tasks and gave us a positive feedback about the usability of our solution.

Second, we found that we could break-down the intermediate steps of the RAR technique for the analysis of probabilistic forecasts. We introduced a new interactive visual workflow, also following the visual analytics mantra, and providing them a three-step loops solution. Albero unfolds the internal steps and data used in the RAR technique by means of linked-views and interactive visualization methods combined with automatic processing. Our approach was also designed and evaluated as part of the participatory design process.

Finally, we performed a survey of some of the most popular game engines and proposed an integral solution of flight data models, standard game engine features, and scientific visualization to build geovisualization applications. We found that this approach can shorten the prototyping and development times, and could be used with great advantages by our research community to build visual tools.

1.3 Limitations

The precision of the data could be improved by using local data-sets coming from radars and weather stations in Argentina. Most of the issues come from the quality of the data and their scarcity. In the case of radar data, a standard still needs to be defined for data interfaces.

We also recognize the need for more evaluation of serious games applications in geovisualization. We are currently considering building 3D visualizations of storm cells and cloud tracking using game engines, and this would be a very interesting application of game technologies.

1.4 Future work

Future work will include a new series of evaluation tasks with the users for VIDa and its components. We will make user studies to have quantitative measures of usability, efficiency, and user experience.

Another upcoming challenge is to apply visual analytics for the tracking and analysis of forecasted phenomena such as intense precipitation, hail and severe storms.

In particular, for Albero, we plan to include the design of a new module for RAR technique validation that will be able to calculate quality measures, known as “skill”, using different verification scores. Additionally, we will include variations of the similarity function used by RAR, the addition of more meteorological variables, and new visualizations such as histograms, and extreme values summaries. Furthermore, a new module will be developed for parameter optimization, using for example genetic algorithms, to optimize the parameterization for a given region and situation. This shall reduce the time it takes researchers to optimize the RAR technique for that region and situation.

1.5 Open problems and perspective

We are facing a time of information overload, with huge amounts of data coming from different sources that are heterogeneous and have different resolutions in time and space. We have to cope with these enormous datasets, called “big data”, which poses enormous challenges for certain analyses and domains of application.

Big data will continue its tremendous growth, with technologies like wearable devices, Internet-of-Things, and the increasing number of computer devices available all around the world. However there are gaps in the data, missing and erroneous information with significant associated uncertainties. The next big challenge is to confront these data complexities. We have to take into account important aspects such as quality and scalability.

In particular, in meteorology this challenging undertaking could be tackled using promising and powerful computational analysis methods such as artificial intelligence techniques like deep learning, virtualization in the cloud, and non-relational databases. Faster supercomputers and optimized numerical models allow for better weather forecasts, but equally important is the availability of improved observational data, as these are the input of all the computational models.

Some of the complexities and issues that are present in current data are related to their source of origin:

- Data coming from simulations produced by numerical models: during the next years ensemble-numerical weather prediction will grow in precision and in size. Intelligent agents will help to tune forecasting models. Additionally visualizations will reveal new insights to meteorology experts, as well as reveal hidden data relationships to forecasters.
- Observations coming from:
 - Weather stations (surface and upper air) that have a good precision, but a coarse coverage. Interconnected networks of weather stations will become more affordable and usual in the near future.
 - Private weather stations networks that have to be integrated with the official networks.
 - Satellites that have denser coverage, less precision and complex relationships with variables of interest.
 - Radars that have an even denser coverage, good precision, but complex relationships with variables of interest.
 - Users' observations coming for example from social media. These kind of data sources provide subjective information that is denser but less structured and very heterogeneous.

Another important factor to take into account is the complexity associated with workflows and procedures resulting from the users' tasks. There

is a need for more studies in areas such as perception and cognition to understand how visualization influences the users' analysis and decision making capabilities. To find the appropriate balance between visual analysis and automated reasoning is a hot research topic for the visualization discipline. Important topics are to determine what is the relevant part of the data and how to show it.

1.6 Summary

We presented several design cases to the visualization community, where we have demonstrated the application of a participatory and interdisciplinary design process, centered in users' needs, following the concept of the visual analytics loop of knowledge and the visual analytics mantra. Our research shows evidence of efficient designs of geospatial visual analytics tools for meteorology, in particular for operational weather forecasting where quick decision-making is critical. The proposed guidelines for the design process of efficient visual analytics solutions are general and can be reproduced for other domains of application.

Part V

Appendices

Chapter **1**

Appendix A

The following table describes the meaning of various abbreviations, acronyms and symbols used throughout the thesis. The page on which each one is defined or first used is also given.

