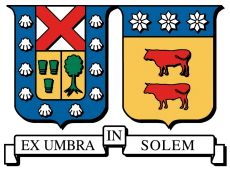
UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE ELECTRÓNICA

SANTIAGO – CHILE



SISTEMA DE OBTENCIÓN Y PROCESAMIENTO DE DATOS CLIMÁTICOS PARA LA PREDICCIÓN DE HELADAS EN AGRICULTURA INTELIGENTE

VICENTE JOSÉ OLMOS RUIZ

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN TELEMÁTICA

PROFESOR GUÍA

NICOLÁS SEBASTIÁN GÁLVEZ RAMÍREZ

PROFESOR CORREFERENTE

NICOLÁS TORRES RUDLOFF

2026

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ■ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Sistema de obtención y procesamiento de datos climáticos para la predicción de heladas en agricultura inteligente

Nombre del candidato(a): Vicente José Olmos Ruiz

Carrera / Grado: Ingeniería Civil Telemática

Campus: San Joaquin **Departamento:** Electrónica

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Nicolás Sebastián Gálvez Ramírez, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

X El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por **(marcar una opción)**:

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

_____

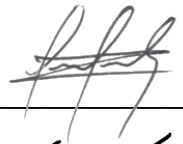
_____

_____

4.- FIRMAS

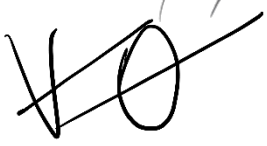
Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 29/10/2026 **Firma:** _____



Estudiante o Candidato(a):

Fecha: 26/10/2026 **Firma:**



Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Índice general

1. Introducción **1**

1.1. Propósito 4

1.2. Objetivos 4

1.2.1. Objetivo general 4

1.2.2. Objetivos específicos 4

1.3. Contribuciones 5

1.4. Estructura del informe 5

2. Contextualización y definición del problema **7**

2.1. Eventos meteorológicos y riesgo agrícola a nivel global 7

2.2. Caracterización operacional de las heladas 10

2.3. De los datos climáticos a la decisión agrícola 12

2.4. Brecha digital rural y adopción tecnológica 14

2.5. Situación en América Latina 16

2.6. Contexto chileno: datos disponibles, fragmentación y brecha de uso . . 18

2.7. Dimensiones técnicas del problema en Chile 21

2.8. Causas y consecuencias del problema 23

2.8.1. Consecuencias de no abordar el problema 23

2.9. Delimitación del problema 24

3. Estado del arte **26**

3.1. Obtención e integración de datos meteorológicos 26

3.2. Procesamiento y generación de predicciones meteorológicas 27
3.2.1. Pronóstico global, predicción inmediata y fenómenos localizados 28
3.2.2. Predicción de heladas en la literatura internacional 29
3.3. Interpretación y métricas de evaluación 30
3.4. Comunicación de predicciones y alertas meteorológicas 33
3.5. Sistemas existentes en Chile 37
3.6. Hacia la integración de soluciones 39

4. Propuesta y diseño de solución 40
4.1. Requisitos del sistema 40
4.1.1. Requisitos funcionales 41
4.1.2. Requisitos no funcionales 41
4.1.3. Requisitos de la interfaz 42
4.1.4. Requisitos de ambiente 43
4.2. Diseño de la arquitectura del sistema 44
4.2.1. Diagrama de contexto del sistema 44
4.2.2. Diagrama de arquitectura 45
4.3. Módulos del sistema 45
4.3.1. M1: Recolección de datos climáticos 46
4.3.2. M2: Procesamiento y limpieza de datos 46
4.3.3. M3: Modelo predictivo 47
4.3.4. M4: Integración conversacional con un servicio de mensajería . . 49
4.3.5. M5: Selección de parámetros 50
4.3.6. M6: Validación del sistema 50
4.3.7. Descripción general de módulos 51
4.3.8. Matriz de requisitos funcionales y módulos 52
4.3.9. Matriz de requisitos no funcionales y módulos 53
4.4. Diseño de interfaces 53
4.4.1. Modelo de navegación 54

4.4.2. Diseño de interfaces usuarias 56
4.5. Síntesis de la propuesta 56

5. Implementación de la solución 57
5.1. Marco General 57
5.1.1. Hitos principales del proyecto 57
5.2. Obtención de datos 58
5.2.1. Pronóstico de referencia de MeteoChile 59
5.3. Estrategia de extracción: API, archivos y programas de extracción . . . 59
5.3.1. Llamadas a API 59
5.3.2. Descarga de archivos masivos 60
5.3.3. Programas de extracción 60
5.4. Flujo de datos 61
5.5. Limpieza y estandarización 62
5.6. Construcción del conjunto de datos para heladas 64
5.7. Predicción e integración con el sistema 65
5.7.1. Integración con Miido 70
5.7.2. Integración con WhatsApp 71
5.7.3. Prototipos de interacción del caso de estudio 73
5.8. Resultados preliminares 76
5.8.1. Verificación y Validación 76
5.9. Decisiones de arquitectura 80
5.10. Limitaciones actuales 81
5.10.1. Revisión de requerimientos 81
5.10.2. Revisión del diseño 81
5.10.3. Análisis crítico 82
5.11. Gestión de riesgos 82
5.11.1. Supuestos 82
5.11.2. Dependencias 83

5.11.3. Restricciones 83
5.11.4. Evaluación de riesgos 83

6. Conclusiones **84**
6.1. Cumplimiento de los objetivos 84
6.2. Limitaciones y trabajo futuro 85
6.3. Comentarios del autor 86

Anexo **87**

Referencias **89**

Capítulo 1

Introducción

La agricultura es una actividad productiva profundamente expuesta a la variabilidad climática. A diferencia de otros sectores, donde muchas decisiones operacionales pueden aislarse parcialmente del entorno, la producción agrícola depende de condiciones meteorológicas que cambian en escalas horarias, diarias y estacionales. La temperatura, la humedad, el viento, la radiación, la precipitación y la disponibilidad hídrica influyen directamente en el desarrollo de los cultivos, la aparición de enfermedades, la planificación de labores y la ocurrencia de eventos meteorológicos extremos. Las sequías reducen la disponibilidad de agua; las lluvias intensas y las inundaciones pueden dañar suelos, caminos e infraestructura; los granizos y vientos fuertes provocan daño físico en plantas y frutos; las olas de calor aumentan el estrés hídrico y térmico; y las heladas pueden afectar tejidos vegetales sensibles durante pocas horas. Estos fenómenos constituyen amenazas capaces de reducir rendimientos, interrumpir cadenas productivas y afectar la seguridad alimentaria [1, 2].

Por esta razón, una parte importante de la gestión agrícola consiste en tomar decisiones bajo incertidumbre, con información incompleta y, muchas veces, con poco tiempo para reaccionar. Dentro de estos eventos, las heladas constituyen una amenaza particularmente crítica porque pueden ocurrir en ventanas temporales acotadas y causar daños severos durante etapas fenológicas sensibles, como la floración, la brotación o la formación de frutos. Una caída de temperatura bajo determinados umbrales puede dañar células, tejidos, flores o frutos en pocas horas y afectar el rendimiento esperado de una temporada completa. Además, la temperatura crítica cambia según la especie, la variedad, el estado de desarrollo y las condiciones locales [3].

El impacto de una helada no depende únicamente de la temperatura mínima registrada, sino también de la duración del evento, la humedad, la nubosidad, el viento, la topografía local y la resolución espacial de los datos disponibles. En [4, 5] se muestra que la predicción requiere combinar variables meteorológicas, representar mejor la variación espacial de la temperatura mínima y evaluar distintos horizontes temporales. Por ello, una lectura aislada de temperatura no basta para anticipar el riesgo agrícola real.

En la práctica, el agricultor no necesita únicamente saber que una estación meteorológica registró una temperatura determinada, sino conocer si debe activar riego antiheladas, preparar sistemas de protección, reorganizar labores, revisar la disponibilidad de insumos o comunicar una alerta a su equipo. Aunque hoy existen estaciones climáticas, sensores, registros históricos y servicios de pronóstico, el problema no es solamente la disponibilidad de datos. El desafío es transformar esos datos en apoyo oportuno, confiable y comprensible para la operación agrícola. El acceso a redes agrometeorológicas, como la Red Agrometeorológica del Instituto de Investigaciones Agropecuarias (INIA) [6], constituye una base relevante, pero su valor aumenta cuando los datos pueden procesarse, contextualizarse y entregarse en formatos útiles para la acción.

En este trabajo, la agricultura inteligente se entiende de manera acotada y práctica: el uso de datos, conectividad y herramientas digitales para apoyar decisiones agrícolas concretas. No se plantea como incorporación tecnológica por sí misma, sino como una forma de resolver una necesidad operacional: capturar datos climáticos desde distintas fuentes, procesarlos, estimar riesgos y entregar información accionable al agricultor. La inteligencia del sistema no reside solo en el modelo predictivo, sino también en su capacidad de convertir información dispersa en una alerta comprensible y útil.

La brecha digital rural corresponde a las diferencias entre territorios, grupos sociales y actores productivos respecto del acceso, uso efectivo y aprovechamiento de tecnologías digitales. No se limita a la conectividad: también incluye disponibilidad de dispositivos, costo de acceso, habilidades digitales, nivel socioeconómico, confianza, pertinencia del contenido y utilidad percibida. En [7] se plantea que la conectividad de banda ancha en zonas rurales y remotas es una condición habilitante para la participación económica y social, pero que la infraestructura no asegura por sí sola la adopción de servicios digitales. Este aspecto resulta especialmente relevante para los sistemas agrícolas: una plataforma técnicamente correcta puede fracasar si exige interpretar datos complejos, utilizar portales poco familiares o modificar radicalmente las rutinas de trabajo.

En América Latina, la adopción de tecnologías digitales en el agro enfrenta restricciones adicionales. La agricultura tiene un peso económico y social relevante en la región, pero tecnologías como sensores, Internet de las cosas (IoT, del inglés *Internet of Things*), automatización o analítica avanzada siguen siendo inaccesibles para numerosos entornos rurales. En [8] se señala que estas tecnologías pueden apoyar el monitoreo, la trazabilidad y la eficiencia productiva, pero que su despliegue se encuentra limitado por barreras de inversión, conocimiento técnico y escala productiva, especialmente entre pequeños y medianos agricultores. En consecuencia, el diseño de soluciones agrícolas digitales debe considerar su accesibilidad, costo de adopción y forma de interacción con el usuario final, además de su capacidad técnica.

La digitalización agrícola se ha instalado como una vía relevante para mejorar la productividad, la trazabilidad y la capacidad de respuesta frente a riesgos productivos. Sin embargo, su adopción no depende solo de contar con infraestructura tecnológica. En Chile, las políticas públicas han transitado desde una agenda centrada en conectividad

hacia desafíos de uso, capacidades digitales y aplicaciones sectoriales, especialmente para la agricultura familiar campesina y la pequeña y mediana agricultura [9]. La brecha digital, por tanto, también comprende diferencias de frecuencia de uso, habilidades, apropiación, nivel socioeconómico y capacidad para transformar la tecnología en beneficios concretos [10].

El contexto chileno confirma esta tensión entre disponibilidad tecnológica y adopción efectiva. Existen iniciativas públicas, redes agrometeorológicas, plataformas de información territorial, programas de extensión y esfuerzos de digitalización sectorial. No obstante, en la Región del Maule se identifican redes fragmentadas, baja participación de productores en procesos de innovación, capacidades digitales insuficientes y brechas de conectividad que retrasan el acceso a tecnologías digitales [11]. Estas limitaciones muestran que el éxito de una solución no depende únicamente del modelo, sino también de su inserción en un ecosistema productivo real.

Los servicios de mensajería móvil forman parte de prácticas cotidianas de coordinación en el sector rural y agrícola chileno. En la encuesta de la Subsecretaría de Telecomunicaciones y CADEM, el 99,1 % de los hogares que respondieron la pregunta sobre dispositivos de conexión declaró utilizar teléfono móvil o *smartphone* para acceder a Internet en el hogar, tanto en áreas urbanas como rurales [12]. Por separado, [9] describe el uso de mensajería para coordinar labores agrícolas y apoyar sistemas mixtos de extensión rural. Estas condiciones favorecen un canal de comunicación cotidiano sin demostrar por sí solas su adopción por todos los productores. Las plataformas concretas del caso se presentan en el capítulo de implementación.

En este contexto, se propone un sistema de apoyo a decisiones agrícolas basado en la obtención automática de datos climáticos, su procesamiento y su uso posterior en modelos predictivos de heladas. La información generada debe integrarse con canales cercanos al usuario —un servicio de mensajería y una interfaz de gestión territorial— para facilitar una reacción temprana ante escenarios de riesgo. El foco del proyecto es el flujo completo de información: obtención, preparación, integración, verificación y comunicación. El entrenamiento de modelos se aborda en la medida necesaria para explicar sus salidas y su evaluación, no como eje aislado de la solución.

Este proyecto se sitúa, por tanto, en la intersección entre agricultura, datos climáticos e ingeniería de software. Su aporte no consiste solamente en aplicar un modelo de aprendizaje automático, sino en describir las decisiones necesarias para que ese modelo opere sobre datos reales, heterogéneos y cambiantes. La calidad del resultado depende tanto de la selección del algoritmo como de la disciplina con que se construye el flujo completo de datos. Esta mirada sistémica evita limitar la solución a condiciones controladas que no representan un entorno agrícola operativo.

1.1. Propósito

La relevancia del trabajo se justifica en tres dimensiones. Primero, desde la dimensión agrícola, porque las heladas son eventos de alto impacto que requieren anticipación y coordinación. Segundo, desde la dimensión tecnológica, porque la disponibilidad creciente de datos climáticos permite construir servicios predictivos siempre que exista una arquitectura capaz de integrarlos y depurarlos. Tercero, desde la dimensión social y territorial, porque una solución digital solo tiene valor si puede ser utilizada por quienes toman decisiones en terreno. La integración de estas dimensiones permite avanzar desde una demostración algorítmica hacia una herramienta aplicable a necesidades reales del sector.

1.2. Objetivos

Los objetivos delimitan el alcance técnico y aplicado del trabajo. Primero se presenta el objetivo general; luego se detallan objetivos específicos orientados a las fuentes de datos, la automatización, el procesamiento, la evaluación de modelos y la integración con el sistema.

1.2.1. Objetivo general

Desarrollar y documentar la implementación de un sistema de obtención, procesamiento y uso de datos climáticos para apoyar la predicción de heladas en un contexto de agricultura inteligente, mediante la integración de fuentes externas, procesos automáticos de extracción, modelos predictivos y canales de interacción accesibles para los agricultores.

1.2.2. Objetivos específicos

- Identificar y seleccionar fuentes de datos climáticos útiles para la predicción de heladas, considerando estaciones agrometeorológicas, servicios meteorológicos y registros históricos.
- Diseñar un flujo automatizado de obtención de datos mediante programas de extracción, consultas a servicios externos y tareas periódicas.
- Estandarizar y limpiar los datos recolectados para construir un conjunto de datos consistente y reutilizable.
- Evaluar modelos de clasificación y regresión capaces de estimar la ocurrencia de heladas y la evolución horaria de la temperatura.

- Integrar los resultados del modelo con una arquitectura modular que permita su uso en una interfaz de gestión territorial y en un servicio de mensajería.

1.3. Contribuciones

Las contribuciones se organizan en torno a resultados concretos de ingeniería: un flujo de datos climáticos, una base procesable para modelamiento, la evaluación de modelos predictivos y una arquitectura de integración orientada al uso agrícola. El proyecto documenta las decisiones técnicas, los desafíos encontrados y las soluciones implementadas, con énfasis en la trazabilidad y la calidad del flujo de datos.

La elección de canales cercanos al usuario no es secundaria. En contextos donde existen diferencias de alfabetización digital y disponibilidad de tiempo, una alerta enviada por un canal cotidiano puede ser más útil que una plataforma que exige navegación manual. La solución busca reducir la fricción entre el dato técnico y la acción agrícola: el agricultor no debería revisar múltiples fuentes, descargar archivos climáticos ni interpretar métricas de clasificación, sino recibir una señal clara, contextualizada y suficientemente confiable para apoyar una decisión de manejo.

La predicción de heladas combina variables meteorológicas sensibles, como temperatura mínima, humedad relativa, viento, radiación solar y precipitación, y requiere trabajar tanto con datos históricos como con datos cercanos al tiempo real. En [4] se muestra que el aprendizaje automático y los ensambles de modelos pueden mejorar la detección de estos eventos cuando se consideran múltiples variables y horizontes de predicción.

La construcción de esta base técnica presenta desafíos adicionales. Las fuentes climáticas no siempre publican información con la misma estructura, frecuencia o nivel de completitud. Los datos pueden contener valores faltantes, duplicados, unidades inconsistentes, desfases horarios o cambios en la nomenclatura de variables. Estas dificultades exigen procesos de recolección, limpieza y validación repetibles y mantenibles. En sistemas predictivos aplicados, el flujo de datos constituye la infraestructura que condiciona la confiabilidad de la predicción final.

1.4. Estructura del informe

El resto del informe se organiza de manera progresiva: primero se define el problema; luego se revisan antecedentes técnicos y bibliográficos; posteriormente se presenta la propuesta de solución y su implementación; finalmente, se sintetizan las conclusiones y el trabajo futuro.

- El capítulo 2 contextualiza y define el problema, con énfasis en las fuentes de datos, los

desafíos técnicos, los eventos meteorológicos relevantes y las necesidades del usuario.

- El capítulo 3 detalla el estado del arte sobre predicción de heladas, digitalización agrícola, eventos meteorológicos y brecha digital rural.
- El capítulo 4 describe la propuesta de solución, incluidos la recolección de datos, el procesamiento, los modelos predictivos y la integración con las interfaces de consulta, configuración y alertas.
- El capítulo 5 presenta el caso de estudio de la implementación, con las plataformas y prototipos concretos.
- El capítulo 6 expone las conclusiones, las limitaciones y las líneas futuras de trabajo.
- El anexo describe los modelos de voto y la regresión horaria del caso.

Capítulo 2

Contextualización y definición del problema

En este capítulo se aborda la dificultad de transformar datos climáticos dispersos en información accionable para anticipar heladas en contextos agrícolas. La problemática no se limita a la existencia de eventos meteorológicos adversos; también involucra la forma en que los datos se obtienen, integran, procesan e interpretan, y cómo llegan al agricultor en un momento útil para decidir.

2.1. Eventos meteorológicos y riesgo agrícola a nivel global

La producción agrícola depende directamente de condiciones ambientales que no pueden ser controladas por el productor. La temperatura, la humedad, la radiación, el viento, las precipitaciones, la disponibilidad de agua y las propiedades del suelo influyen en el desarrollo de los cultivos y la planificación de labores. Eventos como sequías, inundaciones, tormentas de viento y lluvia, olas de calor, granizos y heladas pueden afectar la disponibilidad hídrica, dañar infraestructura, reducir rendimientos, interrumpir faenas, aumentar costos de operación y generar pérdidas económicas durante una temporada completa. Los desastres y las crisis climáticas tienen impactos significativos sobre la agricultura, la seguridad alimentaria y los medios de vida rurales, especialmente cuando los productores no cuentan con herramientas oportunas de anticipación y respuesta [1, 2].

Tabla 2.1: Relación indicativa entre eventos meteorológicos, exposición agrícola y necesidad de respuesta [1–4].

Evento	**Cultivos expuestos**	**Riesgo principal**	**Variable espacial relevante**	**Tiempo de acción**	**Pérdida potencial**
Sequía	Cereales, frutales y hortalizas	Déficit hídrico y estrés prolongado	Humedad del suelo y déficit acumulado	Semanas a meses	Reducción progresiva del rendimiento y la calidad
Inundación	Hortalizas, cereales y frutales	Anegamiento, erosión y daño radicular	Precipitación acumulada y zonas inundables	Horas a días	Daño de cultivos, suelo e infraestructura
Ola de calor	Frutales, hortalizas y cereales	Estrés térmico, deshidratación y golpe de sol	Temperatura máxima y duración	Horas a días	Menor rendimiento y pérdida de calidad
Granizo	Frutales, vides y hortalizas	Daño físico directo	Trayectoria e intensidad de la tormenta	Minutos a horas	Pérdida inmediata y localizada de producción
Helada	Frutales, vides y hortalizas en etapas sensibles	Daño de flores, brotes, frutos y tejidos	Temperatura mínima y probabilidad por sector	24 a 72 horas	Pérdida de rendimiento o de producción estacional

Los tiempos y las pérdidas de la Tabla 2.1 son referenciales: cambian según la intensidad y duración del evento, la ubicación del predio, el cultivo y su etapa fenológica. Las variables espaciales permiten representar la intensidad o probabilidad del riesgo y distinguir sectores prioritarios dentro de una zona agrícola. Una alerta útil debe combinar información meteorológica, espacial y productiva, en lugar de comunicar solamente un valor climático aislado.

El riesgo agrícola presenta una característica que vuelve compleja su gestión: un pronóstico no describe un evento seguro, sino una probabilidad de ocurrencia. Para decidir, el agricultor necesita conocer la intensidad estimada, el sector potencialmente afectado, el cultivo expuesto, su etapa fenológica y las medidas preventivas posibles dentro del tiempo disponible. Una misma condición meteorológica puede tener consecuencias distintas según la especie, la variedad, la humedad, el viento, la topografía y el estado de desarrollo del cultivo. Por lo tanto, el riesgo surge de la combinación entre la probabilidad de la amenaza, la exposición, la vulnerabilidad productiva y la capacidad de reacción.

Dentro de los eventos meteorológicos adversos, las heladas representan un caso particularmente exigente. Una helada puede ocurrir durante pocas horas de la noche o la madrugada y provocar daños severos si coincide con etapas sensibles del cultivo. El daño puede manifestarse en flores, brotes, frutos o tejidos vegetales, y depende de temperaturas críticas que varían según la especie, la variedad, el estado fenológico, la duración del evento y las condiciones microclimáticas [3]. Esta sensibilidad temporal obliga a que la información llegue antes del evento, pues una alerta tardía pierde gran parte de su valor operativo.



Figura 2.1: Daño por bajas temperaturas en frutos de cerezo. Fuente: [13].

La Figura 2.1 ilustra el daño en frutos de cerezo. El riesgo no depende solo de la temperatura mínima; también inciden variables como la humedad relativa, el viento, la nubosidad, la radiación, la topografía y la resolución espacial de los datos meteorológicos [5]. En zonas con relieve complejo, pequeñas diferencias de altura, exposición o circulación de aire pueden generar variaciones locales relevantes. Por esto, una predicción útil requiere observar varias variables en conjunto, trabajar con datos suficientemente representativos y convertir el resultado en una señal comprensible para quien debe actuar en terreno.

Este panorama global permite acotar el problema abordado por el proyecto. Aunque los eventos meteorológicos presentan escalas y efectos diferentes, todos exigen transformar datos en información situada: probabilidad, zona afectada, cultivo expuesto y margen de reacción. A partir de este marco general, se profundiza en el fenómeno de las heladas, cuya breve ventana de ocurrencia vuelve especialmente crítica la anticipación.

2.2. Caracterización operacional de las heladas

Para comprender el problema que enfrenta el agricultor, es necesario distinguir entre la helada como fenómeno meteorológico y como evento operacional. Desde el punto de vista meteorológico, la helada describe un descenso de temperatura bajo un umbral de referencia. Desde la perspectiva agronómica, en cambio, el umbral relevante es la temperatura crítica del cultivo, que varía según la especie, la variedad y el estado fenológico [3]. El problema operacional aparece cuando la probabilidad de alcanzar ese umbral obliga a decidir si se activan medidas de protección, se reorganizan labores o se moviliza personal. Una medición puede ser suficiente para describir el fenómeno, pero insuficiente para apoyar la decisión.



Figura 2.2: Características de una helada radiativa. Fuente: INIA, Boletín 417 (2020) [13].

La Figura 2.2 representa condiciones de una helada radiativa. Las heladas pueden presentarse bajo condiciones locales muy específicas. La acumulación de aire frío en sectores bajos, la ausencia de viento, la humedad del ambiente, la nubosidad y la radiación durante la noche influyen en el comportamiento de la temperatura cerca del cultivo. Por ello, una medición tomada en una estación cercana puede no representar completamente la situación del predio. Esta diferencia entre medición puntual y condición local obliga a considerar la escala espacial de los datos y la forma en que estos se vinculan con el territorio agrícola.



Figura 2.3: Estados fenológicos del cerezo: (a) yema hinchada, (b) puntas blancas, (c) plena flor y (d) fruto cuajado. Fuente: [13].

La Figura 2.3 distingue etapas desde la yema hinchada hasta el fruto cuajado. En [3] se establece que las temperaturas críticas varían entre cultivos, variedades y estados fenológicos. Una temperatura tolerable durante la *dormancia* (estado de reposo de la planta) puede resultar dañina durante la floración o la cuaja. Por ello, una alerta genérica de baja temperatura puede servir como advertencia inicial, pero no siempre representa el riesgo productivo real.

El tiempo disponible para reaccionar es otro elemento central. Medidas como activar riego antiheladas, preparar equipos, revisar combustibles, coordinar personal o proteger sectores sensibles requieren anticipación. Si la alerta se emite cuando la temperatura ya alcanzó el punto crítico, el margen de acción puede ser insuficiente. La predicción no debe limitarse a detectar condiciones actuales, sino estimar la probabilidad de que el evento ocurra dentro de un horizonte operativo útil.

La gestión de heladas también implica balancear costos y riesgos. Una falsa alarma puede llevar al uso innecesario de agua, energía, horas de trabajo u otros recursos. En cambio, una helada no detectada puede dañar directamente la producción y afectar los ingresos de la temporada. Esta asimetría vuelve relevante la selección de métricas de evaluación y umbrales de alerta. No siempre conviene optimizar solo la exactitud global del modelo; también importa reducir la omisión de eventos relevantes y comprender el

costo práctico de cada tipo de error.

La predicción de heladas debe entenderse, por tanto, como un problema compuesto. Requiere datos meteorológicos, procesamiento, contextualización, evaluación de incertidumbre y comunicación clara. La dificultad no consiste solo en saber si la temperatura bajará, sino en construir un sistema que transforme esa información en una señal temprana, localizada y comprensible para apoyar decisiones de manejo.

2.3. De los datos climáticos a la decisión agrícola

En un sistema agrícola real, los datos climáticos recorren una cadena de valor antes de transformarse en una decisión. Primero se obtienen desde una fuente confiable; luego se integran, limpian, normalizan y almacenan; después se interpretan mediante reglas, modelos estadísticos o modelos de aprendizaje automático; finalmente, se comunican al usuario mediante un mensaje claro, oportuno y contextualizado. Cada tramo puede fallar y reducir el valor del sistema.

La primera brecha se produce en la obtención e integración. Las fuentes climáticas pueden publicarse mediante interfaces de programación de aplicaciones (API, del inglés *Application Programming Interface*), archivos descargables, tablas web, mapas interactivos o páginas diseñadas para consulta humana. Esta diversidad exige mecanismos diferentes de extracción, integración y mantenimiento. Cuando una fuente cambia su estructura, formato o nombres de variables, o interrumpe temporalmente el acceso, el sistema debe detectarlo y responder sin comprometer todo el flujo de datos.

La segunda brecha se produce en la calidad y consistencia de los datos. Para que un modelo predictivo funcione, no basta con acumular registros. Es necesario revisar fechas, unidades, valores faltantes, duplicados, registros inconsistentes entre estaciones, desfases horarios y cambios de nomenclatura. También se requiere conservar trazabilidad entre el dato original y el dato procesado, porque una alerta agrícola debe poder explicarse a partir de una fuente y un procesamiento verificables.

La tercera brecha aparece en la interpretación. Un agricultor no toma decisiones a partir de una matriz de variables meteorológicas, sino de preguntas concretas: si existe riesgo de helada, cuándo ocurrirá, dónde afectará, qué tan confiable es la alerta y qué acciones debería considerar. El sistema no debe limitarse a exponer datos crudos; debe traducirlos en información accionable y reducir el esfuerzo necesario para construir una conclusión.

Finalmente, existe una brecha de comunicación. Una predicción útil puede perder impacto si se entrega por un canal poco utilizado, con lenguaje excesivamente técnico o sin relación con la rutina del agricultor. La alerta debe integrarse a herramientas conocidas en el sector agrícola y presentar un formato directo. Por ello, el problema incluye tanto el modelo predictivo como la inserción de sus resultados en la operación diaria.

Estas brechas no operan de forma aislada. La **integración** constituye una dimensión transversal: la continuidad entre obtención, normalización, vinculación territorial, interpretación y comunicación determina si un dato que existe en una fuente puede convertirse en una señal comprensible para decidir. Una ruptura en cualquier tramo —por ejemplo, datos normalizados que no se asocian al predio, o una predicción interpretable que no alcanza el canal habitual del usuario— devuelve el sistema al punto de partida, aunque las etapas previas hayan funcionado correctamente.

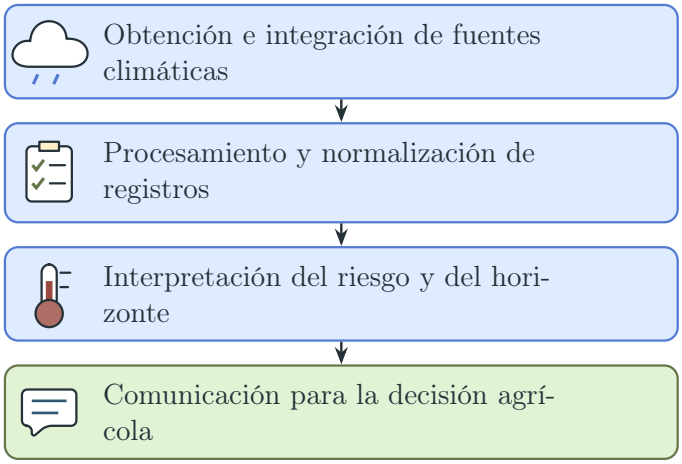


Figura 2.4: Flujo desde la obtención de fuentes climáticas hasta la decisión agrícola.

La Figura 2.4 presenta esas transformaciones como una secuencia, no como etapas intercambiables. Solo el último bloque llega al agricultor; los anteriores preparan el dato, pero no sustituyen la comunicación del riesgo.

Tabla 2.2: Niveles de transformación desde el dato climático hasta la alerta.

Nivel	Transformación	Riesgo si falla
Obtención e integración	Captura datos desde fuentes externas, los vincula y registra su origen.	Información incompleta, desactualizada o sin trazabilidad.
Procesamiento	Limpia, normaliza y almacena variables comparables.	Entradas inconsistentes y predicciones poco confiables.
Interpretación	Estima probabilidad, nivel de riesgo y horizonte temporal.	Resultados técnicos sin significado operacional.
Comunicación	Vincula el resultado con ubicación, cultivo y mensaje de alerta.	Información tardía, ambigua o ignorada por el usuario.

La conexión entre los niveles de la Tabla 2.2 constituye una práctica general en soluciones tecnológicas basadas en datos: el resultado final depende de la continuidad y

calidad del proceso completo. En desarrollos agrícolas debe aplicarse el mismo principio, pues optimizar una etapa aislada no resuelve el problema. Recopilar datos sin estandarizarlos, entrenar modelos sin asegurar la calidad de entrada o generar alertas sin considerar cómo las entiende el usuario produce soluciones frágiles, aunque alguno de sus componentes funcione correctamente.

La decisión agrícola se produce bajo incertidumbre. Incluso con datos adecuados, una predicción de helada no entrega certeza absoluta. Por ello, el sistema debe comunicar riesgo y no prometer determinismo. Esto implica diseñar umbrales, niveles de alerta y métricas que permitan interpretar el resultado como apoyo a la decisión, no como reemplazo del criterio del agricultor.

La trazabilidad también es parte del problema. Si una alerta indica riesgo alto de helada, el sistema debe relacionar esa salida con los datos de entrada, la fecha de actualización, la fuente utilizada y el procesamiento aplicado. Esta capacidad permite depurar errores, comparar fuentes, ajustar modelos y generar confianza mediante resultados comprensibles, verificables y consistentes en el tiempo.

2.4. Brecha digital rural y adopción tecnológica

Las tecnologías digitales permiten automatizar procesos, optimizar recursos y responder con mayor rapidez a cambios del entorno. En la agricultura, sin embargo, esta transformación enfrenta mayores dificultades, especialmente en territorios rurales y entre pequeños y medianos productores. En [7] se indica que la conectividad rural es una condición habilitante para el desarrollo económico y social, pero que la infraestructura no asegura la adopción efectiva: también influyen la asequibilidad, la alfabetización digital, la confianza, la pertinencia del contenido y la percepción de utilidad.

Esta distinción es central para definir el problema. Una solución agrícola puede ser técnicamente correcta y aun así fracasar si exige revisar múltiples portales, descargar archivos, interpretar gráficos meteorológicos o comprender métricas de modelos predictivos. La brecha digital rural no se limita a la posibilidad de conectarse a Internet; también incluye la capacidad de utilizar la información disponible para tomar mejores decisiones. En Chile, el acceso no garantiza un uso homogéneo ni un aprovechamiento equivalente entre distintos grupos sociales y territoriales [10].

En el contexto agrícola, esta brecha se vuelve más compleja porque el usuario no solo necesita acceso a tecnología, sino también pertinencia productiva. Una plataforma puede entregar datos meteorológicos correctos, pero si estos no se relacionan con el predio, el cultivo, el umbral de riesgo o la acción esperada, el agricultor seguirá enfrentando incertidumbre. La adopción depende de que la solución reduzca trabajo, simplifique decisiones y entregue beneficios visibles en la rutina productiva.

Tabla 2.3: Dimensiones acumulativas de la brecha digital rural.

Dimensión	Condición necesaria
Acceso	Disponer de conectividad, dispositivos y datos relevantes.
Uso	Integrar la herramienta en procesos agrícolas reales.
Comprensión	Interpretar alertas y resultados sin conocimientos técnicos especializados.
Apropiación	Reconocer la herramienta como parte útil y confiable de la rutina de trabajo.

Las dimensiones de la Tabla 2.3 son acumulativas: el acceso habilita el uso; el uso requiere comprensión; y la comprensión sostenida favorece la apropiación. Si una dimensión falla, la solución corre el riesgo de ser técnicamente interesante, pero operativamente poco utilizada.

En agricultura, la brecha se expresa de manera distinta que en entornos más estandarizados. Cada temporada cambia, cada predio presenta condiciones particulares, los cultivos no responden de la misma forma y el clima introduce incertidumbre. Una herramienta digital agrícola debe adaptarse al contexto productivo y no exigir que el usuario adapte completamente su operación a la herramienta.

La adopción también depende de la confianza. Si una plataforma contradice la experiencia acumulada del agricultor sin explicar el origen de una alerta, es probable que sea ignorada. En cambio, el uso de fuentes reconocibles, la entrega anticipada, la consistencia y la vinculación con una ubicación concreta pueden favorecer su adopción. Esto refuerza la necesidad de diseñar sistemas transparentes, trazables y orientados a decisiones reales.

El canal de comunicación influye directamente en la adopción. En el sector agrícola chileno, la coordinación mediante aplicaciones de mensajería móvil forma parte de prácticas reportadas en estudios de agricultura digital, junto con una alta disponibilidad de teléfono móvil [9].

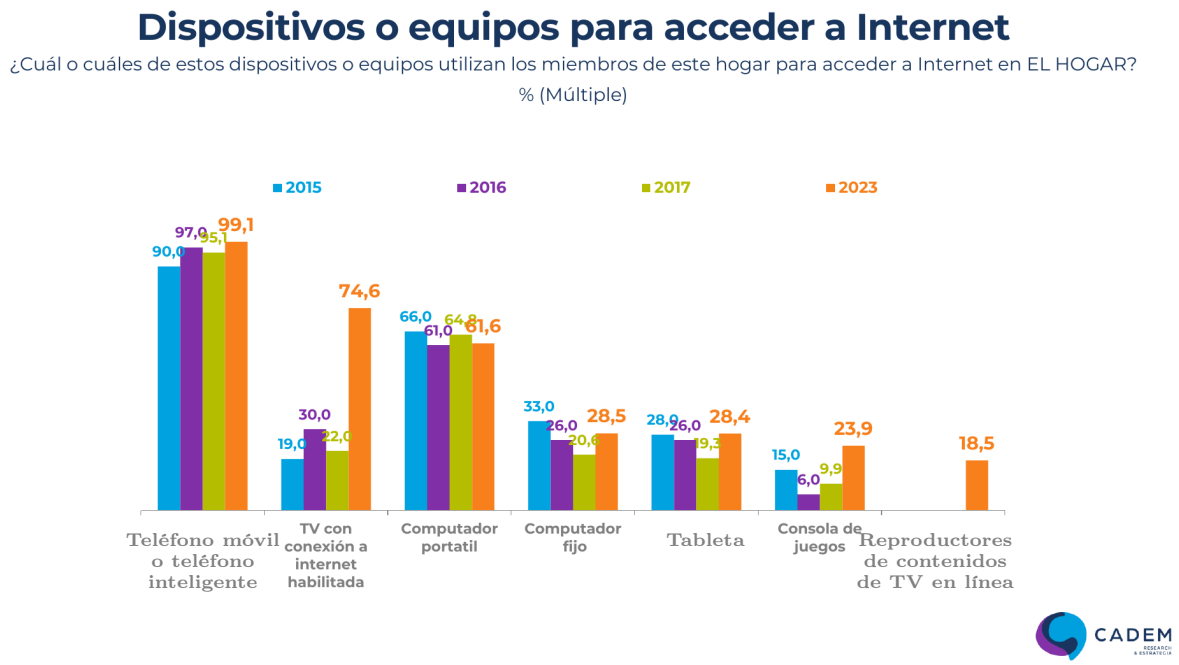


Figura 2.5: Dispositivos o equipos utilizados para acceder a Internet en hogares chilenos; la muestra representa hogares de Chile y no exclusivamente al sector agrícola. Fuente: [12]; adaptación al español de los anglicismos en las categorías, con cifras originales.

La Figura 2.5 distingue los dispositivos de acceso en hogares chilenos, no exclusivamente agrícolas. Aprovechar un canal cotidiano puede reducir la fricción de adopción respecto de portales o formatos ajenos al trabajo diario. El canal de comunicación permite consultar y recibir alertas; una interfaz de gestión territorial organiza campos, grupos y parámetros. El problema consiste en que la información llegue al agricultor por un medio cercano a su rutina y se relacione con el predio que administra.

Por lo tanto, la brecha digital rural no debe entenderse como una carencia pasiva del usuario, sino como una responsabilidad de diseño. Si la herramienta no considera las condiciones de conectividad, el tiempo disponible, el lenguaje, los hábitos de comunicación y las necesidades productivas, el problema persistirá aunque existan datos, modelos y aplicaciones.

2.5. Situación en América Latina

En América Latina, la adopción de tecnologías digitales en el agro enfrenta barreras estructurales. La presencia significativa de pequeños productores y agricultura familiar limita la incorporación de soluciones costosas o complejas que requieren inversión inicial, conocimiento técnico especializado y soporte permanente. En [8] se señala que

la sensorización, el IoT y la agricultura de precisión pueden aportar al monitoreo, la trazabilidad y la eficiencia productiva, pero que su adopción se restringe por barreras financieras, escala productiva y falta de conocimiento técnico. Esta condición exige soluciones capaces de aprovechar fuentes externas, datos públicos y canales cotidianos sin depender exclusivamente de infraestructura avanzada instalada en el predio.

Más allá de estas barreras generales, algunos países de la región han desarrollado iniciativas públicas de información agroclimática y territorial que permiten contrastar distintos grados de centralización y acceso. Los casos siguientes se presentan como ejemplos orientativos y no constituyen un inventario exhaustivo de América Latina [14].

En Argentina, INTA Digital Geo ofrece mapas temáticos, datos espaciales, un visor, geoportales y geoservicios de libre acceso, con información sobre agricultura y recursos naturales, incluidos suelo, agua, clima y meteorología [15]. En Brasil, la plataforma AgroAPI de Embrapa agrupa APIs orientadas a la agricultura digital. El catálogo público incluye APIs como Agritec, ClimAPI y SATVeg, entre otras [16]. En Colombia, el observatorio Agronet de la UPRA concentra información rural y agropecuaria, con estadísticas, plataformas y boletines agroclimáticos [17]. En Uruguay, el SNIA del MGAP proporciona información y visualizadores orientados a decisiones, políticas públicas y gestión de riesgo agropecuario [18].

Tabla 2.4: Iniciativas públicas de información agroclimática en países seleccionados de América Latina (contraste ilustrativo).

País	Iniciativa	Aporte verificable	Brecha relevante
Argentina	INTA Digital Geo [15]	Mapas temáticos, datos espaciales, visor, geoportales y geoservicios libres (suelo, agua, clima).	La oferta pública no garantiza por sí sola integración predial ni alerta accionable.
Brasil	AgroAPI (Embrapa) [16]	Plataforma de APIs con catálogo que incluye Agritec, ClimAPI y SATVeg, entre otros.	Disponibilidad de APIs no equivale a adopción ni a interoperabilidad operativa.
Colombia	Agronet (UPRA) [17]	Observatorio con estadísticas, plataformas y boletines agroclimáticos.	La publicación agregada no asegura uso en predio ni adopción por productores.
Uruguay	SNIA (MGAP) [18]	Información y visualizadores para decisiones y gestión de riesgo agropecuario.	La centralización institucional no sustituye integración territorial ni canales de uso.

La Tabla 2.4 sintetiza estos casos. En conjunto, muestran distintos grados de centralización y de acceso a información pública, pero la mera disponibilidad de portales,

observatorios o APIs no equivale automáticamente a integración predial, interoperabilidad, adopción efectiva ni entrega de una alerta accionable en el predio. Los ejemplos son ilustrativos y no agotan la diversidad regional; tampoco implican que dichas plataformas aborden por sí solas la anticipación de heladas u otros eventos de escala operativa.

En complemento a lo anterior, la predicción de heladas presenta desafíos meteorológicos regionales. En [4] se evalúan soluciones para el sur y sureste de Brasil, Uruguay, Paraguay, norte de Argentina y sureste de Bolivia mediante la combinación de variables meteorológicas y horizontes de 24 y 72 horas. Estos casos muestran potencial técnico, pero también confirman que el valor final depende de datos consistentes, actualizados y correctamente procesados.

Ambas dimensiones impiden resolver el problema mediante una solución genérica. El sistema debe adaptarse a la disponibilidad real de datos, los cultivos relevantes, las capacidades digitales de los usuarios, el tipo de conectividad y las instituciones que publican información climática. Algunas fuentes regionales ofrecen plataformas estructuradas; otras publican documentos, mapas o páginas destinadas a consulta humana. Asimismo, existen predios con sensores propios y conectividad estable, mientras otros dependen de datos públicos y canales móviles. Estos contrastes exigen una arquitectura flexible capaz de incorporar fuentes heterogéneas y operar con distintos niveles de calidad y frecuencia de actualización.