1.1 Abbreviations

Abbrev.	Meaning	Page
2D	2-dimensional	59
3D	3-dimensional	5
CIMA	Centro de Investigación del Mar y de la Atmósfera	53
DBMS	database management system	90
ETL	Extract-Transform-Load	89
FTP	File Transfer Protocol	89
GDAS	Global Data Assimilation System	102
GFS	Global Forecast System	81

Abbrev.	Meaning	Page
GI	Geographic Information	43
GPU	Graphics Processing Unit	50
GP GPU	General-Purpose computation on Graphics Processing Units	90
GrADS	Grid Analysis and Display System	56
GUI	Graphical User Interface	56
HPC	High Performance Computing	90
ICA	International Cartographic Association	43
IDV	Integrated Data Viewer	56
LCC	Lambert Conformal Conic	88
MCV	Multiple Coordinated Views	63
NSF	National Science Foundation	43
NWP	Numerical Weather Prediction	61
RAR	Reforecast Analog Regression	81
SMN	Servicio Metereológico Nacional	32
SOM	Self-Organizing Map	59
TAC	Time-Activity-Curves	60
UCAR	University Corporation for Atmospheric Research	56
UTC	Coordinated Universal Time	72
UTM	Universal Transverse Mercator	71
UV-CDAT	Ultrascale Visualization Climate Data Analysis Tools	56
VAPOR	Visualization and Analysis Platform for Ocean, Atmosphere, and Solar Researchers	56
VIDa	Visual Interactive Dashboard	50
WRF	Weather Research and Forecasting	64
WWT	Microsoft World Wide Telescope	48
YIQ	Luma In-phase Quadrature	100

1.2 Nomenclature

Symbol	Meaning	Page
<i>bias</i>	It is a measure of the forecast error commonly used in forecast verification analysis	80
Δ	It is a measure of the distance between the sketched curve-pattern and the temporal curves on each geographic location of the 2D scalar-field associated to the selected time-steps used in the curve-pattern selector	76
<i>lat</i>	latitude	78
<i>long</i>	longitude	78

Bibliography

- [1] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, *Visualization of Time-Oriented Data*. London, UK: Springer, 2011.
- [2] L. Aldeco, J. Ruiz, and C. Saulo, “Probabilistic forecasts during the monsoon season: the analogs technique as a tool for precipitation prediction over southeastern-south america.” in *X International Conference on Southern Hemisphere Meteorology and Oceanography*. American Meteorological Society, 2012.
- [3] E. F. Anderson, S. Engel, P. Comninou, and L. McLoughlin, “The case for research in game engine architecture,” in *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, ser. Future Play ’08. New York, NY, USA: ACM, 2008, pp. 228–231. [Online]. Available: <http://doi.acm.org/10.1145/1496984.1497031>
- [4] J. Andersson, “Terrain rendering in frostbite using procedural shader splatting,” in *ACM SIGGRAPH 2007 Courses*, ser. SIGGRAPH ’07. New York, NY, USA: ACM, 2007, pp. 38–58. [Online]. Available: <http://doi.acm.org/10.1145/1281500.1281668>
- [5] G. Andrienko, N. Andrienko, P. Jankowski, D. Keim, M. J. Kraak, A. MacEachren, and S. Wrobel, “Geovisual analytics for spatial decision support: Setting the research agenda,” *International Journal of Geographical Information Science*, vol. 21, no. 8, pp. 839–857, Jan. 2007. [Online]. Available: <http://dx.doi.org/10.1080/13658810701349011>

- [6] G. Andrienko, N. Andrienko, D. Keim, A. M. MacEachren, and S. Wrobel, "Editorial: Challenging problems of geospatial visual analytics," *Journal of Visual Languages and Computing*, vol. 22, no. 4, pp. 251–256, Aug. 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.jvlc.2011.04.001>
- [7] G. L. Andrienko, N. V. Andrienko, S. Bremm, T. Schreck, T. von Landesberger, P. Bak, and D. A. Keim, "Space-in-time and time-in-space self-organizing maps for exploring spatiotemporal patterns," *Computer Graphics Forum*, vol. 29, no. 3, pp. 913–922, Jun 2010. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-8659.2009.01664.x>
- [8] N. V. Andrienko and G. L. Andrienko, *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*, 2006th ed. Springer, 2006.
- [9] A. Asirvatham and H. Hoppe, "Terrain rendering using GPU-based geometry clipmaps," in *GPU Gems 2*, M. Pharr, Ed. Addison-Wesley, 2005, pp. 27–45.
- [10] BAMS, "Guidelines for using color to depict meteorological information: IIPS subcommittee for color guidelines," *Bulletin of the American Meteorological Society*, vol. 74, no. 9, pp. 1709–1713, Sep. 1993. [Online]. Available: [http://dx.doi.org/10.1175/1520-0477\(1993\)074<1709:GFUCTD>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1993)074<1709:GFUCTD>2.0.CO;2)
- [11] M. Basler, M. Spott, S. Buchanan, J. Berndt, B. Buckel, C. Moore, C. Olson, D. Perry, M. Selig, and D. Walisser, *The FlightGear Manual*, FlightGear, Mar 2015.
- [12] J. E. Bendis, "Developing educational virtual worlds with game engines," in *ACM SIGGRAPH 2007 Educators Program*, ser. SIGGRAPH '07. New York, NY, USA: ACM, 2007. [Online]. Available: <http://doi.acm.org/10.1145/1282040.1282068>