Frente a este escenario, la modularidad constituye una respuesta concreta. Si una fuente deja de estar disponible, el sistema debería poder incorporar otra; si un modelo no funciona bien en una zona, debería poder evaluarse o reemplazarse; si un canal no es adecuado para determinado grupo, debería poder complementarse con otro. La modularidad no es solo una preferencia de ingeniería, sino una respuesta al carácter variable del entorno agrícola y digital latinoamericano.

2.6. Contexto chileno: datos disponibles, fragmentación y brecha de uso

Chile dispone de redes agrometeorológicas, portales públicos, servicios meteorológicos, registros históricos e iniciativas institucionales orientadas a la digitalización agrícola. La Red Agrometeorológica del INIA entrega datos provenientes de estaciones distribuidas en el territorio nacional [6], mientras que la Dirección Meteorológica de Chile mantiene un portal institucional de acceso público [19]. Esta disponibilidad constituye una ventaja, pero los datos necesarios para anticipar heladas no se encuentran necesariamente en un único lugar ni en formatos directamente utilizables por un sistema predictivo.

En este contexto conviene distinguir tres niveles que no coinciden necesariamente. El **dato existente** corresponde a registros publicados por una institución en un repositorio o portal. El **dato técnicamente accesible** puede ser consultado o extraído por un

especialista que conoce el mecanismo de acceso, aunque ello exija cruzar sitios, formatos o criterios heterogéneos. El **dato utilizable para decidir** llega al agricultor con la escala territorial, el horizonte temporal y el formato que permiten actuar antes del evento. La fragmentación chilena afecta sobre todo los dos últimos niveles: la información puede existir sin ser fácilmente integrable ni traducible en una alerta comprensible en terreno.

Tabla 2.5: Fuentes climáticas públicas relevantes en Chile (según portales institucionales verificables).

Fuente	Información o servicio observado	Forma pública de consulta	Implicación para el problema
Red Agrometeorológica INIA [6]	Estaciones agrometeorológicas, variables meteorológicas, resumen y pronóstico, sección de heladas; consulta mediante tablas, gráficos y descarga en Excel o CSV según el portal.	Portal web institucional (*agrometeorologia.cl*).	La información existe y es consultable, pero su uso para anticipar heladas exige seleccionar estaciones, variables y formatos de forma explícita.
Dirección Meteorológica de Chile [19]	Portal institucional público de información meteorológica oficial.	Sitio web de la institución (*meteochile.gob.cl*).	Constituye una fuente de referencia climática reconocida; su aprovechamiento operativo dependerá de identificar qué productos resultan pertinentes para el contexto agrícola considerado.

La Tabla 2.5 sintetiza la información y las formas públicas de consulta de estas dos fuentes, junto con sus implicaciones para el problema.

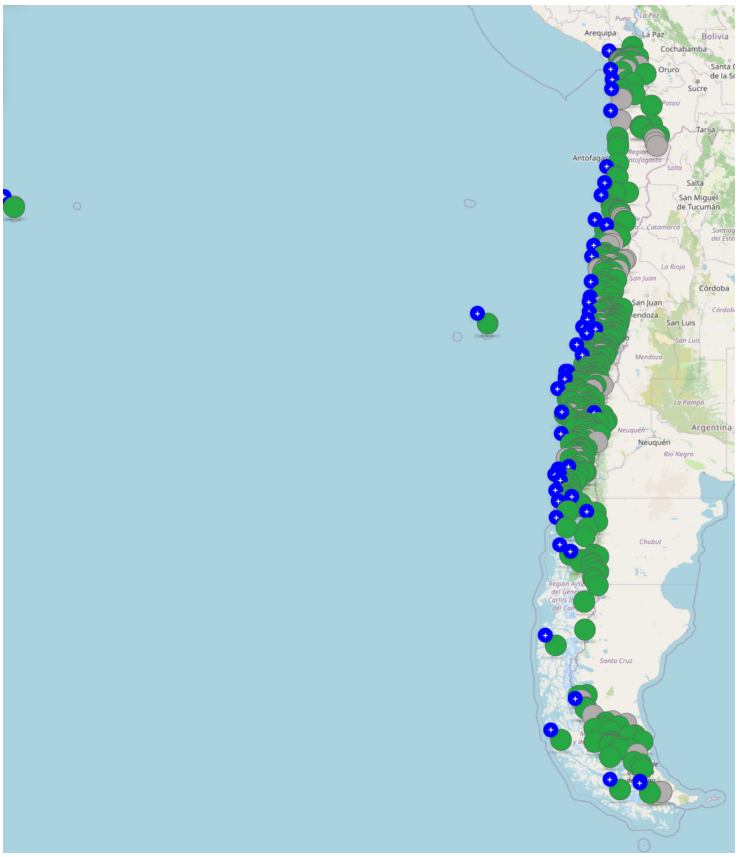


Figura 2.6: Distribución territorial de las estaciones de la Red de Agrometeorología del INIA. Fuente: [6].

La Figura 2.6 sitúa las estaciones en el territorio; su distribución no implica que cada predio disponga de una medición representativa.

La política pública chilena reconoce que el desafío de la agricultura digital no se reduce a la conectividad. En [9] se describe una transición desde políticas centradas en infraestructura hacia desafíos de uso, capacidades digitales, aplicaciones sectoriales y digitalización de procesos productivos, especialmente en la agricultura familiar campesina y la pequeña y mediana agricultura. Aunque Chile presenta altos niveles de acceso declarado a Internet y telefonía móvil, persisten diferencias territoriales y necesidades de apropiación y pertinencia de los servicios [10].

La brecha de uso también se expresa territorialmente. En la Región del Maule, cuya actividad agrícola incluye rubros frutícolas y agroalimentarios sensibles a las heladas, se identifican redes fragmentadas, capacidades digitales insuficientes, baja participación de productores y brechas de conectividad que retrasan la adopción tecnológica [11]. Estas condiciones son relevantes porque la utilidad de una alerta depende tanto del modelo como de la capacidad del sistema para insertarse en un ecosistema productivo real.

A estas barreras se suma el impacto productivo de los eventos climáticos. En [20] se informan pérdidas por US$1.932 millones para el sector agrícola chileno durante 2023, asociadas a incendios, inundaciones y lluvias tardías. Esta cifra agrega distintos tipos de eventos y, por ello, no permite atribuir ese monto a las heladas. Sin embargo, evidencia la magnitud económica del problema general y la necesidad de mejorar la anticipación y coordinación de medidas preventivas.

En síntesis, Chile combina fuentes de datos climáticos y un ecosistema institucional que habilita la digitalización agrícola con barreras que dificultan transformar esos recursos en decisiones oportunas en terreno.

2.7. Dimensiones técnicas del problema en Chile

El contexto chileno descrito se traduce en un conjunto de dimensiones técnicas que separan la disponibilidad de datos de la entrega de una alerta accionable. Cada dimensión condiciona a las demás, por lo que conviene revisarlas en conjunto antes de plantear una solución.

La primera dimensión corresponde a la **disponibilidad distribuida de datos**. Las fuentes entregan observaciones, pronósticos, antecedentes históricos o información territorial bajo distintos mecanismos de acceso. Esta dispersión exige criterios de selección relacionados con cobertura geográfica, frecuencia de actualización, variables disponibles, estabilidad, granularidad temporal, confiabilidad institucional y posibilidad de automatización.

La segunda dimensión es la **automatización de la extracción**. Un sistema de alerta no puede depender de descargas manuales o revisiones diarias de portales. Debe ejecutar procesos periódicos que obtengan datos, detecten errores y actualicen la información, considerando que las fuentes externas pueden cambiar o fallar.

La tercera dimensión es la **normalización e integración**. Distintas fuentes pueden representar fechas, horas, estaciones, comunas, coordenadas y variables meteorológicas de maneras diferentes. La integración requiere transformar esa variedad en una estructura común, vincular registros equivalentes y conservar el origen de cada dato.

La cuarta dimensión se relaciona con la **calidad y completitud**. Los registros pueden contener valores faltantes, mediciones atípicas, cortes temporales, estaciones sin actualización reciente o datos duplicados. Estos problemas pueden sesgar el entrenamiento, generar falsas alertas o impedir la detección de eventos reales, por lo que la limpieza debe considerarse una etapa crítica.

La quinta dimensión corresponde a la **vinculación territorial**. Para que la predicción sea útil, los datos deben asociarse con ubicaciones relevantes para el agricultor. La escala de publicación de una fuente no siempre coincide con la escala de decisión del usuario;

el sistema debe reconocer esta diferencia y explicitar qué datos representan cada zona.

La sexta dimensión es la **integración con los canales de uso**. Los resultados no deben quedar aislados en archivos o reportes técnicos, sino conectarse con los servicios que entregan información al agricultor. Esto exige contratos de datos, identificadores, campos, umbrales y reglas de alerta. El uso de mensajería móvil en tareas agrícolas documentado en [9] favorece un canal de comunicación de baja fricción para consultas y alertas. La interfaz de gestión territorial, por su parte, organiza el mapa, los campos, los grupos y la configuración. La interfaz sitúa la información en el predio; el canal lleva el mensaje. Ambas funciones se complementan y no se sustituyen.

Estas dimensiones no operan de forma independiente. La **integración transversal** expresa la continuidad requerida entre disponibilidad, extracción automatizada, normalización, calidad, vinculación territorial y canales de uso: una ruptura en la articulación entre ellas impide que el dato deje de ser solo existente o accesible y se convierta en información utilizable para decidir, aun cuando etapas aisladas funcionen correctamente.

Tabla 2.6: Dimensiones técnicas del problema climático en Chile.

Dimensión	Síntesis del problema
Disponibilidad distribuida	Datos relevantes publicados en fuentes múltiples, con criterios heterogéneos de cobertura, actualización y acceso.
Automatización de extracción	Dependencia de consultas manuales frente a fuentes que pueden cambiar, interrumpirse o requerir actualización periódica.
Normalización e integración	Convenciones distintas entre fuentes dificultan comparar, vincular y conservar el origen de cada registro.
Calidad y completitud	Faltantes, atípicos, cortes y duplicados pueden sesgar el análisis o inducir alertas incorrectas.
Vinculación territorial	La escala de publicación de la fuente no siempre coincide con la escala de decisión del predio o sector.
Integración con canales de uso	Resultados técnicos que no alcanzan canales cotidianos del agricultor pierden utilidad operacional.
Integración transversal	Fallas de continuidad entre dimensiones devuelven el flujo al punto de partida, aunque etapas aisladas funcionen.

La Tabla 2.6 resume sintéticamente lo expuesto anteriormente. Estas dimensiones muestran que el núcleo del problema se encuentra en la coordinación de múltiples capas: el flujo debe comportarse como una cadena y, si uno de sus eslabones falla, el resultado final se debilita. Resolver el problema exige poder observar, mantener y ajustar cada etapa sin perder la continuidad del flujo completo.

2.8. Causas y consecuencias del problema

Las causas descritas pueden organizarse según su origen y efecto principal. Esta síntesis evita tratar el problema como una sola falla técnica y muestra por qué la solución debe conectar datos, procesamiento, predicción, comunicación y adopción.

Tabla 2.7: Causas principales del problema y sus efectos sobre la solución.

Causa	Efecto inmediato	Consecuencia para la solución
Fuentes climáticas fragmentadas	Datos distribuidos y formatos heterogéneos.	Mayor complejidad de extracción, integración y mantenimiento.
Calidad y cobertura variables	Registros faltantes, desactualizados o poco representativos.	Predicciones menos confiables y menor trazabilidad.
Riesgo dependiente del contexto	Un mismo evento afecta de forma distinta según zona, cultivo y etapa.	Necesidad de vinculación territorial y parámetros productivos.
Comunicación poco accesible	El usuario recibe información técnica, tardía o fuera de su rutina.	Baja adopción y menor capacidad de reacción.

La Tabla 2.7 evidencia dependencias cruzadas. Una fuente incompleta no solo afecta el procesamiento, sino también la confianza del usuario si genera una alerta incorrecta. Del mismo modo, una comunicación deficiente puede volver inútil una predicción técnicamente correcta. La propuesta debe evaluarse como un sistema completo y no únicamente por el desempeño aislado de un componente.

2.8.1. Consecuencias de no abordar el problema

Cuando los datos climáticos no se transforman en información accionable, el agricultor queda expuesto a decisiones tardías o basadas en antecedentes incompletos. Esto puede generar activación innecesaria de medidas preventivas, falta de reacción ante eventos reales, uso ineficiente de recursos o pérdidas productivas evitables. En el caso de una helada, el costo de una alerta omitida puede ser particularmente alto, porque el daño puede ocurrir en pocas horas y afectar etapas sensibles del cultivo [3].

Desde la perspectiva tecnológica, una integración deficiente produce sistemas difíciles de mantener. Si cada fuente se procesa manualmente o mediante programas aislados, incorporar nuevas estaciones, corregir errores o auditar resultados se vuelve complejo.

Sin trazabilidad, también resulta difícil explicar por qué se generó una alerta o por qué el modelo produjo determinado resultado.

Desde la perspectiva de adopción, una solución que no considera la rutina del usuario puede quedar subutilizada. El agricultor podría continuar consultando fuentes dispersas o recibiendo información mediante canales informales, aun cuando existan datos más completos. Esto refuerza la brecha entre disponibilidad tecnológica y beneficio productivo real.

Las consecuencias pueden agruparse en tres niveles. En el nivel productivo, una helada no anticipada puede afectar el rendimiento, la calidad de la fruta, la planificación de la cosecha y la rentabilidad. En el nivel operativo, la falta de información integrada obliga a reaccionar con poca anticipación o coordinar recursos de forma improvisada. En el nivel tecnológico, la ausencia de datos limpios y trazables limita cualquier intento de modelamiento.

2.9. Delimitación del problema

A partir del contexto anterior, el problema central puede formularse mediante la siguiente pregunta:

> ***¿Cómo transformar fuentes climáticas dispersas en Chile en un sistema capaz de apoyar la predicción de heladas y entregar alertas comprensibles para agricultores?***

Existen datos climáticos relevantes para anticipar heladas, pero se encuentran distribuidos, presentan formatos heterogéneos y no están preparados directamente para alimentar un sistema de alertas agrícolas. Además, su interpretación exige conocimientos técnicos que no siempre están presentes en los usuarios finales. El problema no es la ausencia absoluta de información, sino la falta de un flujo integrado que la convierta en una alerta útil, trazable y accionable.

Esta brecha produce una desconexión entre las fuentes de datos climáticos, los modelos predictivos y la decisión agrícola concreta. Si los datos no se recolectan de forma confiable, el modelo no puede operar adecuadamente. Si el modelo entrega resultados difíciles de interpretar, el agricultor no puede convertirlos en acción. Si la alerta llega mediante un canal ajeno a su rutina, disminuye la probabilidad de uso efectivo. Por lo tanto, el problema no consiste solo en predecir heladas, sino en construir un flujo completo que transforme datos dispersos en información oportuna, contextualizada y comprensible.

El problema también presenta una dimensión temporal. Una alerta de helada es útil solo si llega con suficiente anticipación para ejecutar medidas preventivas. Un sistema

que detecta correctamente un evento después de ocurrido no resuelve la necesidad operacional. Los horizontes de 24 y 72 horas evaluados en [4] ofrecen distintos márgenes de reacción, aunque también introducen desafíos de precisión y confiabilidad.

Desde el punto de vista técnico, se requiere una arquitectura que soporte fuentes heterogéneas, procesamiento repetible, almacenamiento consistente, ejecución de modelos, evaluación de resultados e integración con canales de comunicación. Cada etapa debe ser mantenible, porque un sistema de alertas agrícolas no puede depender de procedimientos manuales frágiles. La automatización no busca reemplazar el criterio del agricultor, sino reducir el tiempo entre la aparición de una condición de riesgo y la entrega de una señal interpretable.

La solución debe contemplar además la escalabilidad. El número de estaciones, fuentes, variables, usuarios o predios no es fijo: un sistema que funciona con pocas fuentes en una zona piloto puede fallar al incorporar más estaciones, más regiones o más destinatarios. Por ello, los procesos deben ser repetibles e idempotentes, el almacenamiento debe crecer con el volumen de registros, la actualización de modelos debe ser incremental y los canales deben poder distribuir alertas a una población mayor sin degradar la oportunidad de entrega. La escalabilidad forma parte del problema porque condiciona la arquitectura desde el inicio y no puede postergarse como un detalle de despliegue.

Desde el punto de vista del usuario, la alerta debe indicar el nivel de riesgo, el momento estimado, la ubicación relevante y una recomendación general de atención. Si el mensaje es demasiado técnico, el usuario debe invertir tiempo en interpretarlo; si es demasiado genérico, no ayuda a decidir; si llega tarde, pierde utilidad. La calidad de la comunicación es, por tanto, una parte constitutiva del problema y no un detalle posterior de interfaz.

Capítulo 3

Estado del arte

Este capítulo revisa los antecedentes para un sistema de alerta temprana de heladas agrícolas siguiendo el flujo que transforma datos en decisiones: obtención e integración, procesamiento y generación de predicciones, interpretación de resultados y comunicación de alertas. Se examinan las fuentes meteorológicas y sus escalas, los métodos de predicción, las métricas para interpretar su desempeño y los mecanismos operativos que traducen pronósticos en alertas accionables.

3.1. Obtención e integración de datos meteorológicos

La adquisición reúne fuentes con escalas y propósitos distintos. Entre las fuentes globales en rejilla se encuentra el reanálisis de quinta generación del Centro Europeo de Predicción Meteorológica a Plazo Medio (ECMWF, del inglés *European Centre for Medium-Range Weather Forecasts*), denominado ERA5 (del inglés *ECMWF Reanalysis v5*), utilizado para entrenar GraphCast sobre una rejilla de 0,25°, niveles atmosféricos y pasos de seis horas [21]. Para aplicaciones estacionales, el Sistema de Pronóstico Estacional 5 (SEAS5, del inglés *Seasonal Forecast System 5*) aporta repronósticos que pueden reducirse de escala, como en el estudio de valles alpinos que pasó de 12 km a 250 m [5]. Estas fuentes favorecen cobertura espacial y continuidad temporal, pero su rejilla no sustituye observaciones del entorno inmediato.

Las observaciones de alta frecuencia y las estaciones locales complementan esos productos. Las composiciones de radar empleadas para precipitación pueden disponer de resolución de 1 km y frecuencia de cinco minutos [22]; para heladas, las series de estaciones aportan temperatura, radiación, viento, precipitación y punto de rocío cercanos al cultivo [23]. La elección de fuente depende del fenómeno y del horizonte: cobertura global para condiciones atmosféricas de gran extensión, observación frecuente para evolución de corto plazo y estaciones locales para condiciones de superficie.

La integración requiere verificar de forma explícita la resolución espacial, frecuencia temporal, unidades, sistema de coordenadas, período disponible, valores faltantes y procedencia de cada registro. También exige documentar las transformaciones aplicadas —por ejemplo, filtros de rango e interpolación de faltantes en series de estaciones [23]— para no tratar fuentes heterogéneas como equivalentes. La automatización depende, además, de la forma de acceso: una consulta web, una descarga manual y un servicio estructurado no ofrecen las mismas garantías de actualización ni trazabilidad.

En Chile, el modelo regional de Investigación y Pronóstico del Tiempo de la Dirección Meteorológica de Chile (WRF-DMC; WRF, del inglés *Weather Research and Forecasting*; DMC, Dirección Meteorológica de Chile) utiliza condiciones iniciales y de borde del Sistema Global de Pronóstico (GFS, del inglés *Global Forecast System*) y, para la zona centro, también del ECMWF [24]. La Red de Agrometeorología del Instituto de Investigaciones Agropecuarias (INIA) reúne estaciones de distintas instituciones, permite consultas históricas y publica pronósticos de heladas a 24 y 48 horas basados en información de la DMC [25]. Como ejemplo de acceso estructurado, el servicio de alertas del Servicio Meteorológico Nacional de Estados Unidos (NWS, del inglés *National Weather Service*) publica productos mediante una interfaz de programación de aplicaciones (API, del inglés *Application Programming Interface*) [26].

En comparación, las fuentes globales en rejilla privilegian cobertura y continuidad; el NWS ejemplifica acceso estructurado a productos; los estudios regionales y locales latinoamericanos dependen de variables y estaciones situadas para casos específicos; y WRF-DMC e INIA aportan, respectivamente, productos regionales y observaciones para Chile [4, 5, 21, 23–26]. EW4All aporta un marco de alerta temprana de extremo a extremo, no una fuente de datos [27].

Así, la adquisición e integración condicionan el flujo posterior: un método puede mostrar buen desempeño experimental, pero no ser operable si sus variables no están disponibles, la cobertura de estaciones es insuficiente o el acceso requiere intervención manual. La comparación entre trabajos debe considerar tanto el algoritmo como los datos, su preparación, su procedencia y su viabilidad operativa.

3.2. Procesamiento y generación de predicciones meteorológicas

Los métodos de predicción responden a escalas y horizontes diferentes. Distinguir el pronóstico global de mediano plazo, la predicción inmediata y los fenómenos localizados de superficie evita atribuir a una misma técnica una resolución u horizonte que no le corresponde.