- [13] J. S. Berndt, *JSBSim: An open source, platform-independent, flight dynamics model in C++*, Jon S. Berndt, Jun 2011. [Online]. Available: <http://www.jsbsim.org/JSBSimReferenceManual.pdf>
- [14] L. Bishop, D. Eberly, T. Whitted, M. Finch, and M. Shantz, "Designing a PC game engine," *Computer Graphics and Applications, IEEE*, vol. 18, no. 1, pp. 46–53, Jan 1998.
- [15] S. Bruckner and T. Möller, "Result-driven exploration of simulation parameter spaces for visual effects design," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1467–1475, Oct 2010. [Online]. Available: <http://dblp.uni-trier.de/db/journals/tvcg/tvcg16.html#BrucknerM10>
- [16] P. Buono, A. Aris, C. Plaisant, A. Khella, and B. Shneiderman, "Interactive pattern search in time series," *Proceedings of SPIE*, vol. 5669, no. 2, pp. 175–186, Jan 2005. [Online]. Available: <http://www.cs.umd.edu/hcil/timesearcher/>
- [17] Case Western Reserve University, "Freedman Center website, accessed in Jul. 2015." [Online]. Available: <http://library.case.edu/ksl/freedmancenter/>
- [18] M. Cavazza, F. Charles, and S. J. Mead, "Emergent situations in interactive storytelling," in *Proceedings of the 2002 ACM Symposium on Applied Computing*, ser. SAC '02. New York, NY, USA: ACM, 2002, pp. 1080–1085. [Online]. Available: <http://doi.acm.org/10.1145/508791.509003>
- [19] R. W. Clark and R. Mauer, "Visual terrain editor: An interactive editor for real terrains," *Journal of Computing Sciences in Colleges*, vol. 22, no. 2, pp. 12–19, Dec. 2006. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1181901.1181904>
- [20] Climate Prediction Center, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, "NOAA CPC Morphing Technique (CMORPH) Global

- Precipitation Analyses," Boulder, CO, USA, 2011. [Online]. Available: <http://dx.doi.org/10.5065/D6CZ356W>
- [21] J. Clyne, P. Mininni, A. Norton, and M. Rast, "Interactive desktop analysis of high resolution simulations: application to turbulent plume dynamics and current sheet formation," *New Journal of Physics*, vol. 9, no. 8, p. 301, 2007. [Online]. Available: <http://stacks.iop.org/1367-2630/9/i=8/a=301>
- [22] E. Coto, S. Grimm, S. Bruckner, M. E. Gröller, A. Kanitsar, and O. Rodriguez, "Mammoexplorer: An advanced CAD application for breast DCE-MRI," in *Proceedings of Vision, Modelling, and Visualization 2005*, Nov 2005, pp. 91–98.
- [23] S. Deitrick and R. Edsall, "The influence of uncertainty visualization on decision making: An empirical evaluation," in *Progress in Spatial Data Handling*. Springer Berlin Heidelberg, 2006, pp. 719–738.
- [24] delta3D, "delta3d, accessed in Feb. 2016." [Online]. Available: <http://delta3d.org/>
- [25] Department of Geography, Hunter College, "Department of Geography, Hunter College, accessed in Feb. 2016." [Online]. Available: <http://www.geography.hunter.cuny.edu/>
- [26] A. Diehl, H. Abbate, C. Delrieux, and J. Gambini, "Integration of flight simulators and geographic databases," *CLEI Electronic Journal*, vol. 12, no. 3, pp. 1–8, 2009. [Online]. Available: <http://www.clei.org/cleiej/paper.php?id=173>
- [27] J. Dykes, S. Fabrikant, and J. Wood, "Preface," in *Exploring Geovisualization*, ser. International Cartographic Association, J. D. M. M.-J. Kraak, Ed. Oxford: Elsevier, 2005, pp. ix–xii. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B9780080445311504152>
- [28] J. A. Fails, A. Karlson, L. Shahamat, and B. Shneiderman, "A visual interface for multivariate temporal data: Finding patterns of events

- across multiple histories," in *IEEE Symposium on Visual Analytics Science and Technology*. IEEE Computer Society Press, Oct 2006, pp. 167–174.
- [29] Z. Fang, T. Möller, G. Hamarneh, and A. Celler, "Visualization and exploration of time-varying medical image data sets," in *Proceedings of Graphics Interface 2007*, ser. GI '07. New York, NY, USA: ACM, 2007, pp. 281–288. [Online]. Available: <http://doi.acm.org/10.1145/1268517.1268563>
- [30] M. Finch, "Effective water simulation from physical models," in *GPU Gems*, R. Fernando, Ed. Addison-Wesley, 2004, pp. 5–29.
- [31] FlightGear, "The FlightGear flight simulator, accessed in Jul. 2015." [Online]. Available: <http://www.flightgear.org/>
- [32] L. Gimeno, "Grand challenges in atmospheric science," *Frontiers in Earth Science*, vol. 1, no. 1, 2013. [Online]. Available: http://www.frontiersin.org/atmospheric_science/10.3389/feart.2013.00001/full
- [33] M. Glatter, J. Huang, S. Ahern, J. Daniel, and A. Lu, "Visualizing temporal patterns in large multivariate data using modified globbing," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1467–1474, Nov 2008. [Online]. Available: <http://dx.doi.org/10.1109/TVCG.2008.184>
- [34] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, "Visual comparison for information visualization," *Information Visualization*, vol. 10, no. 4, pp. 289–309, Oct 2011. [Online]. Available: <http://dx.doi.org/10.1177/1473871611416549>
- [35] Global Data Assimilation System, "Global Data Assimilation System, accessed in Nov. 2014." [Online]. Available: <http://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-data-assimilation-system-gdas>
- [36] Google, "Keyhole Markup Language, accessed in Aug. 2015." [Online]. Available: <https://developers.google.com/kml/?hl=en>

- [37] Google Earth, "Google Earth, accessed in Jul. 2015." [Online]. Available: <http://www.google.com/earth/index.html>
- [38] Google Maps, "Google maps, accessed in Jul. 2015." [Online]. Available: <https://maps.google.com/>
- [39] T. M. Hamill and J. S. Whitaker, "Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application," *Monthly Weather Review*, vol. 134, no. 11, pp. 3209–3229, 2006. [Online]. Available: http://www.esrl.noaa.gov/psd/people/tom.hamill/reforecast_analog_v2.pdf
- [40] C. Han and H. Hoppe, "Optimizing continuity in multiscale imagery," *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2010)*, vol. 29, no. 5, pp. 171:1–171:9, 2010.
- [41] M. Hasan, M. Sazzad Karim, and E. Ahmed, "Generating and rendering procedural clouds in real time on programmable 3D graphics hardware," in *9th International Multitopic Conference, IEEE INMIC 2005*, Dec 2005, pp. 1–6.
- [42] J. Hildebrandt, "Flight simulator as geospatial visualisation platform," in *6th International Command and Control Research and Technology Symposium*, 2001. [Online]. Available: http://www.dodccrp.org/events/6th_ICCRTS/Tracks/Papers/Track2/055_tr2.pdf
- [43] H. Hochheiser and B. Shneiderman, "Dynamic query tools for time series data sets: Timebox widgets for interactive exploration," *Information Visualization*, vol. 3, no. 1, pp. 1–18, Mar 2004. [Online]. Available: <http://dx.doi.org/10.1145/993176.993177>
- [44] id Software, "Quake III Arena, accessed in Jul. 2015." [Online]. Available: <http://www.idsoftware.com/en-gb>
- [45] International Cartographic Association, "International Cartographic Association (ICA), accessed in Aug. 2015." [Online]. Available: <http://icaci.org/>

- [46] H. Jänicke, M. Böttinger, and G. Scheuermann, "Brushing of attribute clouds for the visualization of multivariate data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1459–1466, Nov 2008. [Online]. Available: <http://dx.doi.org/10.1109/TVCG.2008.116>
- [47] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science, 2nd Edition.* John Wiley and Sons Ltd, 2011.
- [48] A. Jones and D. Cornford, "Advanced data driven visualisation for geo-spatial data," in *Computational Science – ICCS 2006*, ser. Lecture Notes in Computer Science, V. Alexandrov, G. van Albada, P. Sloot, and J. Dongarra, Eds. Springer Berlin Heidelberg, 2006, vol. 3993, pp. 586–592. [Online]. Available: http://dx.doi.org/10.1007/11758532_77
- [49] JSBSim, "JSBSim, accessed in Jan. 2016." [Online]. Available: <http://jsbsim.sourceforge.net/>
- [50] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 2003. [Online]. Available: <https://books.google.com.ar/books?id=Uqc7zC7NULMC>
- [51] F.-E. W. Karl-Ingo Friese, Marc Herrlich, "Using game engines for visualization in scientific applications," in *New Frontiers for Entertainment Computing*, ser. IFIP International Federation for Information Processing, P. Ciancarini, R. Nakatsu, M. Rautenberg, and M. Roccati, Eds., vol. 279. Boston, USA: Springer, 2008, pp. 11–22. [Online]. Available: http://www.springerlink.com/content/v1736j5754533h74/_blank
- [52] J. Kehrer and H. Hauser, "Visualization and visual analysis of multifaceted scientific data: A survey," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 3, pp. 495–513, Mar 2013. [Online]. Available: <http://dx.doi.org/10.1109/TVCG.2012.110>