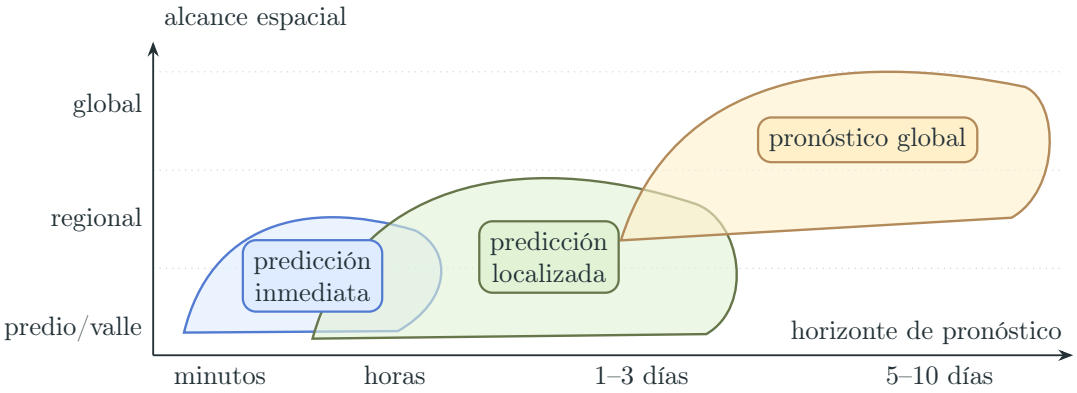


Figura 3.1: Escalas conceptuales de horizonte y alcance espacial para predicción inmediata, predicción localizada y pronóstico global.

La Figura 3.1 sitúa las decisiones en un continuo desde el corto plazo y el entorno local hasta horizontes más extensos y alcance global. Los métodos localizados actúan como puente entre ambos extremos; por ello, ningún método único cubre con igual pertinencia todas las escalas de decisión.

3.2.1. Pronóstico global, predicción inmediata y fenómenos localizados

GraphCast es un modelo de aprendizaje profundo para pronóstico meteorológico global de mediano plazo. Fue evaluado con variables atmosféricas de superficie y niveles superiores, para horizontes de varios días, y obtuvo desempeño comparable o superior al de sistemas físicos operativos de referencia en múltiples variables y plazos dentro de ese rango [21].

La predicción inmediata (*nowcasting*) estima la evolución de un fenómeno durante los minutos u horas siguientes a partir de observaciones recientes de alta frecuencia. En [22] se desarrollan modelos generativos profundos para generar campos de precipitación desde imágenes de radar y se evalúan frente a métodos operativos en ese horizonte corto.

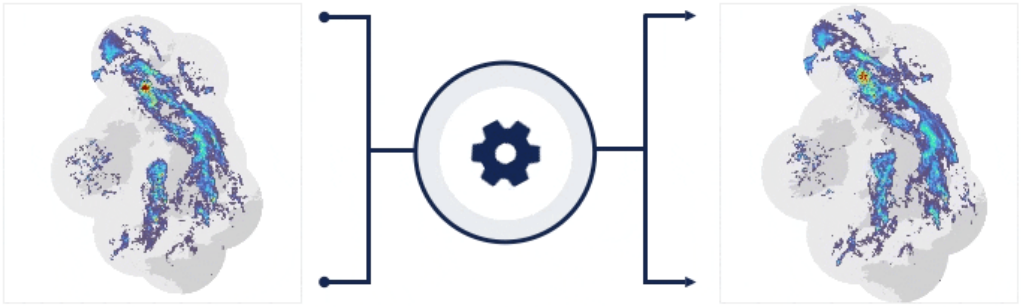


Figura 3.2: Ejemplo de generación de *nowcasting* de precipitación a partir de contexto de radar mediante un modelo generativo profundo. Fuente: [22]; etiquetas adaptadas al español.

La Figura 3.2 separa el contexto observado de la secuencia futura generada, lo que ilustra procesamiento para horizontes cortos. El caso de precipitación muestra esa capacidad de anticipación inmediata, pero no valida por sí mismo la predicción de heladas, cuyas señales y criterios de evaluación son distintos.

Para fenómenos localizados de superficie, [28] construye modelos estadísticos y de aprendizaje automático para anticipar niebla radiativa en Japón a partir de variables meteorológicas de superficie. El caso ilustra que los fenómenos de predio o valle requieren señales y resolución distintas a las del pronóstico sinóptico.

3.2.2. Predicción de heladas en la literatura internacional

La predicción de heladas puede estimar una variable continua, como la temperatura mínima, o clasificar la ocurrencia de un evento bajo un umbral. Los trabajos revisados incluyen índices derivados de modelos regionales, comités de aprendizaje automático y métodos de reducción de escala. Para comparar sus aportes se consideran los datos utilizados, la escala, el horizonte y las limitaciones de cada estudio.

Tabla 3.1: Estudios internacionales sobre predicción de heladas.

Referencia	**Enfoque**	**Datos y escala**	**Horizonte o salida**	**Limitación**
[29]	Índice de helada sobre variables de modelo regional	Salidas de modelo atmosférico regional	Pronóstico de evento de helada en áreas evaluadas	Desempeño variable según región; índice requiere variables de modelo
[4]	Aprendizaje automático; comité por votación	Variables meteorológicas; Sudamérica	24 y 72 h	Transferencia a Chile no demostrada; necesita datos consistentes
[30]	DNN, CNN y RF a escala de punto	Series locales; relieve complejo	6–48 h	Error local; transferencia entre sitios limitada
[5]	Reducción de escala espacial (ANN, RF, CNN)	Rejilla fina; valles alpinos	$T_{\text{mín}}$ diaria y heladas	Contexto europeo; no operación institucional chilena

La Tabla 3.1 evidencia diversidad de salidas, horizontes y escalas entre los estudios revisados. Antes de emplear estos antecedentes en una operación chilena, resulta necesaria una validación local que compruebe su comportamiento bajo condiciones, datos y decisiones del territorio.

El índice de helada de [29] mejora la discriminación respecto del uso directo de la temperatura de un modelo regional en las áreas evaluadas. El comité por votación de [4] evalúa horizontes de 24 y 72 horas en Sudamérica, mientras que [30] y [5] comparan redes neuronales profundas (DNN, del inglés *Deep Neural Network*), redes neuronales convolucionales (CNN, del inglés *Convolutional Neural Network*), redes neuronales artificiales (ANN, del inglés *Artificial Neural Network*) y bosques aleatorios (RF, del inglés *Random Forest*) para escalas locales o reducción de escala. Los productos globales revisados no establecen por sí mismos habilidad para heladas a escala de predio en Chile; su uso puede requerir reducción de escala o calibración local.

3.3. Interpretación y métricas de evaluación

La salida de un modelo necesita interpretarse antes de convertirse en una alerta. Conviene distinguir tres niveles. La **observación y el monitoreo** registran o visualizan condiciones ya ocurridas o en curso —por ejemplo, temperaturas medidas en estaciones

o mapas de eventos pasados— sin estimar por sí solos la probabilidad futura del fenómeno. El **pronóstico meteorológico** proyecta variables atmosféricas —temperatura, humedad, viento y nubosidad— sobre un horizonte definido. La **predicción de riesgo de helada** contrasta esas variables, índices o probabilidades con umbrales relevantes, que pueden diferir del daño agrícola efectivo según cultivo y fenología [3]. Una temperatura mínima bajo 0 °C describe una helada meteorológica; el impacto productivo depende además de la especie, la variedad y el estado del cultivo.

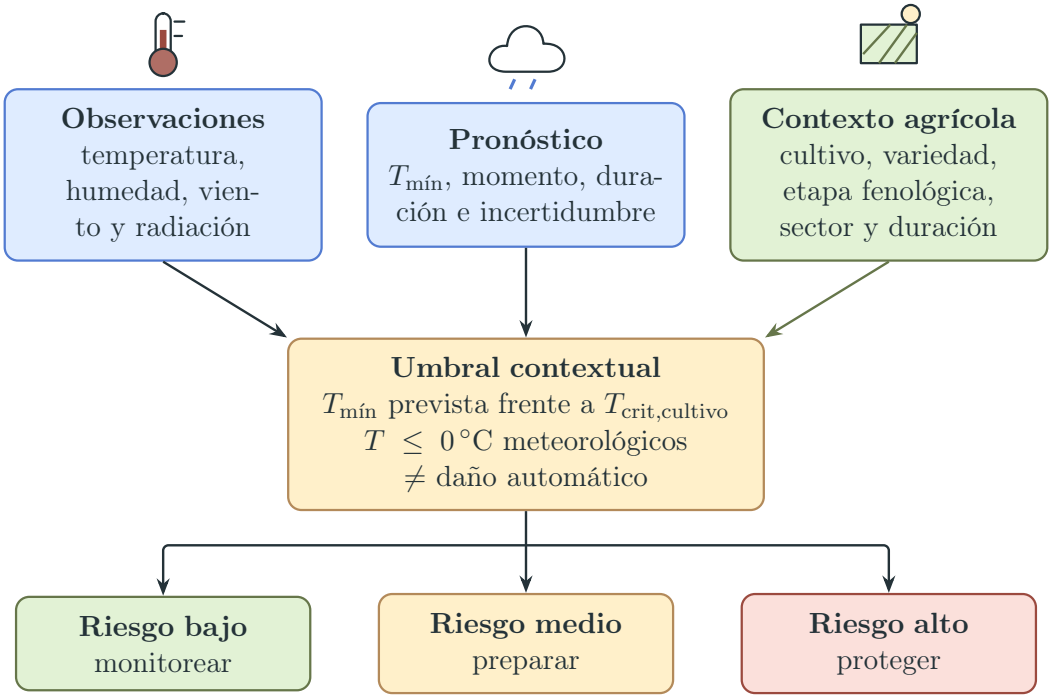


Figura 3.3: Conversión de observaciones y pronóstico en riesgo de helada agrícola mediante umbral térmico y contexto del cultivo.

La Figura 3.3 establece que las observaciones y el pronóstico son insumos, mientras el contexto agrícola les asigna significado operativo. El resultado es riesgo diferenciado con acciones asociadas, no una afirmación automática de daño al cultivo.

Cuando la salida es una temperatura, las métricas cuantifican magnitud y dirección del error. El error absoluto medio (MAE, del inglés *Mean Absolute Error*) resume la desviación en las unidades de la variable; la raíz del error cuadrático medio (RMSE, del inglés *Root Mean Squared Error*) penaliza con mayor fuerza los errores grandes; el error cuadrático medio (MSE, del inglés *Mean Squared Error*) expresa el promedio de los errores al cuadrado; y el sesgo medio (MBE, del inglés *Mean Bias Error*) permite identificar una tendencia sistemática a sobreestimar o subestimar. También se emplean coeficientes de correlación o determinación para describir asociación, aunque una correlación alta no garantiza errores pequeños [5, 23, 30].

Cuando la salida corresponde a la ocurrencia de una helada, la evaluación se construye a partir de aciertos, omisiones y falsas alarmas. La probabilidad de detección (POD, del inglés *Probability of Detection*) indica qué proporción de los eventos reales fue anticipada. La familia de falsas alarmas incluye la razón de falsas alarmas (FAR, del inglés *False Alarm Ratio*); bajo la convención de clasificación binaria adoptada, la tasa de falsos descubrimientos (FDR, del inglés *False Discovery Rate*) es equivalente, pues ambas se calculan como la proporción de falsos positivos entre predicciones positivas, $FP/(TP + FP)$. La medida F1 y el índice crítico de éxito (CSI, del inglés *Critical Success Index*) equilibran parcialmente detecciones, omisiones y falsas alarmas. La proporción correcta (PC, del inglés *Proportion Correct*) y la métrica de sesgo (BIAS, del inglés *Bias*) describen el acuerdo global y la tendencia a sobrealertar o subalertar; deben interpretarse junto con las métricas de evento cuando las heladas son infrecuentes.

Tabla 3.2: Familias de métricas utilizadas para interpretar modelos meteorológicos y de heladas.

Familia	Métricas representativas	Pregunta que permiten responder
Error continuo	MAE, RMSE, MSE, MBE	¿Cuánto se desvía la temperatura estimada y existe sesgo sistemático?
Detección de eventos	POD	¿Qué proporción de las heladas observadas fue detectada?
Falsas alarmas	FAR, FDR	¿Qué proporción de las predicciones positivas no correspondió a una helada observada?
Balance de clasificación	F1, CSI	¿Cómo se equilibra la detección frente a omisiones y falsas alarmas?
Desempeño global	PC, BIAS	¿Cuál es el acuerdo general y existe tendencia a sobrealertar o subalertar?

La Tabla 3.2 muestra por qué no conviene seleccionar una sola métrica: las familias de error, detección, falsas alarmas y desempeño global responden preguntas complementarias. En Chile, [23] reporta 98 % de exactitud y 86 % de sensibilidad, junto con una precisión media de 0,82 y una medida F1 de 0,84: la diferencia entre exactitud y sensibilidad muestra que un valor global alto no implica detectar todos los eventos. Para el índice de helada de [29], la evaluación a 72 horas obtuvo un CSI de 0,56 y una POD de 0,69, frente a un CSI de 0,37 y una POD de 0,44 con la temperatura directa del modelo regional. En [4], las métricas se calcularon por región y para horizontes de 24 y 72 horas, mostrando que el desempeño cambia según ubicación y anticipación.

El horizonte y la transferencia espacial también forman parte de la interpretación. En [30], el RMSE del RF aumentó de 1,72 °C a seis horas a 3,16 °C a 48 horas, mientras la medida F1 disminuyó de 0,95 a 0,92; el resultado muestra menor precisión térmica y menor equilibrio de clasificación al extender el plazo. En [5], los errores aumentaron

para plazos extendidos y se evaluaron por separado la temperatura y la ocurrencia de heladas. Por ello, cada resultado debe leerse con su horizonte, estación o región y umbral; comparaciones numéricas fuera de ese contexto pueden ser engañosas.

Finalmente, las métricas académicas no deben atribuirse automáticamente a servicios operativos. Las fuentes institucionales consultadas para WRF-DMC e INIA describen datos, productos y horizontes, pero no publican en esas páginas una evaluación operacional equivalente [24, 25]. Por ello, los resultados de [23] constituyen evidencia de un modelo evaluado con diez estaciones entre 2010 y 2017, no del desempeño actual del portal de INIA. Para una alerta agrícola, la evaluación debe combinar error de temperatura, detección de eventos, falsas alarmas, horizonte útil y comportamiento por zona.

3.4. Comunicación de predicciones y alertas meteorológicas

Más allá de la capacidad predictiva, los sistemas operativos deben transformar salidas técnicas en mensajes comprensibles y accionables para distintos destinatarios. En la cadena típica, una salida predictiva —variable numérica, campo espacial o índice— se contrasta con umbrales o categorías de riesgo; el resultado se dirige a un destinatario definido mediante un canal específico y, cuando corresponde, incluye instrucciones de mitigación. Conviene distinguir el **pronóstico meteorológico**, que describe condiciones futuras en términos físicos; el **aviso institucional**, emitido por un servicio oficial bajo criterios normados y cobertura geográfica delimitada; y la **alerta personalizada**, orientada a un usuario o sector con umbrales y acciones adaptadas a su contexto operativo.

En [31, 32] se describe el Sistema Europeo de Conciencia de Inundaciones (EFAS, del inglés *European Flood Awareness System*), sistema de alerta temprana para crecidas con pronósticos hidrológicos hasta diez días de anticipación. Sus productos se orientan a servicios hidrológicos nacionales, protección civil y consulta mediante portal institucional; no constituyen alertas directas a productores agrícolas ni implican canales de comunicación no documentados en la fuente.

Para calor extremo, [33] presenta HeatRisk, producto experimental del NWS con valor de riesgo a 24 horas y proyección a siete días. El producto clasifica niveles de 0 a 4 con codificación cromática, describe impacto esperado, grupos expuestos y recomendaciones de protección. El nivel de riesgo sintetiza severidad e impacto; no debe interpretarse necesariamente como probabilidad explícita del evento.

NWS HeatRisk: riesgo por calor

Producto experimental que complementa los avisos oficiales y orienta a quienes toman decisiones y a las personas sensibles al calor.

0 · Verde: escaso o nulo
Riesgo escaso o nulo por el calor previsto.

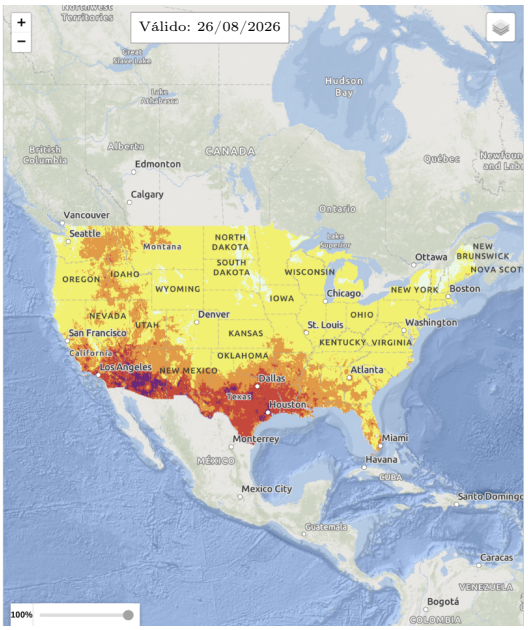
1 · Amarillo: menor
Afecta principalmente a personas muy sensibles al calor y sin refrigeración o hidratación.

2 · Naranja: moderado
Afecta a personas sensibles al calor, especialmente sin refrigeración o hidratación, y a algunos sistemas de salud y sectores productivos.

3 · Rojo: mayor
Afecta a cualquier persona sin refrigeración o hidratación, además de sistemas de salud y sectores productivos.

4 · Magenta: extremo
Calor excepcional o prolongado, sin alivio nocturno. Afecta a personas sin refrigeración o hidratación, sistemas de salud, sectores productivos e infraestructura.

Pronóstico válido para el 26/08/2026
Actualización: 26/08/2026, 11:48 (GMT-0400).



Población estimada por categoría en la captura: nivel 1, 212,0 millones; nivel 2, 53,8 millones; nivel 3, 48,6 millones; nivel 4, 17,1 millones. No se muestra una cifra para el nivel 0, lo que no equivale a población nula.

Figura 3.4: NWS HeatRisk: niveles de riesgo de impactos por calor de 0 a 4. Fuente: [33], captura del 26 de agosto de 2026. Adaptación al español de la leyenda y los rótulos; mapa recortado de la interfaz original, sin modificar el campo de riesgo.

La Figura 3.4 muestra la progresión desde el pronóstico físico hasta categorías de impacto y recomendaciones de respuesta. Estas categorías comunican impacto esperado, no una probabilidad explícita de ocurrencia del evento.

En [34,35] se documenta la experiencia canadiense en calificación de peligro de incendio forestal: variables meteorológicas y de combustible se combinan en índices operativos del Sistema Canadiense de Clasificación del Peligro de Incendio Forestal (CFFDRS, del inglés *Canadian Forest Fire Danger Rating System*). El Sistema Canadiense de Información sobre Incendios Forestales (CWFIS, del inglés *Canadian Wildland Fire Information System*) difunde información operativa y mapas de peligro a organismos y público general. Los índices expresan condiciones favorables al fuego; no equivalen a probabilidad de ignición.

En el caso de heladas, el NWS emite avisos diferenciados durante la temporada de crecimiento: *Frost Advisory*, cuando se espera helada generalizada habitualmente en el rango medio de los 30 °F (aproximadamente 1 °C–3 °C), y *Freeze Warning*, cuando se

prevén temperaturas bajo cero sostenidas en área extensa [36]. Los avisos geográficos se distribuyen mediante servicios web, API, JSON-LD, Atom y CAP versión 1.2 [26]. La documentación consultada describe estos canales y la redistribución por terceros, pero no identifica un servicio propio de mensajes de texto para estos productos.

En contraste, el Sistema de Alertas de Heladas por Mensajes de Texto de la Red Meteorológica Automatizada de Florida (FAWN, del inglés *Florida Automated Weather Network*) dirige mensajes mediante servicio de mensajes cortos (SMS, del inglés *Short Message Service*) a productores agrícolas [37]. El sistema entrega retroalimentación meteorológica en tiempo real y orientación sobre cuándo iniciar o detener el riego de protección contra heladas; constituye un canal directo y accionable hacia el usuario final, distinto del aviso institucional de cobertura territorial amplia.

CAP es un estándar de interoperabilidad para alertas de emergencia, no un modelo predictivo ni un canal de difusión. Especifica un formato estructurado para codificar evento, área afectada, vigencia, urgencia, severidad, certeza e instrucciones, y permite consistencia multicanal cuando servicios como el NWS u otras plataformas publican productos derivados de pronósticos distintos [38].

La iniciativa Alertas Tempranas para Todos (EW4All, del inglés *Early Warnings for All*) plantea un sistema de alerta temprana de extremo a extremo que articula conocimiento del riesgo, observación y pronóstico, difusión y comunicación, y preparación para responder [27]. La comunicación forma parte de ese proceso, no es una etapa aislada al final del pronóstico.



Figura 3.5: Componentes de la iniciativa Alertas Tempranas para Todos (EW4All). Fuente: [27]; título central adaptado al español, con los pictogramas de los cuatro componentes conservados.

La Figura 3.5 reúne conocimiento del riesgo, observación y pronóstico, comunicación y difusión, y preparación para responder como componentes conectados. La comunicación y la preparación no pueden separarse de la observación y el pronóstico, pues ambos

definen qué riesgo se comunica y cómo se prepara la respuesta.