- [53] J. Kehrer, F. Ladstädter, P. Muigg, H. Doleisch, A. Steiner, and H. Hauser, "Hypothesis generation in climate research with interactive visual data exploration," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1579–1586, Oct 2008. [Online]. Available: <http://dx.doi.org/10.1109/TVCG.2008.139>
- [54] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, "Visual analytics: Definition, process, and challenges," *Information Visualization*, pp. 154–175, 2008. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-70956-5_7
- [55] D. A. Keim, F. Mansmann, and J. Thomas, "Visual analytics: How much visualization and how much analytics?" *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 2, pp. 5–8, May 2010. [Online]. Available: <http://doi.acm.org/10.1145/1809400.1809403>
- [56] H. Kelly, K. Howell, E. Glinert, L. Holding, C. Swain, A. Burrowbridge, and M. Roper, "How to build serious games," *Communications of the ACM*, vol. 50, no. 7, pp. 44–49, Jul. 2007. [Online]. Available: <http://doi.acm.org/10.1145/1272516.1272538>
- [57] J. Kienzle, C. Verbrugge, B. Kemme, A. Denault, and M. Hawker, "Mammoth: A massively multiplayer game research framework," in *Proceedings of the 4th International Conference on Foundations of Digital Games*, ser. FDG '09. New York, NY, USA: ACM, 2009, pp. 308–315. [Online]. Available: <http://doi.acm.org/10.1145/1536513.1536566>
- [58] Z. Konyha, A. Lež, K. Matković, M. Jelović, and H. Hauser, "Interactive visual analysis of families of curves using data aggregation and derivation," in *Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies*, ser. i-KNOW '12. New York, NY, USA: ACM, Sep 2012, pp. 24:1–24:8. [Online]. Available: <http://doi.acm.org/10.1145/2362456.2362487>
- [59] P. Köthur, M. Sips, A. Unger, J. Kuhlmann, and D. Dransch, "Interactive visual summaries for detection and assessment of

- spatiotemporal patterns in geospatial time series," *Information Visualization*, vol. 13, no. 3, pp. 283–298, 2013. [Online]. Available: <http://dblp.uni-trier.de/db/journals/ivs/ivs13.html#KothurSUKD14>
- [60] M. Krstajic, E. Bertini, and D. A. Keim, "CloudLines: Compact Display of Event Episodes in Multiple Time-Series," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2432–2439, Dec 2011. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22034364>
- [61] R. Laramee and R. Kosara, "Challenges and unsolved problems," in *Human-Centered Visualization Environments*, ser. Lecture Notes in Computer Science, A. Kerren, A. Ebert, and J. Meyer, Eds. Springer Berlin Heidelberg, 2007, vol. 4417, pp. 231–254. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-71949-6_5
- [62] G. Lepouras and C. Vassilakis, "Virtual museums for all: Employing game technology for edutainment," *Virtual Reality*, vol. 8, no. 2, pp. 96–106, Sep. 2004. [Online]. Available: <http://dx.doi.org/10.1007/s10055-004-0141-1>
- [63] M. Lewis, J. Jacobson, M. Anandarajan, and Association for Computing Machinery, *Game Engines in Scientific Research*, ser. Communications of the ACM. ACM, 2001.
- [64] F. Losasso and H. Hoppe, "Geometry clipmaps: Terrain rendering using nested regular grids," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 769–776, Aug. 2004. [Online]. Available: <http://doi.acm.org/10.1145/1015706.1015799>
- [65] A. M. MacEachren and M.-J. Kraak, "Research challenges in geovisualization," *Cartography and Geographic Information Science*, vol. 28, no. 1, pp. 3–12, 2001. [Online]. Available: <http://dx.doi.org/10.1559/152304001782173970>

- [66] M. M. Malik, C. Heinzl, and M. E. Gröller, "Comparative visualization for parameter studies of dataset series," *IEEE Transaction on Visualization and Computer Graphics*, vol. 16, no. 5, pp. 829–840, Sep 2010.
- [67] S. Mantler, G. Hesina, M. Greiner, and W. Purgathofer, "GEARViewer: A state of the art real-time geospatial visualization framework," in *Proceedings of CORP 2011*, P. Z. Manfred SCHRENK, Vasily V. POPOVICH, Ed., 2011, pp. 345–354. [Online]. Available: <https://www.cg.tuwien.ac.at/research/publications/2011/Mantler-2011-GEAR/>
- [68] S. Mariani, M. Casaioli, M. Calza, and G. Futura, *Forecast Verification: A Summary of Common Approaches and Examples of Application*, ser. Foralps Technical Reports. Trento (Italy): Università di Trento. Dipartimento di ingegneria civile e ambientale, 2008.
- [69] D. M. Mark, "Geographic information science: Critical issues in an emerging cross-disciplinary research domain," *URISA Journal*, vol. 12, no. 1, pp. 45–54, Jan 2000. [Online]. Available: http://dusk.geo.orst.edu/Pickup/AAG2010/Mark_URISA.pdf
- [70] R. Marroquim and A. Maximo, "Introduction to GPU programming with GLSL," in *Proceedings of the 2009 Tutorials of the XXII Brazilian Symposium on Computer Graphics and Image Processing*, ser. SIBGRAPI-TUTORIALS '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 3–16. [Online]. Available: <http://dx.doi.org/10.1109/SIBGRAPI-Tutorials.2009.9>
- [71] P. Martz, *OpenSceneGraph Quick Start Guide: A Quick Introduction to the Cross-platform Open Source Scene Graph API*, ser. The OpenSceneGraph programming series. Skew Matrix Software, 2007.
- [72] S. McDonald and R. D. Stevenson, "Visualizing geo-referenced data with the eco flight simulator," in *Proceedings of the 2002 Annual National Conference on Digital Government Research*, ser. dg.o '02.

- Digital Government Society of North America, 2002, pp. 1–4. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1123098.1123128>
- [73] A. Meyer, *X-Plane Operation Manual*, Laminar Research, Nov 2014.
- [74] Microsoft Bing Maps, “Microsoft Bing Maps, accessed in Jul. 2015.” [Online]. Available: <https://www.bing.com/maps/>
- [75] Microsoft Corporation, “Bing Maps Tile System, accessed in Aug. 2015.” [Online]. Available: <https://msdn.microsoft.com/en-us/library/bb259689.aspx>
- [76] Microsoft Flight, “Microsoft Flight, accessed in Jul. 2015.” [Online]. Available: <http://www.microsoft.com/games/flight/>
- [77] Microsoft Research, “Microsoft FetchClimate, accessed in Apr. 2014.” [Online]. Available: <http://fetchclimate.cloudapp.net/>
- [78] M. S. Monmonier, *How to Lie with Maps*. University of Chicago Press, 1991. [Online]. Available: <https://books.google.gl/books?id=7BS0QgAACAAJ>
- [79] T. Munzner, *Visualization Analysis and Design*, ser. A.K. Peters visualization series. A K Peters, 2014. [Online]. Available: <http://www.crcpress.com/product/isbn/9781466508910>
- [80] D. Murray, J. McWhirter, S. Wier, and S. Emmerson, “The integrated data viewer: a web-enabled application for scientific analysis and visualization,” in *19th Conference on International Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*. American Meteorological Society, 2003. [Online]. Available: <http://doi.org/10.5065/D6RN35XM>
- [81] NASA Shuttle Radar Topography Mission, “SRTM, accessed in Feb. 2016.” [Online]. Available: <http://www2.jpl.nasa.gov/srtm/>
- [82] National Geographic, “National Geographic, accessed in Aug. 2015.” [Online]. Available: <http://www.nationalgeographic.com/>

- [83] National Science Foundation (NSF), "National Science Foundation (NSF), accessed in Aug. 2015." [Online]. Available: <http://www.nsf.gov/>
- [84] NOAA Earth System Research Lab, Physical Sciences Division, Boulder, Colorado, USA, "A Description of the 2nd-Generation NOAA Global Ensemble Reforecast Data Set, accessed in Dec. 2015." [Online]. Available: <http://www.esrl.noaa.gov/psd/forecasts/reforecast2/>
- [85] T. Nocke, M. Flechsig, and U. Bohm, "Visual exploration and evaluation of climate-related simulation data," in *Simulation Conference, 2007 Winter*, S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, Eds. WSC, Dec 2007, pp. 703–711. [Online]. Available: <http://dblp.uni-trier.de/db/conf/wsc/wsc2007.html#NockeFB07>
- [86] J. Olsen, "Real time procedural terrain generation," Department of Mathematics and Computer Science, University of Southern Denmark, Tech. Rep., Dec 2004. [Online]. Available: <http://web.mit.edu/cesium/Public/terrain.pdf>
- [87] Open Geospatial Consortium, "OGC, accessed in Aug. 2015." [Online]. Available: <http://www.opengeospatial.org/>
- [88] ——, *The OGC Abstract Specification*, Open Geospatial Consortium Inc, Oct 2004. [Online]. Available: <http://www.opengeospatial.org/docs/as>
- [89] OpenSceneGraph, "The OpenSceneGraph Project website, accessed in Jan. 2016." [Online]. Available: <http://www.openscenegraph.org/>
- [90] P. Paar, "Landscape visualizations: Applications and requirements of 3D visualization software for environmental planning," *Computers, Environment and Urban Systems*, vol. 30, no. 6, pp. 815–839, 2006. [Online]. Available: <http://dblp.uni-trier.de/db/journals/urban/urban30.html#Paar06>

- [91] T. Palmer and R. Hagedorn, *Predictability of Weather and Climate*. Cambridge University Press, 2006.
- [92] K. Perlin, "Implementing improved perlin noise," in *GPU Gems*, R. Fernando, Ed. Addison-Wesley, 2004, pp. 73–85.
- [93] C. Portele, *OpenGIS Geography Markup Language (GML) Encoding Standard*, Open Geospatial Consortium Inc., 2007. [Online]. Available: <http://www.opengeospatial.org/standards/gml>
- [94] PostGIS, "PostGIS: Spatial and Geographic objects for PostgreSQL, accessed in Aug. 2015." [Online]. Available: <http://postgis.net/>
- [95] K. Potter, P. Rosen, and C. R. Johnson, "From quantification to visualization: A taxonomy of uncertainty visualization approaches," in *Uncertainty Quantification in Scientific Computing*. Springer, 2012, pp. 226–249.
- [96] K. Potter, A. Wilson, P.-T. Bremer, D. Williams, C. Doutriaux, V. Pasucci, and C. R. Johhson, "Ensemble-vis: A framework for the statistical visualization of ensemble data," in *IEEE Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes.*, Oct 2009, pp. 233–240.
- [97] Presagis, *OpenFlight Scene Description Database Specification*, Presagis, Jun 2009. [Online]. Available: <http://www.presagis.com/files/standards/OpenFlight16.4.pdf>
- [98] P. S. Quinan and M. Meyer, "Visually comparing weather features in forecasts," *IEEE Transactions on Visualization and Computer Graphics (InfoVis '15)*, vol. 22, no. 1, pp. 389–398, Jan 2016. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/26390490>
- [99] T. M. Rhyne and A. MacEachren, "Visualizing geospatial data," in *ACM SIGGRAPH 2004 Course Notes*, ser. SIGGRAPH '04. New York, NY, USA: ACM, 2004. [Online]. Available: <http://doi.acm.org/10.1145/1103900.1103931>