Dos patrones de comunicación coexisten en la literatura y la práctica operativa revisada. El patrón **institucional e interoperable** —EFAS, avisos del NWS y encapsulamiento CAP— prioriza estandarización, trazabilidad y distribución a organismos o público mediante portales y API [26, 31, 32, 38]. El patrón **agrícola, directo y accionable** —FAWN— acorta la cadena hacia el productor e incorpora instrucciones operativas concretas [37]; la fuente consultada no documenta un formato interoperable equivalente a CAP.

Tabla 3.3: Transformación de pronósticos en comunicación para eventos meteorológicos o hidrometeorológicos seleccionados.

Sistema / evento	**Transformación del pronóstico**	**Destinatario / distribución**	**Limitación**
EFAS / crecidas [31, 32]	Pronóstico hidrológico → productos de alerta temprana	Servicios hidrológicos, protección civil; portal institucional	No alerta directa a agricultores; horizonte y resolución institucional
HeatRisk / calor [33]	Índice de riesgo 0–4 con impacto y recomendaciones	Público y planificadores; productos web del NWS	Nivel de riesgo ≠ probabilidad explícita del evento
CWFIS / incendios [34, 35]	Variables meteorológicas y de combustible → índices de peligro	Organismos y público; mapas e información operativa	Índice de peligro ≠ probabilidad de ignición
NWS / heladas [26, 36]	Umbrales térmicos → *Frost Advisory* o *Freeze Warning*	Cobertura geográfica; sitio web y API; formatos JSON-LD, Atom y CAP v1.2	Servicio propio de SMS del NWS no documentado en la fuente; redistribución por terceros
FAWN / heladas [37]	Condiciones locales → alerta con orientación de riego	Productores agrícolas; SMS	Alcance regional; interoperabilidad con CAP no documentada en la fuente

La Tabla 3.3 muestra que ningún sistema revisado demuestra simultáneamente cobertura institucional amplia, personalización agrícola directa e interoperabilidad completa. Esta comparación motiva combinar sus fortalezas como antecedente, sin afirmar todavía una implementación determinada.

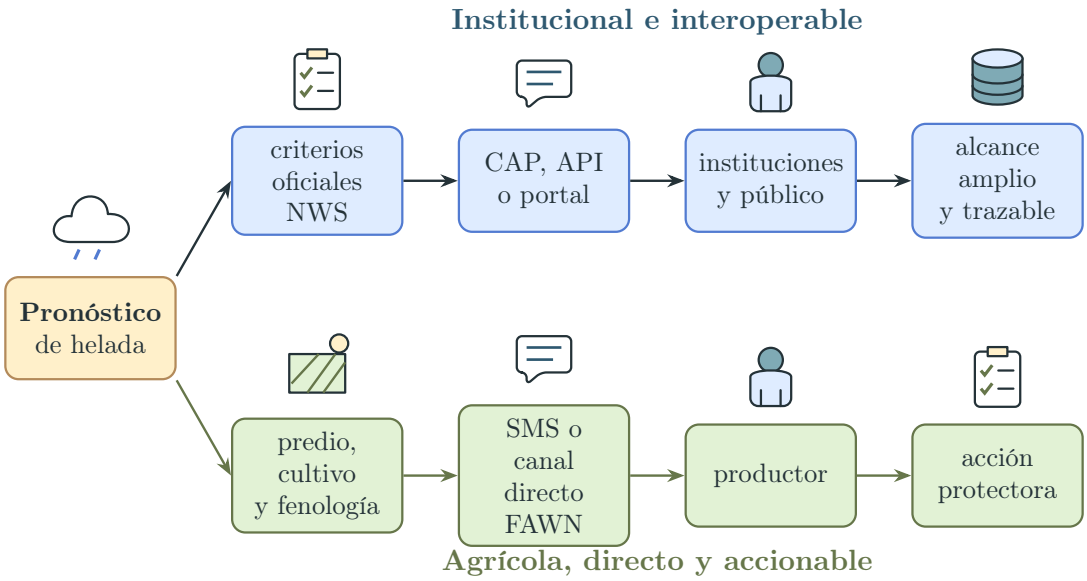


Figura 3.6: Dos rutas desde un pronóstico de helada: distribución institucional e interoperable y alerta agrícola directa y accionable.

La Figura 3.6 contrapone una ruta institucional de alcance amplio y trazable con otra directa, centrada en el contexto agrícola y la acción. Ambas son complementarias: privilegian beneficios distintos y no constituyen alternativas mutuamente excluyentes.

3.5. Sistemas existentes en Chile

En Chile, la predicción y el monitoreo de heladas se articulan principalmente a través de servicios meteorológicos institucionales y redes agrometeorológicas, complementados por estudios académicos evaluados en contextos locales.

La DMC, a través de MeteoChile, opera el modelo numérico WRF-DMC para pronóstico regional. Según la documentación institucional, las simulaciones utilizan condiciones iniciales y de borde del GFS; para la zona centro del país, las corridas de WRF-DMC se generan además con condiciones del ECMWF [24]. El horizonte operativo alcanza los próximos cinco días y se mantienen las últimas cinco corridas disponibles para contraste con observaciones. Entre los productos orientados a heladas figuran la distinción entre heladas blancas y negras, mapas y series del número de horas con temperatura a 2 m (T_{2m}) menor o igual a 0 °C, y estimaciones de la tasa máxima de enfriamiento nocturno [24,39]. WRF-DMC constituye un modelo de pronóstico atmosférico; no debe clasificarse como inteligencia artificial ni confundirse con un predictor de daño agrícola por sí solo.

El portal de monitoreo de heladas de MeteoChile integra estos productos para consulta pública [39]. Su aporte principal es cobertura territorial, continuidad operativa y respaldo institucional, aunque la resolución espacial del pronóstico puede no representar

microclimas de predios en relieve complejo.

La Red de Agrometeorología del INIA publica datos de estaciones de diversas instituciones, historial consultable por estación y fecha, exportación de registros y pronósticos a 24 y 48 horas en su sección de heladas [25]. El portal declara desplegar pronósticos basados en información de la DMC. Las funciones de consulta histórica y visualización de registros pasados corresponden a **monitoreo y archivo**; no demuestran que el algoritmo de pronóstico actual ejecute validación retrospectiva ni que use registros pasados para emitir la predicción del día siguiente.

En [23] se evaluaron ANN con retropropagación para la zona central de Chile. El trabajo empleó diez estaciones entre 2010 y 2017 y predijo la temperatura mínima del día siguiente y la ocurrencia de helada radiativa a partir de temperaturas horarias, radiación, viento, precipitación y punto de rocío. Reportó 98 % de exactitud, 86 % de sensibilidad y una tasa de error de clasificación de 2 % [23]. Este resultado evidencia un método evaluado en ese período, pero no prueba que el portal agrometeorológico del INIA en su versión actual ejecute esa red ni que reproduzca esas métricas en operación.

Además de las heladas, la DMC mantiene un Sistema de Alerta Meteorológica multiamenaza que distingue avisos, alertas y alarmas según severidad. La página institucional consultada muestra ejemplos para viento, altas temperaturas y tormentas eléctricas [40].

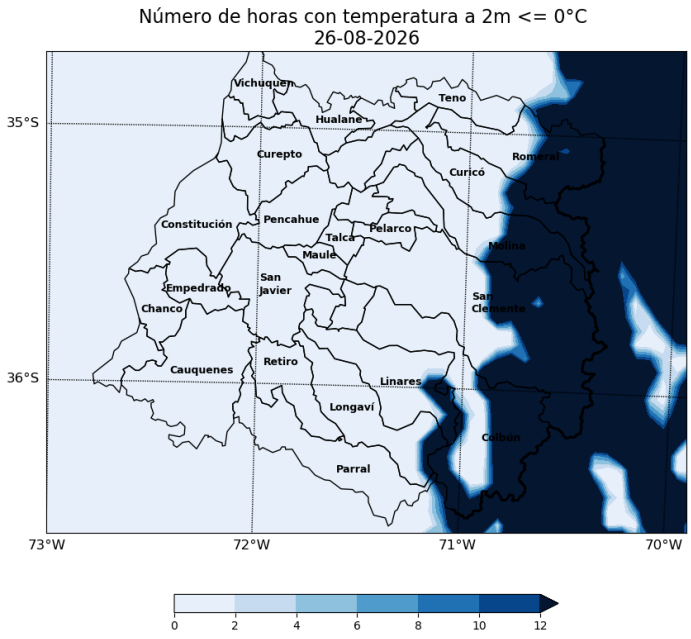


Figura 3.7: Pronóstico del número de horas con temperatura del aire a 2 m (T_{2m}) menor o igual a 0 °C para la Región del Maule, válido para el 26 de agosto de 2026. La escala cromática inferior de 0 a 12 representa la duración pronosticada en horas. Fuente: Dirección Meteorológica de Chile (DMC) [41].

La Figura 3.7 hace visible la heterogeneidad espacial del pronóstico y permite priorizar observación y respuesta en el territorio. Las horas bajo umbral describen duración térmica pronosticada, pero no equivalen a probabilidad de daño en el cultivo.

3.6. Hacia la integración de soluciones

Los antecedentes muestran aportes complementarios, pero no intercambiables. Las fuentes institucionales permiten acceder a observaciones y pronósticos; los estudios de aprendizaje automático evalúan métodos bajo conjuntos de datos, ubicaciones y horizontes determinados; y los sistemas de alerta organizan la comunicación hacia distintos destinatarios. Disponer de una de estas capacidades no resuelve automáticamente las restantes [4, 23, 25, 27].

Para el problema definido en este proyecto, la brecha se encuentra en la continuidad entre esas etapas. Los datos deben obtenerse de forma repetible, conservar su procedencia, normalizarse y vincularse con el contexto agrícola antes de alimentar un modelo. A su vez, la salida predictiva necesita una evaluación explícita y una forma de comunicación que el agricultor pueda comprender y utilizar. Una mejora del algoritmo no compensa por sí sola datos insuficientes, una asociación territorial incorrecta o una alerta que llega tarde.

Esto fundamenta el diseño de una herramienta que integre fuentes, procesamiento, predicción y comunicación, en lugar de presentar cada capacidad como una solución aislada. Su aporte esperado es reducir la distancia entre el dato disponible y la decisión agrícola, sin atribuirle desempeños obtenidos por otros estudios. El capítulo siguiente formula esa propuesta en términos generales; el caso de estudio posterior documenta las tecnologías y los resultados de su aplicación concreta.

Capítulo 4

Propuesta y diseño de solución

Los capítulos anteriores mostraron que disponer de datos meteorológicos no basta para apoyar una decisión agrícola: es necesario integrarlos, evaluar su calidad, producir una estimación de riesgo y comunicarla de manera comprensible. A partir de esa necesidad, este capítulo presenta el diseño de un sistema orientado a la predicción de heladas y a la entrega de información contextualizada al agricultor mediante una interfaz web y un servicio de mensajería.

La propuesta organiza el recorrido de la información desde las fuentes climáticas hasta la configuración de alertas por usuario. En concordancia con los objetivos del trabajo, se busca tanto la clasificación de eventos como la estimación de temperatura mediante regresión. La separación entre recolección, procesamiento, predicción e interacción permite identificar qué responsabilidad corresponde a cada componente y de qué información depende. La validación se incorpora como una actividad transversal: una arquitectura que conecta datos y usuarios no demuestra, por sí sola, la calidad de las predicciones ni su utilidad en terreno.

El capítulo describe los requisitos, la arquitectura, los módulos y las formas de interacción previstas, sin imponer una plataforma comercial ni una herramienta de desarrollo concreta. El Capítulo 5 presenta la implementación, sus prototipos y los riesgos asociados. Los requisitos y diagramas de esta propuesta no deben interpretarse como evidencia de que todas las capacidades ya estén operativas.

4.1. Requisitos del sistema

Los requisitos expresan las capacidades y condiciones de operación que orientan el diseño. Los identificadores RF1 a RF6 conservan la correspondencia del planteamiento original y acompañan a nombres descriptivos. Se utilizan en la matriz de trazabilidad para relacionar cada necesidad con los componentes responsables.

4.1.1. Requisitos funcionales

Los siguientes requisitos reúnen las necesidades de consulta, monitoreo y comunicación del sistema propuesto. Las consultas de inventario, aplicaciones y tareas forman parte de su contexto de gestión agrícola, siendo el foco predictivo la ocurrencia de heladas. Enumerar estas capacidades generales no implica que se hayan desarrollado o evaluado modelos para todas las amenazas mencionadas. En la implementación del Capítulo 5, la salida operativa del ensamble es una clase (voto) o un puntaje no calibrado; la «probabilidad» de RF5 y RF6 se entiende como objetivo de diseño y no como magnitud ya acreditada para disparar alertas.

RF1 Registro de información del campo: el usuario debe poder ingresar información de su campo y configuraciones mediante una conversación comprensible por un servicio de mensajería.

RF2 Consulta de información agrícola: el usuario debe poder consultar información disponible sobre inventario, aplicaciones, tareas e historial mediante el servicio de mensajería, sin conocer la estructura interna de los datos.

RF3 Monitoreo de datos meteorológicos: el sistema debe consultar periódicamente los datos de sensores o estaciones y los pronósticos de servicios externos para mantener actualizado el análisis de las condiciones climáticas.

RF4 Detección y registro de amenazas: el sistema debe identificar y registrar condiciones que puedan requerir una acción del usuario. El diseño general contempla amenazas meteorológicas y de gestión, como escasez de productos; la evaluación predictiva de este trabajo se concentra en las heladas.

RF5 Estimación del riesgo: el sistema debe estimar el riesgo de ocurrencia de la amenaza —como clase, score o probabilidad calibrada, según lo que el modelo y su validación permitan— y asociarlo al lugar y al horizonte temporal de la predicción. La resolución espacial y temporal que pueda ofrecer depende de las fuentes disponibles y del modelo evaluado.

RF6 Notificación de alertas: el sistema debe informar a los usuarios correspondientes mediante el servicio de mensajería cuando el criterio de alerta configurado —umbral sobre probabilidad, score o voto de clase— se cumpla.

4.1.2. Requisitos no funcionales

Estas condiciones orientan la calidad de servicio, la protección de los datos y la accesibilidad de la solución. Se formulan como requisitos de diseño, no como resultados de pruebas de rendimiento o disponibilidad.

RNF1 Capacidad de procesamiento: procesar y analizar datos históricos y actualizaciones operativas para alimentar el entrenamiento y la generación de predicciones.

RNF2 Escalabilidad: admitir múltiples fuentes de datos, como sensores, servicios meteorológicos e imágenes satelitales, sin una pérdida significativa de rendimiento. La carga de referencia y el deterioro aceptable requieren especificación antes de verificar este requisito.

RNF3 Disponibilidad: alcanzar una disponibilidad de al menos el 99,9 %, considerando la dependencia de servicios externos. Esta meta requiere definir el período de medición y el tratamiento de interrupciones externas; no se presenta como disponibilidad ya demostrada.

RNF4 Seguridad: transmitir datos mediante HTTPS con TLS y controlar el acceso a información sensible mediante autenticación y autorización basada en roles.

RNF5 Usabilidad: ofrecer consultas y alertas comprensibles para agricultores con experiencia técnica limitada, mediante lenguaje claro y acciones reconocibles.

RNF6 Calidad predictiva: evaluar las estimaciones de helada mediante detecciones, omisiones y falsas alarmas, y las estimaciones de temperatura mediante errores en grados. Los criterios de aceptación requieren justificarse según los datos, el horizonte y el uso previsto; no se adopta un porcentaje único de precisión.

RNF7 Personalización: incorporar configuraciones por usuario sin comprometer el funcionamiento general del sistema.

RNF8 Tolerancia a entradas ambiguas: solicitar correcciones o confirmaciones ante información incompleta o ambigua, en lugar de asumir datos no proporcionados.

RNF9 Persistencia de la configuración: almacenar de forma segura los parámetros asociados a la identidad del usuario y recuperarlos en sesiones posteriores.

4.1.3. Requisitos de la interfaz

La interfaz del sistema comprende tanto la interacción con el agricultor como el intercambio de información entre componentes. Los eventos se caracterizan por su iniciador, los datos de entrada y el resultado esperado; esta descripción funcional no sustituye una especificación completa de los contratos de las API ni de sus casos de error.

Evento	**Iniciador**	**Información de entrada**	**Resultado esperado**
Consulta de sensores	Tarea programada	Identidad del sensor, temperatura, humedad y fecha	Datos obtenidos y procesados
Actualización meteorológica	Tarea programada	Datos del servicio externo	Registros importados y almacenados
Recepción de mensaje	Servicio de mensajería	Usuario, texto y fecha	Mensaje entregado al backend
Registro de campo	Servicio de mensajería	Usuario, tipo de registro y contenido estructurado	Confirmación de registro guardado
Consulta agrícola	Servicio de mensajería	Usuario, tipo de consulta y filtros	Datos solicitados
Evaluación de amenaza	Proceso programado	Sensor, amenaza, valor y fecha	Riesgo estimado (clase, score o probabilidad)
Envío de alerta	Proceso por evento	Usuario, tipo de alerta, valor y criterio	Mensaje enviado al usuario
Configuración personal	Mensajería o interfaz web	Usuario, cultivo, ubicación y parámetros	Configuración aplicada y almacenada

Tabla 4.1: Eventos de intercambio de información previstos entre componentes.

La Tabla 4.1 distingue las operaciones iniciadas por tareas automáticas de aquellas originadas por el usuario. La primera ruta actualiza los datos y evalúa amenazas; la segunda aporta contexto agrícola y recupera información. Ambas convergen en la identificación del destinatario y en la configuración que determina cuándo corresponde comunicar una alerta.

4.1.4. Requisitos de ambiente

El entorno de soporte combina infraestructura de ejecución, herramientas de procesamiento y servicios externos. Las siguientes categorías permiten comparar alternativas de implementación sin adoptar una marca específica ni afirmar un orden de popularidad que no se haya documentado.

1. **Infraestructura:** servidores propios o servicios de cómputo en la nube para ejecutar la aplicación y entrenar modelos. La elección depende de la carga, el costo y la capacidad de administración disponible.
2. **Software:**
   - Interfaz web: tecnologías de navegador o bibliotecas de componentes que permitan consultar datos y administrar parámetros.
   - Backend y almacenamiento: servicios de aplicación y una base de datos que permita asociar registros, usuarios y campos y conservar su trazabilidad.

- Procesamiento y predicción: herramientas de análisis de datos y aprendizaje automático compatibles con las variables y los modelos seleccionados.
- Despliegue: ejecución directa o mediante contenedores, acompañada por control de versiones y automatización de pruebas y entregas.

3. **Red:** conectividad a Internet para acceder a fuentes y canales externos, con comunicaciones protegidas mediante HTTPS y TLS.
4. **Condiciones de uso:** licencias, cuotas y suscripciones de los servicios elegidos. Estas condiciones deben revisarse al definir la implementación concreta.

4.2. Diseño de la arquitectura del sistema

La arquitectura se describe en dos niveles complementarios: el contexto delimita las relaciones con el exterior y el esquema interno muestra el recorrido de los datos entre componentes. Esta separación permite distinguir el origen de la información, su transformación y la entrega de resultados al agricultor.

4.2.1. Diagrama de contexto del sistema

El contexto delimita las relaciones entre las fuentes climáticas, el sistema de predicción y el agricultor. La interacción se realiza mediante un servicio de mensajería para consultas y alertas y una interfaz web para consultar información y gestionar configuraciones.

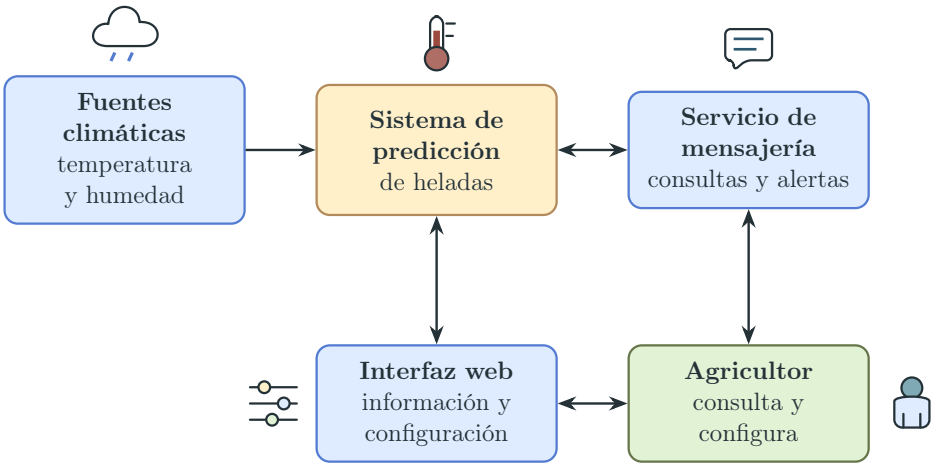


Figura 4.1: Contexto del sistema de predicción y sus intercambios con el agricultor.

La Figura 4.1 muestra que los datos climáticos ingresan al sistema y que el agricultor accede mediante mensajería o web, no directamente al modelo. Las flechas bidireccionales representan intercambios previstos, no conexiones operativas ya verificadas.

4.2.2. Diagrama de arquitectura

La arquitectura organiza las responsabilidades técnicas desde las fuentes hasta los canales de interacción. Los registros históricos y las observaciones se utilizan en la preparación y el entrenamiento; los pronósticos aportan información para generar estimaciones. El backend coordina la persistencia y el intercambio con las interfaces.