- [100] J. C. Roberts, "State of the art: Coordinated & multiple views in exploratory visualization," in *Proceedings of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization*, ser. CMV '07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 61–71. [Online]. Available: <http://dx.doi.org/10.1109/CMV.2007.20>
- [101] R. Rozumalski, *WRF Environmental Modeling System - User's Guide*, National Weather Service SOO Science and Training Resource Center, May 2006.
- [102] J. J. Ruiz, C. Saulo, and J. Nogués-Paegle, "WRF model sensitivity to choice of parameterization over south america: Validation against surface variables," *Monthly Weather Review*, vol. 138, no. 8, pp. 3342–3355, Aug 2010. [Online]. Available: <http://dx.doi.org/10.1175/2010MWR3358.1>
- [103] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. Keim, "Knowledge generation model for visual analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1604–1613, Dec 2014. [Online]. Available: <https://www.computer.org/csdl/trans/tg/preprint/06875967.pdf>
- [104] P. Salio, M. P. Hobouchian, Y. G. Skabar, and D. A. Vila, "Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network," *Atmospheric Research*, vol. 163, no. S1, pp. 146–161, Sep 2015. [Online]. Available: <http://urlib.net/sid.inpe.br/mtc-m21b/2015/09.23.12.47>
- [105] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. Moorhead, "Noodles: A tool for visualization of numerical weather model ensemble uncertainty," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1421–1430, Nov 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20975183>
- [106] J. Schmidt, M. E. Gröller, and S. Bruckner, "VAICo: Visual analysis for image comparison," *IEEE Transactions on Visualization and Computer*

- Graphics*, vol. 19, no. 12, pp. 2090–2099, Dec 2013. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24051775>
- [107] M. Sedlmair, M. Meyer, and T. Munzner, “Design study methodology: Reflections from the trenches and the stacks,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2431–2440, 2012. [Online]. Available: <http://dblp.uni-trier.de/db/journals/tvcg/tvcg18.html#SedlmairMM12>
- [108] B. Shneiderman, “The eyes have it: A task by data type taxonomy for information visualizations,” in *Proceedings of the 1996 IEEE Symposium on Visual Languages*, ser. VL ’96. Washington, DC, USA: IEEE Computer Society, 1996, pp. 336–343. [Online]. Available: <http://dl.acm.org/citation.cfm?id=832277.834354>
- [109] Smithsonian Institute, “Smithsonian Institute website, accessed in Aug. 2015.” [Online]. Available: <http://www.si.edu/>
- [110] Y. Song, J. Ye, N. Svakhine, S. Lasher-Trapp, M. Baldwin, and D. Ebert, “An atmospheric visual analysis and exploration system,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 1157–1164, Sep 2006. [Online]. Available: <http://dx.doi.org/10.1109/TVCG.2006.117>
- [111] R. Stauffer, G. J. Mayr, M. Dabernig, and A. Zeileis, “Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations,” *Bulletin of the American Meteorological Society*, 37 2013. [Online]. Available: <http://EconPapers.repec.org/RePEc:inn:wpaper:2013-37>
- [112] Terra3D, “Terra3D, accessed in Apr. 2014.” [Online]. Available: <http://www.terra3d.de/>
- [113] TerraGear, “TerraGear for FlightGear, accessed in Jul. 2015.” [Online]. Available: <http://wiki.flightgear.org/TerraGear>
- [114] The Lansat Program, “The Lansat Program, accessed in Aug. 2015.” [Online]. Available: <http://www.flightgear.org/>

- [115] The PostgreSQL Global Development Group, "PostgreSQL, accessed in Aug. 2015." [Online]. Available: <http://www.postgresql.org/>
- [116] The Weather Channel, "The Weather Channel, accessed in Apr. 2014." [Online]. Available: <http://www.weather.com/weather/map/interactive>
- [117] J. Thomson, E. Hetzler, A. MacEachren, M. Gahegan, and M. Pavel, "A typology for visualizing uncertainty," in *Electronic Imaging 2005*, vol. 5669. International Society for Optics and Photonics, 2005, pp. 146–157. [Online]. Available: <http://dx.doi.org/10.1117/12.587254>
- [118] T. Torsney-weir, M. Sedlmair, and T. Möller, "Decision making in uncertainty visualization," 2015.
- [119] M. Tory, "User studies in visualization: A reflection on methods," in *Handbook of Human Centric Visualization*, W. Huang, Ed. Springer New York, 2014, pp. 411–426. [Online]. Available: http://dx.doi.org/10.1007/978-1-4614-7485-2_16
- [120] M. Tory and T. Möller, "Human factors in visualization research," *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, no. 1, pp. 72–84, Jan. 2004. [Online]. Available: <http://dx.doi.org/10.1109/TVCG.2004.1260759>
- [121] M. Tory and T. Moller, "Rethinking visualization: A high-level taxonomy," in *Proceedings of the IEEE Symposium on Information Visualization*, ser. INFOVIS '04. Washington, DC, USA: IEEE Computer Society, 2004, pp. 151–158. [Online]. Available: <http://dx.doi.org/10.1109/INFOVIS.2004.59>
- [122] D. Trenholme and S. P. Smith, "Computer game engines for developing first-person virtual environments," *Virtual Reality*, vol. 12, no. 3, pp. 181–187, Aug. 2008. [Online]. Available: <http://dx.doi.org/10.1007/s10055-008-0092-z>
- [123] P. Tsai and B. Doty, "A Prototype Java Interface for the Grid Analysis and Display System (GrADS)," in *Proceedings of 14th Internationalall*

- Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, 1998, pp. 11–16.
- [124] E. R. Tufte, *Beautiful Evidence*. Cheshire, CT: Graphics Press, 2006.
- [125] United Nations Educational Scientific and Cultural Organization, “UNESCO, accessed in Aug. 2015.” [Online]. Available: <http://en.unesco.org/>
- [126] Unreal Engine, “Unreal Engine, accessed in Jan. 2016.” [Online]. Available: <https://www.unrealengine.com/blog>
- [127] Unreal Tournament, “Unreal Tournament, accessed in Jul. 2015.” [Online]. Available: <https://www.unrealtournament.com/>
- [128] Valve, “Source Engine from Valve, accessed in Jul. 2015.” [Online]. Available: <http://www.valvesoftware.com>
- [129] K. Virrantaus, D. Fairbairn, and M.-J. Kraak, “ICA research agenda on cartography and GI science,” *The Cartographic Journal*, vol. 46, no. 2, pp. 63–75, 2009. [Online]. Available: <http://dblp.uni-trier.de/db/journals/cartographica/cartographica44.html#VirrantausFK09>
- [130] N. Wang, “Realistic and fast cloud rendering in computer games,” in *ACM SIGGRAPH 2003 Sketches & Applications*, ser. SIGGRAPH ’03, A. P. Rockwood, Ed. New York, NY, USA: ACM, 2003, pp. 1–1. [Online]. Available: <http://doi.acm.org/10.1145/965400.965539>
- [131] C. Ware, *Information Visualization: Perception for Design*, ser. Information Visualization: Perception for Design. Morgan Kaufmann Publishers Inc., 2013.
- [132] C. Ware and M. D. Plumlee, “Designing a better weather display,” *Information Visualization*, vol. 12, no. 3-4, pp. 221–239, 2013. [Online]. Available: <http://dblp.uni-trier.de/db/journals/ivs/ivs12.html#WareP13>
- [133] Weather Spark, “Weather Spark, accessed in Apr. 2014.” [Online]. Available: <http://weatherspark.com/>

- [134] Weather Underground, "Weather Underground, accessed in Apr. 2014." [Online]. Available: <http://www.wunderground.com/>
- [135] Wetter.de, "Wetter.de, accessed in Apr. 2014." [Online]. Available: <http://www.wetter.de/>
- [136] D. Williams, C. Doutriaux, J. Patchett, S. Williams, G. Shipman, R. Miller, C. Steed, H. Krishnan, C. Silva, A. Chaudhary, P.-T. Bremer, D. Pugmire, W. Bethel, H. Childs, M. Prabhat, B. Geveci, A. Bauer, A. Pletzer, J. Poco, T. Ellqvist, E. Santos, G. Potter, B. Smith, T. Maxwell, D. Kindig, and D. Koop, "The Ultra-scale Visualization Climate Data Analysis Tools (UV-CDAT): Data Analysis and Visualization for Geoscience Data," *IEEE Computer*, vol. 99, p. 1, 2013.
- [137] J. Woodring and H. Shen, "Multiscale time activity data exploration via temporal clustering visualization spreadsheet," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 1, pp. 123–137, 2009. [Online]. Available: <http://dblp.uni-trier.de/db/journals/tvcg/tvcg15.html#WoodringS09>
- [138] M. A. Zielke, M. J. Evans, F. Dufour, T. V. Christopher, J. K. Donahue, P. Johnson, E. B. Jennings, B. S. Friedman, P. L. Ounekeo, and R. Flores, "Serious games for immersive cultural training: Creating a living world," *IEEE Computer Graphics and Applications*, vol. 29, no. 2, pp. 49–60, Mar. 2009. [Online]. Available: <http://dx.doi.org/10.1109/MCG.2009.30>
- [139] M. Zyda, "From visual simulation to virtual reality to games," *IEEE Computer*, vol. 38, no. 9, pp. 25–32, Sep. 2005. [Online]. Available: <http://dx.doi.org/10.1109/MC.2005.297>
- [140] M. Zyskowski, "Aircraft simulation techniques used in low-cost, commercial software," in *AIAA Modeling and Simulation Technologies Conference and Exhibit*. Austin, TX, USA: American Institute of Aeronautics and Astronautics, 2003. [Online]. Available: <http://dx.doi.org/10.2514/6.2003-5818>