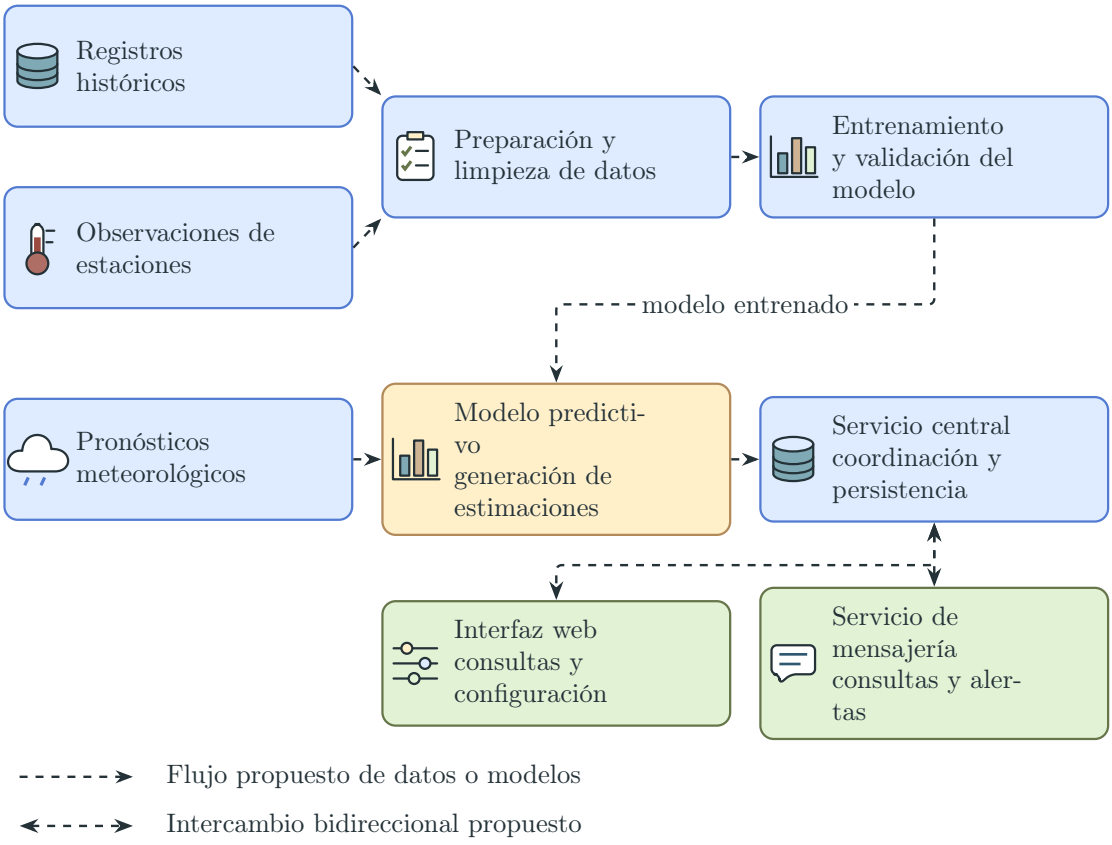


Figura 4.2: Arquitectura general del sistema

La Figura 4.2 distingue el flujo de preparación y entrenamiento del flujo que utiliza el modelo para producir estimaciones. Estas se entregan al backend, que las vincula con consultas, configuraciones y alertas. Las conexiones segmentadas representan intercambios previstos entre responsabilidades técnicas, no un despliegue comprobado ni una comunicación directa entre el modelo y el agricultor.

4.3. Módulos del sistema

Se conservan las seis responsabilidades del diseño original y sus identificadores M1 a M6. Primero se explica cada módulo; después se reúnen sus funciones y su relación con

los requisitos. La validación se entiende como una actividad transversal de evaluación y no como un paso obligatorio en cada envío de alerta.

4.3.1. M1: Recolección de datos climáticos

La recolección constituye el punto de entrada de los datos climáticos. Su propósito es obtener información desde fuentes externas y almacenarla en una base estructurada para que el procesamiento posterior no dependa de consultas manuales. El diseño contempla consultas periódicas, sujetas a la frecuencia de publicación de cada proveedor. Como se indicó al delimitar el problema, el acceso puede requerir una API, archivos descargables o programas de extracción; disponer de un portal público no implica que exista una API para todos sus productos.

- **Dependencias:** conectividad a Internet, servicios externos y credenciales de acceso vigentes.
- **Supuestos:** las fuentes se encuentran activas y publican registros utilizables; su calidad debe revisarse antes del modelamiento.
- **Restricciones:** cuotas de consulta, cambios de formato y diferencias de actualización entre proveedores.

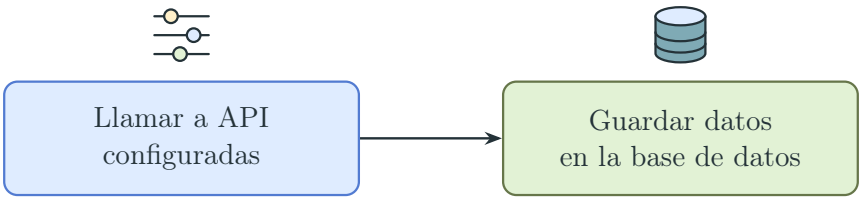


Figura 4.3: Obtención y almacenamiento de datos climáticos.

La Figura 4.3 representa el caso de consulta mediante API y separa el acceso a la fuente del almacenamiento. Guardar un registro permite reutilizarlo, pero no acredita que esté completo o sea consistente con los demás. Por ello, el resultado de esta etapa es una entrada para el procesamiento, no un conjunto listo para entrenar el modelo; otros mecanismos de extracción deben conservar la misma relación entre origen y registro almacenado.

4.3.2. M2: Procesamiento y limpieza de datos

El procesamiento transforma los registros recolectados en un conjunto de datos utilizable para la predicción de heladas. Comprende el tratamiento de valores faltantes, la revisión de inconsistencias y la selección de variables. Su salida debe permitir que

el modelo opere sobre información comparable y no sobre diferencias accidentales de formato. La normalización de fechas, unidades e identificadores y la conservación de la procedencia responden a las brechas de integración y trazabilidad descritas en el capítulo de definición del problema.

- **Dependencias:** registros entregados por M1 y reglas de transformación y selección de variables.
- **Supuestos:** existen variables relevantes y cobertura suficiente para el análisis. La suficiencia de la limpieza automática debe verificarse, no darse por demostrada.
- **Restricciones:** volumen procesable con los recursos disponibles y pérdida de información que pueda producir el tratamiento de registros incompletos o atípicos.

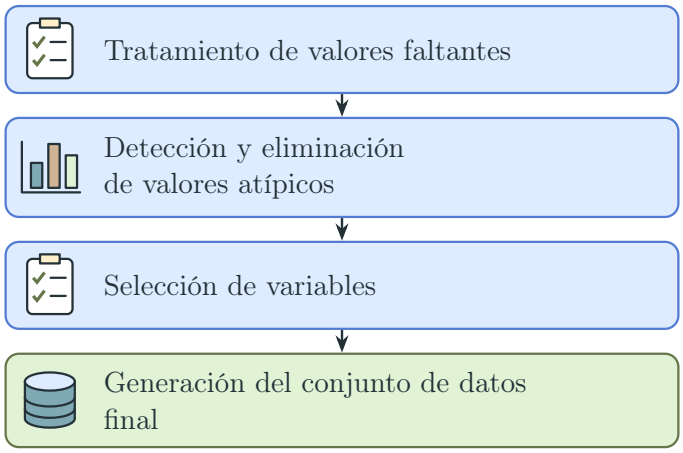


Figura 4.4: Preparación de los datos para el modelamiento; esquema con etiquetas en español.

La Figura 4.4 organiza la limpieza antes de la selección final de variables. El bloque de valores atípicos requiere especial cuidado: una temperatura extrema puede corresponder al fenómeno que se desea predecir y no a un error. Su eliminación no debe ser indiscriminada; la interpretación de cada registro condiciona la representatividad del conjunto resultante.

4.3.3. M3: Modelo predictivo

El modelo predictivo utiliza los datos preparados para aprender relaciones asociadas a la ocurrencia de heladas y a la evolución de la temperatura. El diseño distingue una salida de clasificación, que indica evento o ausencia de evento, y una salida de regresión, que estima una variable continua. La selección de algoritmos y bibliotecas corresponde a la implementación concreta. Una decisión de clase no equivale por sí sola

a una probabilidad calibrada. La estimación probabilística requerida para las alertas debe evaluarse por separado. Del mismo modo, el umbral térmico que define un evento es distinto del umbral de probabilidad que activa una notificación.

- **Dependencias:** conjunto de datos preparado por M2 y herramientas de modelamiento y evaluación compatibles con el problema.
- **Supuestos:** los datos tienen representatividad histórica y las métricas permiten evaluar el evento de interés bajo un horizonte definido.
- **Restricciones:** costo computacional del entrenamiento y disponibilidad de ejemplos suficientes para evaluar la generalización.

La clasificación debe evaluarse mediante detecciones, omisiones y falsas alarmas, mientras que la regresión requiere interpretar errores de temperatura, como el RMSE, en las unidades de la variable. Los criterios de aceptación deben justificarse para el evento, el horizonte, los datos y el uso previsto; no se fija un porcentaje general independiente de esas condiciones.

El umbral operativo elegido para un experimento debe distinguirse de la referencia meteorológica de 0 °C y de las temperaturas críticas de cada cultivo. El desempeño obtenido en un conjunto de datos tampoco se traslada automáticamente a otros predios ni convierte toda clasificación positiva en una predicción de daño agrícola.

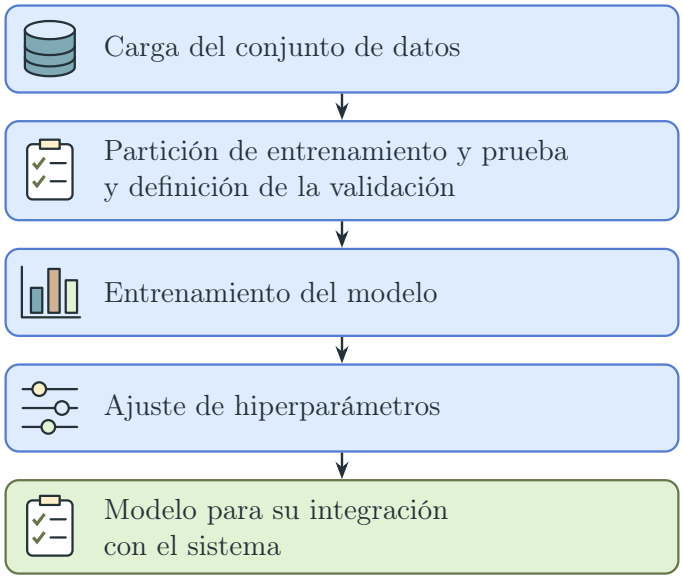


Figura 4.5: Etapas de preparación del modelo predictivo.

La Figura 4.5 representa el paso desde los datos hasta un modelo utilizable por los demás componentes. El ajuste pertenece al proceso de desarrollo: los datos reservados

para prueba no deben utilizarse para seleccionar hiperparámetros. Entregar el modelo a la integración tampoco demuestra que sus resultados sean adecuados en terreno; esa conclusión requiere la evaluación descrita en el capítulo de implementación.

4.3.4. M4: Integración conversacional con un servicio de mensajería

La integración conversacional tiene dos funciones previstas: enviar alertas originadas por el análisis climático y responder consultas del agricultor. El canal debe presentar los resultados con contexto comprensible, sin exigir que el usuario conozca el modelo ni los servicios que procesan sus datos.

- **Dependencias:** interfaz de integración del servicio de mensajería, estimaciones de M3 y backend que coordine la comunicación.
- **Supuestos:** el usuario tiene acceso al canal elegido y su identidad se encuentra correctamente registrada.
- **Restricciones:** políticas y cuotas del servicio de mensajería; consultas acotadas a los casos previstos y a la información disponible.

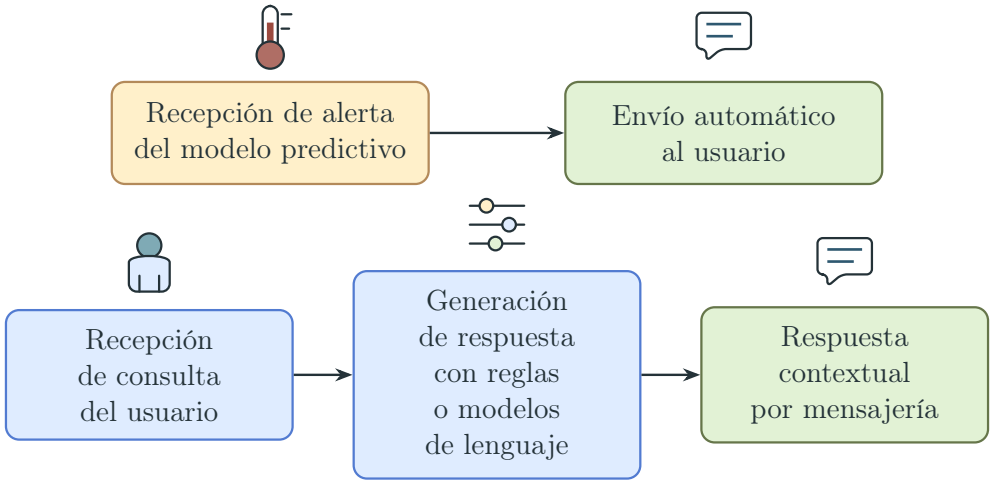


Figura 4.6: Rutas propuestas de alerta y consulta mediante un servicio de mensajería.

La Figura 4.6 muestra dos iniciadores distintos: el sistema produce una alerta sin esperar una pregunta, mientras que el usuario inicia la ruta de consulta. La generación de lenguaje permite presentar información, pero no debe confundirse con el cálculo de la predicción meteorológica. Esta separación facilita identificar qué resultado proviene del modelo y cuál corresponde a su comunicación.

4.3.5. M5: Selección de parámetros

La selección de parámetros asocia el contexto agrícola con la identidad del usuario. Su propósito es conservar información y preferencias que permitan configurar el monitoreo de cada campo, en lugar de aplicar una única configuración a todos los agricultores.

- **Dependencias:** interfaz de configuración, integración conversacional de M4 y resultados de M3 que deban interpretarse con esos parámetros.
- **Supuestos:** el agricultor conoce la información básica de su huerto y puede proporcionarla mediante los canales disponibles.
- **Restricciones:** las respuestas pueden ser ambiguas o incompletas y no siempre existe una fuente independiente para confirmar su exactitud.

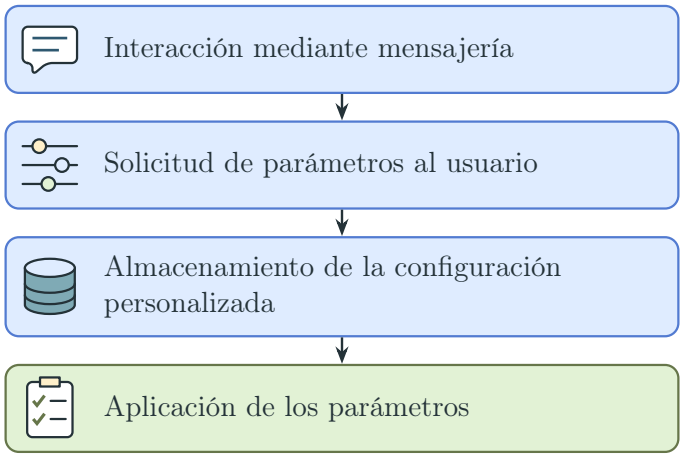


Figura 4.7: Obtención y uso de parámetros personalizados.

La Figura 4.7 distingue obtener una respuesta, conservarla y utilizarla en la configuración. La persistencia evita solicitar nuevamente la misma información en cada sesión, pero no garantiza que el dato sea correcto. Ante ambigüedades, el requisito de confirmación debe aplicarse antes de interpretar una respuesta como una configuración válida.

4.3.6. M6: Validación del sistema

La validación busca determinar si las estimaciones y su forma de entrega responden al propósito del sistema. El diseño considera comparar predicciones con eventos observados y explorar, mediante simulaciones, su relación con decisiones agrícolas. Esta actividad es transversal y su alcance efectivo depende de la evidencia disponible.

- **Dependencias:** resultados de M3, datos históricos y el canal de M4 para recoger información del agricultor.

- **Supuestos:** existen registros climáticos y agrícolas comparables para el mismo período, y los escenarios simulados representan razonablemente las condiciones de interés.
- **Restricciones:** acceso limitado a datos reales de usuarios y dificultad para cuantificar el impacto económico de una decisión.

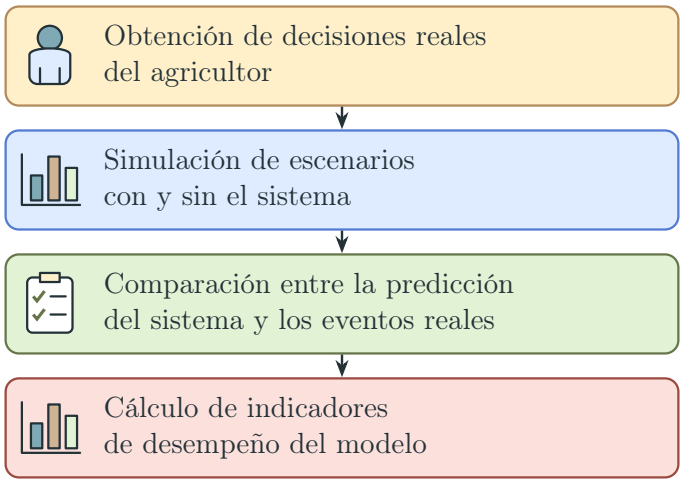


Figura 4.8: Esquema propuesto de validación mediante registros y simulaciones.

La Figura 4.8 relaciona decisiones del agricultor, escenarios simulados y eventos observados con el cálculo de indicadores. La comparación exige datos de referencia: una simulación sin registros suficientes no permite concluir que exista un beneficio económico o una mejora de decisiones en terreno. Los resultados efectivamente obtenidos deben distinguirse de este plan de evaluación.

4.3.7. Descripción general de módulos

Las responsabilidades anteriores forman un conjunto: M1 y M2 preparan los datos, M3 genera estimaciones, M4 y M5 articulan la interacción y M6 evalúa el sistema. Los identificadores permiten relacionar esa síntesis con las subsecciones de detalle sin sustituir sus nombres descriptivos.

ID	Módulo	Propósito	Sección
M1	Recolección de datos climáticos	Obtener y almacenar datos externos.	4.3.1
M2	Procesamiento y limpieza de datos	Depurar registros y preparar variables.	4.3.2
M3	Modelo predictivo	Entrenar y evaluar modelos para generar estimaciones.	4.3.3
M4	Integración conversacional	Comunicar alertas y responder consultas.	4.3.4
M5	Selección de parámetros	Asociar configuraciones al usuario y su contexto.	4.3.5
M6	Validación del sistema	Evaluar desempeño y utilidad con datos de referencia.	4.3.6

Tabla 4.2: Responsabilidades y ubicación de los módulos en el diseño.

La Tabla 4.2 distingue el tratamiento de datos de la interacción con el usuario. La personalización aporta contexto al uso de las estimaciones y la validación examina el comportamiento del conjunto. Esta separación describe responsabilidades, no implica que cada módulo se despliegue como un servicio independiente.

4.3.8. Matriz de requisitos funcionales y módulos

La asignación conserva las relaciones del diseño original y utiliza los requisitos de la Sección 4.1.1. Cada marca indica participación en la implementación o evaluación de una capacidad, no evidencia de cumplimiento.

	M1	M2	M3	M4	M5	M6
RF1				X	X	
RF2				X		
RF3	X	X				
RF4		X	X			
RF5			X			
RF6				X		X

Tabla 4.3: Trazabilidad entre requisitos funcionales y componentes responsables.

La Tabla 4.3 muestra que algunas capacidades requieren colaboración entre módulos. El monitoreo (RF3) combina M1 y M2; el registro de configuraciones (RF1) vincula M4 y M5. La participación de M6 en RF6 corresponde a su evaluación, mientras M4 realiza la comunicación.

4.3.9. Matriz de requisitos no funcionales y módulos

Los requisitos no funcionales de la Sección 4.1.2 atraviesan más de un componente.

	M1	M2	M3	M4	M5	M6
RNF1	X	X	X			
RNF2	X	X				
RNF3	X		X	X	X	
RNF4	X			X	X	
RNF5				X	X	X
RNF6			X			X
RNF7				X	X	
RNF8				X	X	
RNF9				X	X	

Tabla 4.4: Trazabilidad entre requisitos no funcionales y módulos del sistema.

La Tabla 4.4 relaciona la calidad predictiva (RNF6) con el modelo y su validación (M3 y M6), mientras que usabilidad y persistencia dependen de la interacción y la configuración (M4 y M5). Una marca expresa responsabilidad de diseño o evaluación, no una prueba de rendimiento, seguridad o disponibilidad.

4.4. Diseño de interfaces

Las interfaces conectan el resultado técnico con las tareas del agricultor: consultar, reconocer el campo de interés y configurar su monitoreo. Se describe el comportamiento previsto de los canales sin adoptar una plataforma específica. Los prototipos del caso concreto se presentan en el Capítulo 5.

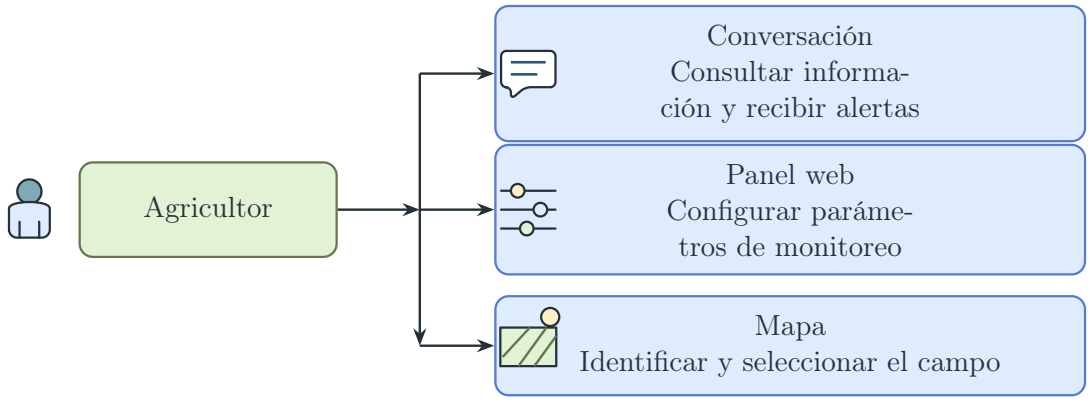


Figura 4.9: Relación entre las tareas del agricultor y las interfaces propuestas.

La Figura 4.9 resume tres tareas complementarias:

1. La conversación facilita consultas de información y recepción de alertas.
2. El panel web permite configurar los parámetros de monitoreo.
3. El mapa apoya la identificación y selección del campo.

Esta organización expresa el diseño propuesto; no acredita plataformas independientes ni una integración operativa.

4.4.1. Modelo de navegación

La navegación comienza con la identificación del usuario y la asociación de sus campos. Una consulta recibida por mensajería necesita ese contexto; si falta, la interfaz web permite iniciar sesión y registrar o seleccionar el campo antes de continuar.

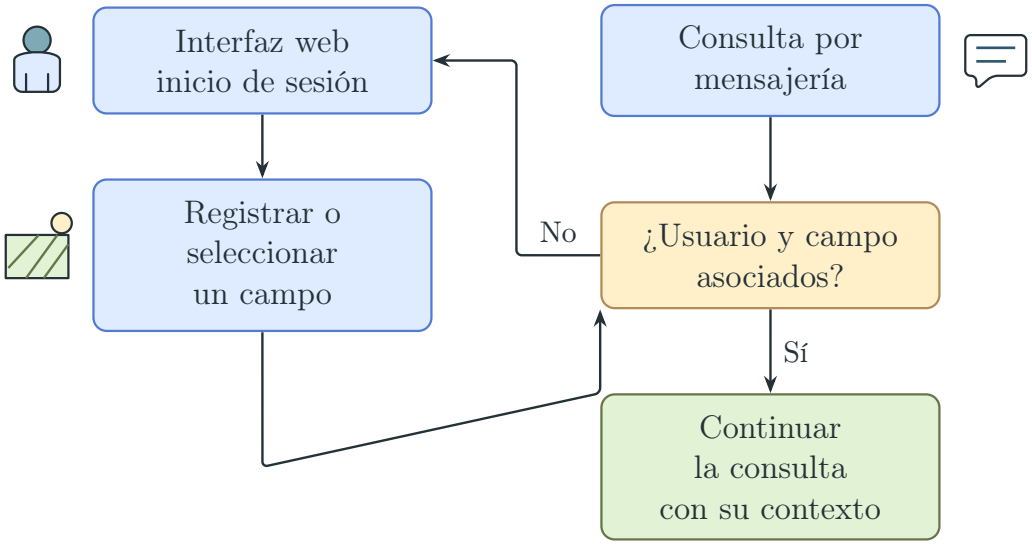


Figura 4.10: Navegación del usuario entre mensajería y configuración del campo.

La Figura 4.10 muestra que una consulta sin usuario o campo asociado requiere completar ese contexto en la interfaz web. Una vez disponible, el servicio de mensajería puede continuar la consulta; el recorrido representa navegación propuesta y no confirma un registro o una respuesta ya verificados.

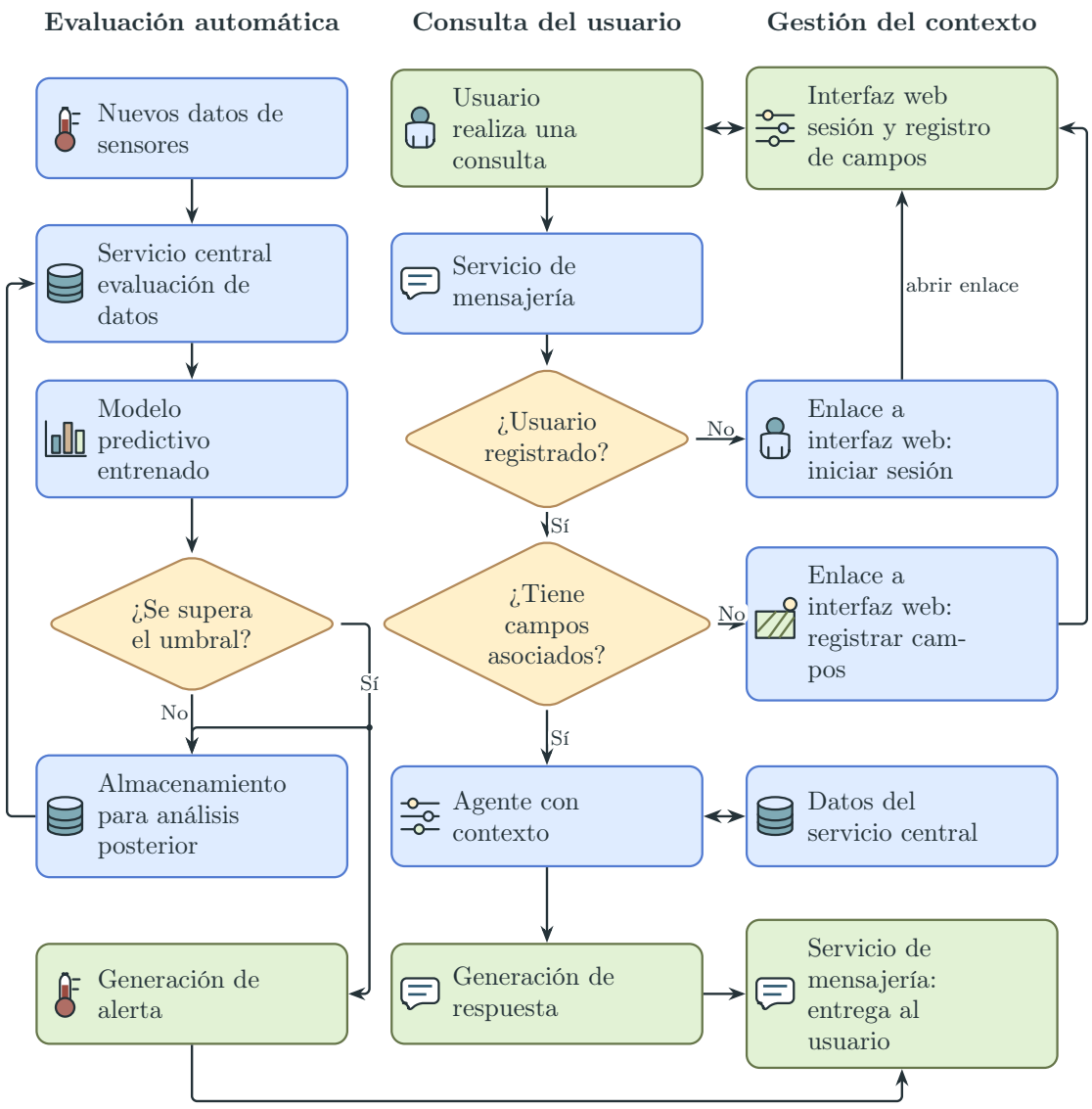


Figura 4.11: Flujo propuesto de consultas y generación de alertas.

La Figura 4.11 distingue las condiciones para preparar una alerta de las verificaciones necesarias para responder una consulta. La alerta solo se genera cuando la evaluación automática supera el umbral de RF6; una consulta que llega al agente produce una respuesta, no un aviso. Si falta identidad o información de campos, se ofrece un enlace para completarla en la interfaz web. Los resultados se conservan para análisis, mientras el agente consulta el backend y prepara la comunicación. El servicio de mensajería aparece al inicio y al final para separar recepción y entrega; ambos bloques representan el mismo canal, no dos servicios distintos.

4.4.2. Diseño de interfaces usuarias

Se distinguen tres formas complementarias de interacción: la conversación para consultar y recibir alertas, el panel web para administrar parámetros y el mapa para asociar la información a sectores agrícolas. La conversación debe comunicar ubicación y momento; el panel debe hacer visible la configuración activa; y el mapa debe distinguir selección de un campo de modificación de su registro.

Esta organización describe necesidades de interacción, no una interfaz comercial obligatoria. Los prototipos correspondientes se presentan en el caso de estudio, donde sus marcas y valores ilustrativos pueden interpretarse sin confundirlos con requisitos generales ni resultados del modelo.

4.5. Síntesis de la propuesta

La propuesta conecta fuentes climáticas, preparación de datos, predicción e interacción mediante M1–M6. Las matrices de requisitos funcionales y no funcionales (Tablas 4.3 y 4.4) relacionan cada necesidad con los módulos que participan en su realización o evaluación. Así, el diseño no se reduce al algoritmo: también explicita de dónde provienen los datos, cómo se utilizan y por qué canal llegan al usuario.

Su aplicación exige seleccionar tecnologías, delimitar los datos de evaluación y comprobar las integraciones. El Capítulo 5 desarrolla esa implementación y distingue los resultados reportados de las capacidades todavía pendientes de validación. La arquitectura propuesta orienta esa implementación, pero no acredita por sí sola desempeño predictivo ni utilidad en terreno.

Capítulo 5

Implementación de la solución

Este capítulo presenta la aplicación del diseño al caso agrícola: la obtención y preparación de datos climáticos, la evaluación de modelos para estimar heladas y la integración de sus resultados con las interfaces del usuario. Se examinan los componentes, los prototipos y la evidencia disponible para distinguir los resultados documentados de las verificaciones necesarias para su operación.

Miido es la plataforma de gestión territorial empleada en este caso: permite organizar predios, grupos de secciones y parámetros de monitoreo. WhatsApp está integrado como canal de comunicación para consultas y alertas. Ambas interfaces cumplen funciones diferentes y conectan el resultado técnico con el contexto del agricultor; su integración no equivale a una validación en terreno. Las fuentes consideradas incluyen al Instituto de Investigaciones Agropecuarias (INIA), el Atlas de Riesgo Climático y MeteoChile. Las cifras de clasificación y la descripción del ensamble provienen del trabajo grupal del proyecto, no de un nuevo experimento realizado para este capítulo.

5.1. Marco General

Este marco resume los avances comprobables antes de describir cómo se obtuvieron los datos, cómo se evaluaron los modelos y cómo se integraron las interfaces. La revisión de requisitos y diseño se presenta después de los resultados, en la sección de limitaciones.

5.1.1. Hitos principales del proyecto

Los hitos documentados corresponden a tareas realizadas; las pruebas con agricultores no formaron parte del alcance de este proyecto.

- Recolección y limpieza de datos climáticos: consolidación documentada de más

de 750.000 registros históricos y estandarización de variables como temperatura, humedad relativa y precipitación.
- Entrenamiento preliminar del modelo predictivo: implementación de algoritmos de aprendizaje automático y resultados iniciales de desempeño.
- Integración de interfaces: Miido permite gestionar campos y grupos de secciones, y WhatsApp aporta el canal de comunicación.

5.2. Obtención de datos

La obtención de datos respondió a la necesidad de construir una base utilizable para anticipar bajas temperaturas, no solo de reunir archivos de distintos proveedores. Las fuentes mencionadas en el caso son la Red Agrometeorológica de INIA, MeteoChile y el Atlas de Riesgo Climático. Se consideran registros de estaciones, antecedentes históricos y productos meteorológicos, cuya disponibilidad y resolución no son necesariamente equivalentes. La selección priorizó temperatura mínima y horaria, humedad relativa, radiación solar y precipitación; el viento se incorporó como variable de interés cuando la fuente lo permitía. Fecha, hora y ubicación asociada son indispensables para interpretar cada medición. No todas las variables deben suponerse presentes en todas las fuentes ni en todos los períodos.

Dato seleccionado	**Función en la preparación**
Temperatura mínima y horaria	Identificar eventos bajo el umbral y describir la evolución térmica.
Humedad relativa, radiación solar y precipitación	Aportar contexto meteorológico cuando la fuente dispone de esas variables.
Viento	Incorporar una variable adicional cuando existe en la fuente consultada.
Fecha, hora y ubicación	Asociar cada registro con un período y lugar para comparar series.

Tabla 5.1: Variables consideradas para preparar los registros climáticos.

La Tabla 5.1 separa las variables climáticas de las claves temporales y espaciales necesarias para interpretarlas. No atribuye cobertura uniforme a todas las fuentes ni confirma que todas las variables se hayan usado en cada modelo.

La documentación reporta la consolidación de más de 750.000 registros climáticos desde múltiples fuentes. Ese volumen corresponde al corpus de trabajo del caso —la base amplia sobre la que se construye el modelamiento— y no al tamaño de la tabla de métricas. Pudahuel aparece más adelante solo como comuna de ejemplo para reportar

la evaluación de clasificación: no agota el corpus ni implica que el modelo se haya entrenado únicamente ahí.

5.2.1. Pronóstico de referencia de MeteoChile

La Dirección Meteorológica de Chile (DMC) publica pronósticos con su implementación regional del modelo de investigación y pronóstico del tiempo (WRF, del inglés *Weather Research and Forecasting*), denominada WRF-DMC. Se trata de un modelo numérico de la atmósfera: simula su evolución a partir de condiciones iniciales y de borde, en lugar de aprender una regla de clasificación a partir del conjunto histórico de este proyecto [42]. Según la documentación institucional, utiliza información del Sistema Global de Pronóstico (GFS, del inglés *Global Forecast System*) y, en la zona centro, también del Centro Europeo de Predicción Meteorológica a Plazo Medio (ECMWF, del inglés *European Centre for Medium-Range Weather Forecasts*). El meteograma de estación presenta variables pronosticadas para los próximos cinco días [24]. Se eligió como referencia externa porque permite consultar una predicción meteorológica institucional asociada a una estación y contrastarla, en principio, con la estimación propia para la misma zona y período. Su cobertura regional no convierte la estación más cercana en una medición del predio ni garantiza que el pronóstico represente su microclima.

5.3. Estrategia de extracción: API, archivos y programas de extracción

La documentación de implementación distingue tres mecanismos de acceso, elegidos según la forma de publicación: solicitudes a servicios estructurados, descarga de archivos y extracción desde interfaces web. La fuente determina el mecanismo; no se asume que todas las instituciones ofrezcan una API ni que cada fuente haya requerido los tres métodos. Los conectores se describen como unidades separadas para contener los cambios de formato y navegación.

5.3.1. Llamadas a API

Cuando existe un servicio estructurado, el conector consulta el período o la entidad de interés y transforma la respuesta en registros procesables. Su responsabilidad es resolver el acceso a la fuente sin trasladar al modelo los detalles del protocolo. Para interpretar una extracción se deben conservar el origen, la fecha consultada y el resultado de la solicitud, distinguiendo una respuesta vacía de una falla de acceso.

Los documentos revisados no especifican endpoints, políticas de reintento ni credenciales

operativas. Por ello, no se atribuyen al prototipo frecuencias efectivamente medidas ni garantías de disponibilidad de una API concreta. La periodicidad diaria aparece como objetivo de automatización en la planificación, no como una medición de continuidad del servicio.

5.3.2. Descarga de archivos masivos

Los archivos permiten incorporar registros históricos sin efectuar una consulta individual por cada observación. Su lectura exige reconocer encabezados, codificación, separadores y representación de fechas y unidades antes de concatenar filas. Una descarga correcta no implica que su contenido ya sea compatible con el conjunto utilizado para entrenar.

La documentación contempla conservar los datos recibidos antes de normalizarlos. Esa separación permite revisar una transformación o repetir la lectura si se detecta un error, sin sustituir silenciosamente el material de origen. No se fija un formato único de archivo para todas las fuentes, dado que el inventario consultado no establece esa correspondencia.

5.3.3. Programas de extracción

Cuando el dato se publica mediante páginas destinadas a consulta humana, los programas de extracción automatizan el acceso y la lectura de los elementos relevantes. Estos conectores dependen de la estructura de la página y de su navegación, por lo que un cambio visual puede impedir la obtención aunque el dato continúe publicado.

El aislamiento por fuente es la decisión documentada para limitar ese impacto. La salida debe mantener la misma información de procedencia que las otras vías de acceso. Los prototipos de extracción emplean Playwright para abrir el meteograma WRF de una estación de MeteoChile, leer las series del gráfico y preparar archivos de valores separados por comas (CSV, del inglés *Comma-Separated Values*) con variable, tiempo y valor. Aunque el producto consultado ofrece hasta cinco días de pronóstico, el cálculo ilustrativo de indicadores de helada se limita a tres días locales; no es una predicción del modelo entrenado en este proyecto. Las direcciones de observaciones recientes que también aparecen en la fuente no deben confundirse con el meteograma pronosticado. Estos prototipos no ejecutan la comparación entre la estimación propia y WRF-DMC.

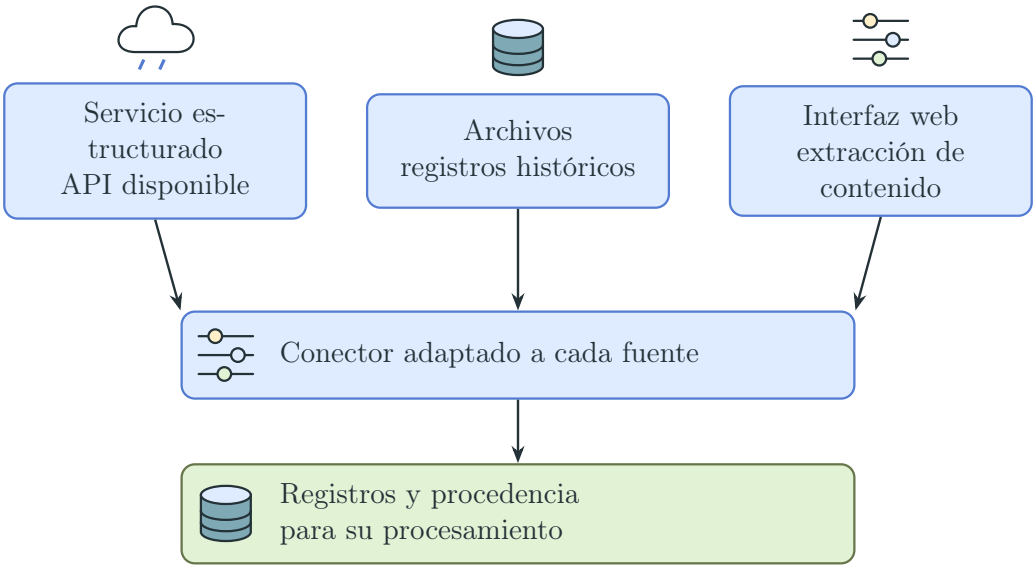


Figura 5.1: Mecanismos de extracción descritos en la documentación del caso.

La Figura 5.1 separa el mecanismo de acceso de la salida que necesita el procesamiento. Cada conector resuelve las particularidades de una fuente, mientras los registros conservan su procedencia. Las tres entradas representan alternativas de acceso, no una asignación verificada de un método a cada proveedor ni una exigencia de ejecutarlos simultáneamente.

5.4. Flujo de datos

El flujo documentado distingue la información recibida, sus transformaciones y los resultados del modelamiento. Conservar una versión bruta permite volver al registro original; normalizar y limpiar genera una base comparable; consolidar prepara las entradas de clasificación y regresión. La publicación posterior debe asociar cada salida con el momento y la ubicación a los que se refiere.

El entorno descrito asigna Node.js con Express al servicio que organiza las solicitudes y MySQL a la persistencia. Node.js permite atender operaciones de entrada/salida mientras se espera una fuente externa o la base de datos [43]; Express estructura las rutas HTTP [44]. Estas tecnologías no garantizan por sí solas la continuidad de la obtención de datos ni la entrega de alertas.

Esta separación evita confundir una serie descargada con una predicción propia o con una alerta entregada. A partir de las coordenadas del campo se consultan el pronóstico WRF-DMC de una estación cercana y la estimación propia para la comuna. Las respuestas se utilizan para decidir, aunque no se dispone de registros de la coordinación completa. Para compararlas deben coincidir la zona, variable, fecha prevista, horizonte

y definición del evento: una clase diaria bajo 3 °C no equivale a una temperatura horaria bajo 0 °C. El acuerdo entre pronósticos tampoco mide su acierto frente a observaciones posteriores.



Figura 5.2: Flujo de obtención, preparación y publicación de resultados climáticos.

La Figura 5.2 distingue registros originales, datos preparados y salidas del modelo. Esa separación permite rastrear una anomalía hasta su fuente; registrar un resultado no confirma que se haya comunicado al agricultor.

5.5. Limpieza y estandarización

La preparación comprende la depuración, transformación y selección de variables para el modelamiento. Incluye estandarizar los formatos, tratar valores faltantes, eliminar duplicados o inconsistencias y consolidar los registros en una estructura común. Ese recorrido corresponde a M2, cuyo resultado alimenta al modelo predictivo.

La estandarización permite comparar registros de distintas fuentes sin confundir su significado. Por ejemplo, una temperatura mínima diaria y una lectura horaria no son intercambiables aunque ambas estén expresadas en grados Celsius. Fecha, ubicación, variable y unidad deben conservarse junto con la procedencia. El tratamiento de faltantes depende del dato y del uso previsto; una misma regla de imputación no debe

aplicarse indiscriminadamente a todas las fuentes.

La eliminación de duplicados exige comparar fuente, ubicación, instante y variable: dos filas con esas mismas claves pueden repetir una medición, mientras una temperatura y una humedad tomadas a la misma hora son registros distintos. Un valor atípico se contrasta con el instrumento y los registros vecinos antes de excluirlo. Una temperatura bajo cero puede ser precisamente la helada buscada y no debe filtrarse solo por ser extrema.

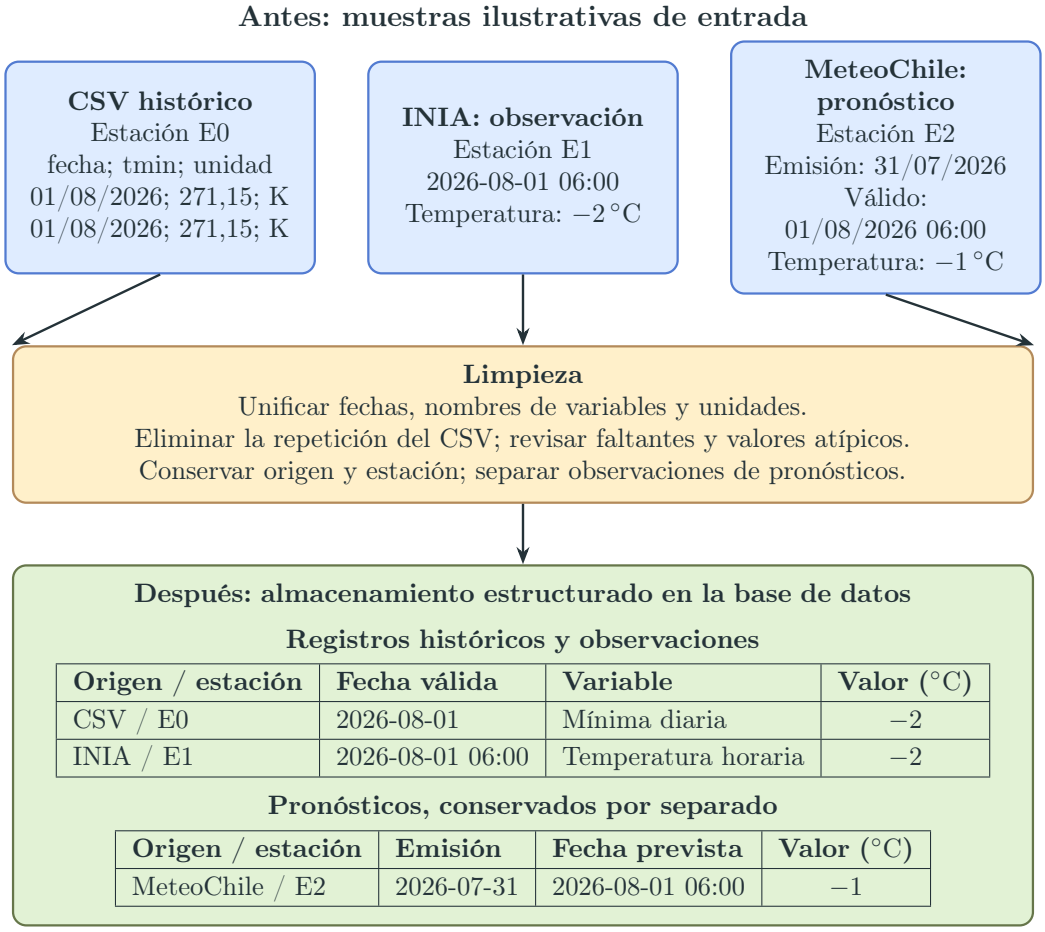


Figura 5.3: Ejemplo ilustrativo de entradas CSV, INIA y MeteoChile, su limpieza y almacenamiento.

La Figura 5.3 muestra cómo una repetición del CSV se elimina y cómo 271,15 K se transforma en −2 °C. Los registros normalizados conservan la fuente, la estación y la referencia temporal, distinguiendo la mínima diaria de una temperatura horaria. El pronóstico se guarda por separado, con su fecha de emisión y el momento previsto. Las muestras y los campos de almacenamiento son ilustrativos: no son respuestas reales de los servicios ni un esquema físico de base de datos verificado.

Para reunir fuentes, cada fila necesita al menos procedencia, tipo de dato, ubicación, instante válido, variable, valor y unidad.

Fuente	**Tipo**	**Referencia conservada**	**Uso**
INIA	Observación	Estación, fecha, variable, valor y unidad	Serie observada
Atlas de Riesgo Climático	Registro histórico	Sector, período, variable, valor y unidad	Antecedente histórico
MeteoChile	Pronóstico	Estación, fecha prevista, variable, valor y unidad	Referencia externa

Tabla 5.2: Esquema ilustrativo de registros normalizados por fuente y tipo.

La Tabla 5.2 muestra campos comunes para integrar registros sin borrar su procedencia. Las series históricas pueden consolidarse después de homologar variables y referencias; el pronóstico conserva su fecha prevista y se contrasta por separado, sin incorporarlo como observación de entrenamiento. No son filas extraídas del conjunto de datos.

Finalmente, la preparación para el modelamiento debe evitar que información futura intervenga en la construcción de entradas del pasado. La evaluación de Pudahuel reportada más adelante comprende 490 casos; no se le atribuye una proporción de partición ni fechas de corte que la documentación disponible no haya vinculado con ese experimento.

5.6. Construcción del conjunto de datos para heladas

El experimento documentado plantea dos preguntas complementarias: si ocurrirá un evento al día siguiente y cómo evolucionará la temperatura. Como ejemplo de evaluación preliminar se reportan resultados con datos de Pudahuel, definiendo la clase positiva mediante una temperatura inferior a 3 °C. Ese umbral operativo vale para la tabla publicada; no reemplaza la definición meteorológica de helada ni expresa una temperatura crítica universal para todos los cultivos. El corpus consolidado de más de 750.000 registros es la base amplia del trabajo; Pudahuel ilustra cómo se midió el desempeño, no el único origen de los datos.

Cada ejemplo de clasificación relaciona la información disponible al emitir la predicción con la clase observada posteriormente. En regresión, el objetivo es la temperatura, con un horizonte descrito de 24 horas e intervalos horarios. La observación posterior permite evaluar el resultado, pero no debe incorporarse como entrada disponible al momento del pronóstico.

La tabla de resultados de clasificación permite reconstruir un conjunto de prueba de

490 casos: 86 positivos y 404 negativos, sumando los conteos observados de cualquiera de sus modelos. Ese subconjunto es el recorte reportado para el ejemplo de Pudahuel y no debe confundirse con los más de 750.000 registros del corpus.

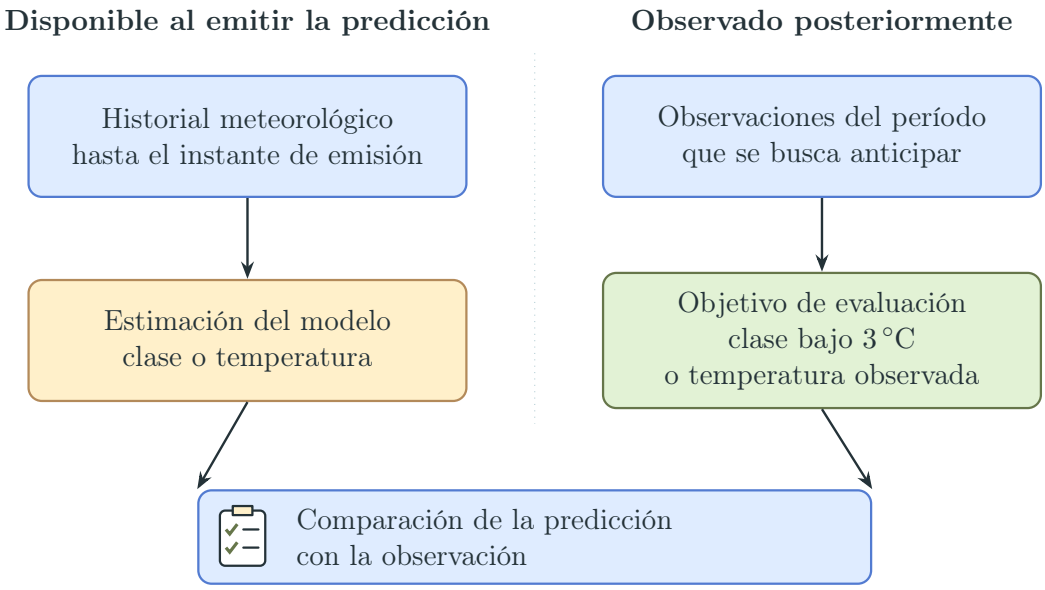


Figura 5.4: Separación conceptual entre entradas disponibles y objetivos observados.

La Figura 5.4 separa la información que puede utilizar el modelo de aquella que permite comprobar su resultado. La clase positiva depende del umbral del experimento y la regresión conserva la temperatura como variable continua. El esquema no asigna ventanas, particiones o fechas que no estén documentadas ni demuestra por sí mismo que se haya evitado toda fuga de información.

5.7. Predicción e integración con el sistema

Contexto del producto

El producto desarrollado se centra en dos ejes principales: (i) la construcción y evaluación de un modelo predictivo inicial de heladas entrenado con datos históricos, y (ii) la implementación en Miido, que permite a los agricultores seleccionar los campos agrícolas que desean monitorear. De esta forma, se establece la base técnica del sistema, aunque aún se encuentra pendiente la validación en condiciones reales de campo.

En el sistema descrito para el caso, el pronóstico WRF-DMC de MeteoChile se integra con la salida predictiva propia para apoyar la decisión de riesgo. La coordinación de las dos consultas y la decisión posterior corresponden al funcionamiento informado por el responsable del proyecto; el código entregado permite revisar cada servicio, pero no

conserva la configuración ni una respuesta conjunta de esa integración. Por ello no se atribuyen a la interfaz Miido resultados numéricos de concordancia, exactitud frente a mediciones o alertas entregadas. La interfaz territorial en producción sí cuenta con evidencia independiente.

Consultas y resultados para evaluar heladas

El sistema recibe la ubicación de un campo. Por un lado, consulta las condiciones previstas por MeteoChile para una estación cercana; por otro, obtiene una estimación propia asociada a la comuna. La primera respuesta permite conocer las temperaturas y los indicadores de helada por hora y por día; la segunda entrega una temperatura mínima estimada y una señal sobre la ocurrencia del evento. Según el funcionamiento informado para el caso, ambas se consideran antes de decidir qué riesgo comunicar. Ninguna respuesta aislada equivale a una alerta entregada.

La predicción propia del código examinado parte de la última fecha disponible en sus registros históricos, que terminan en diciembre de 2016. Por eso, esos archivos no permiten afirmar que la respuesta corresponda al día siguiente de una consulta actual. Esta versión de la consulta tampoco equivale al ensamble de nueve modelos cuya evaluación se presenta más adelante.

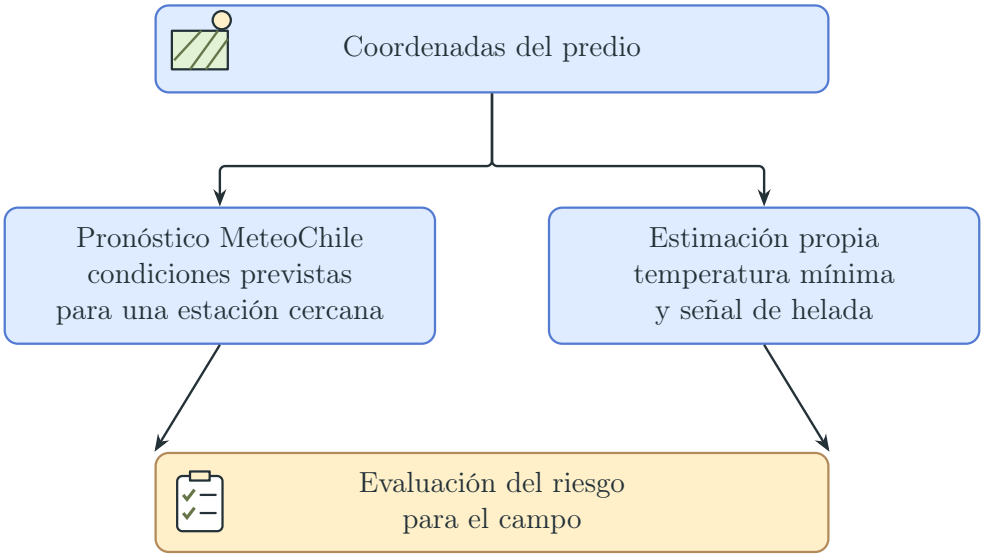


Figura 5.5: Información utilizada para evaluar el riesgo de helada de un campo.

La Figura 5.5 distingue el pronóstico externo de la estimación propia. Ambos aportan información a la evaluación del riesgo; el esquema no supone que una estación represente exactamente las condiciones de todo el campo ni que la coincidencia entre dos pronósticos demuestre su acierto.

Información consultada	**Qué aporta a la evaluación**
Ubicación del campo	Permite identificar la comuna y seleccionar una estación de referencia cercana cuando se dispone de sus datos.
Pronóstico MeteoChile	Entrega temperaturas y humedad previstas, indicadores horarios de helada y un resumen de las horas de riesgo por día. Si la consulta falla, se informa que no se obtuvieron los datos.
Estimación propia	Indica la fecha de los datos utilizados, la temperatura mínima estimada y una señal de ocurrencia del evento para la comuna. Cuando faltan datos, algunas estimaciones pueden no estar disponibles.

Tabla 5.3: Información que aportan las consultas a la evaluación de heladas.

La Tabla 5.3 muestra qué función cumple cada respuesta, sin exigir al lector conocer su formato técnico. La referencia territorial y la fecha de los datos son necesarias para interpretar la información: un pronóstico asociado a una estación no equivale a una medición en cada sector del campo, y una estimación obtenida con registros antiguos no acredita condiciones actuales.

Esquema general

El producto se compone de los siguientes elementos:

- M1 y M2: consolidación y limpieza de registros históricos provenientes de INIA, Atlas de Riesgo Climático y MeteoChile.
- M3: modelos evaluados para clasificar heladas, incluidos dos regresores de temperatura mínima convertidos en votos; el informe también describe regresión horaria a 24 horas, sin publicar resultados de error térmico para esa tarea.
- M5: primera versión de interfaz en Miido, enfocada en la selección de campos y configuración básica de parámetros.
- M4: canal WhatsApp integrado para la comunicación; el comportamiento de cada requisito conversacional y la entrega de mensajes requieren verificaciones propias.

Descripción de componentes implementados

Salida predictiva y evaluación

El modelo permite distinguir si se anticipa una helada y estimar la temperatura. La primera salida es una clase que puede apoyar una decisión de riesgo; la segunda conserva un valor continuo para interpretar su intensidad. Ninguna de ellas equivale, por sí sola, a una alerta entregada al agricultor.

El modelamiento del caso utiliza Python 3.x y bibliotecas de análisis y aprendizaje automático como pandas, numpy, scikit-learn y XGBoost (*eXtreme Gradient Boosting*). Su elección permite preparar datos y evaluar clasificadores; no establece por sí misma el desempeño predictivo.

La combinación de clasificadores busca aprovechar comportamientos complementarios frente a los datos. Los comités de votación cuentan con antecedentes en predicción meteorológica de heladas [4], pero su uso no garantiza por sí solo una mejora. En este caso, la comparación con modelos individuales debe establecerse mediante las métricas de evaluación reportadas.

Para evaluar la predicción de heladas, la clase positiva indica que ocurrió el evento y la negativa que no ocurrió.

	Helada observada	**Sin helada observada**
Predicción positiva	Verdadero positivo (TP)	Falso positivo (FP)
Predicción negativa	Falso negativo (FN)	Verdadero negativo (TN)

Tabla 5.4: Matriz de confusión general para la clasificación de heladas.

La Tabla 5.4 sitúa los aciertos en la diagonal y distingue las falsas alarmas (FP) de los eventos omitidos (FN). Es una guía de lectura, no un resultado experimental.

Con lo anterior se definen las siguientes métricas:

$$CSI = \frac{TP}{TP + FP + FN} \qquad Recall = \frac{TP}{TP + FN} \qquad Precisión = \frac{TP}{TP + FP}$$

El ensamble descrito utiliza votación mayoritaria (*Hard Voting*). Se entrenan clasificadores sobre el conjunto de entrenamiento y cada uno emite una clase para una nueva entrada. La decisión final corresponde a la clase con más votos. La utilidad de esa combinación depende de los errores de sus componentes y se evalúa frente a los modelos individuales; no se presupone que combinar más modelos produzca siempre mejores resultados.

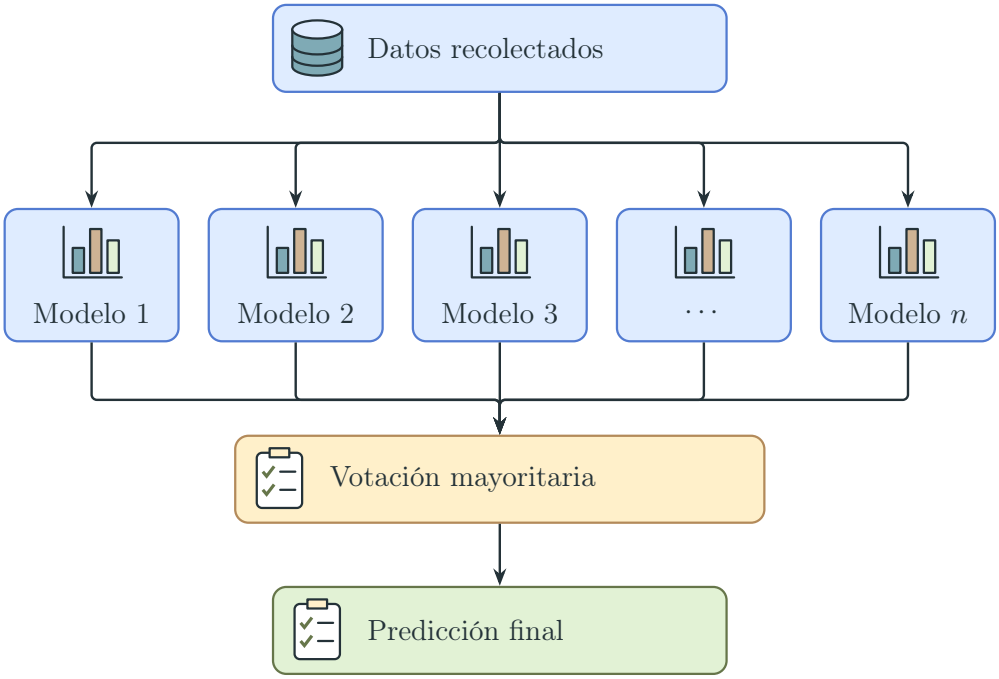


Figura 5.6: Esquema de votación mayoritaria con n clasificadores.

La Figura 5.6 muestra que los clasificadores reciben la misma entrada y aportan sus decisiones a una votación mayoritaria. Las conexiones convergen en la decisión del ensamble; no representan un entrenamiento secuencial de los modelos. La salida es una clase, no una demostración de su calidad predictiva.

La clasificación reúne siete modelos que votan directamente y dos que convierten su temperatura mínima estimada en un voto de helada. La estimación de temperatura hora a hora es una tarea distinta y no se evalúa mediante los conteos de clasificación.

Interacción ML-sistema

Según la documentación del caso, un agente consultaría periódicamente las salidas del modelo —clase del ensamble y, cuando exista, un score asociado— y registraría los resultados en la base de datos, asociándolos a los campos seleccionados por el usuario en Miido. Ese flujo está descrito para alimentar paneles de configuración y estado; no se audita aquí de extremo a extremo salvo la UI territorial en producción. El ensamble documentado decide por votación mayoritaria de clases; no se presenta una probabilidad calibrada lista para disparar alertas.

5.7.1. Integración con Miido

Miido es la plataforma agtech del caso: aporta la capa de configuración territorial. La interfaz en producción permite registrar y seleccionar campos, organizar secciones geográficas y asociarlas a grupos. Ese contexto permite relacionar las salidas del modelo con los sectores que interesan a cada usuario, en vez de presentar una predicción sin referencia al predio. Asociar un campo en Miido no equivale a haber validado el modelo a escala de polígono: la evaluación cuantificada corresponde al ejemplo de Pudahuel.

La interfaz utiliza React para coordinar los componentes visibles al seleccionar campos, grupos y parámetros; Vite forma parte de su entorno de desarrollo [45]. La configuración territorial y la comunicación por WhatsApp mantienen responsabilidades separadas.

La coordinación informada con WRF-DMC es una etapa adicional a esa asociación territorial; el esquema siguiente representa solo el vínculo entre modelo, backend y campos, no la lógica de decisión basada en las dos consultas.

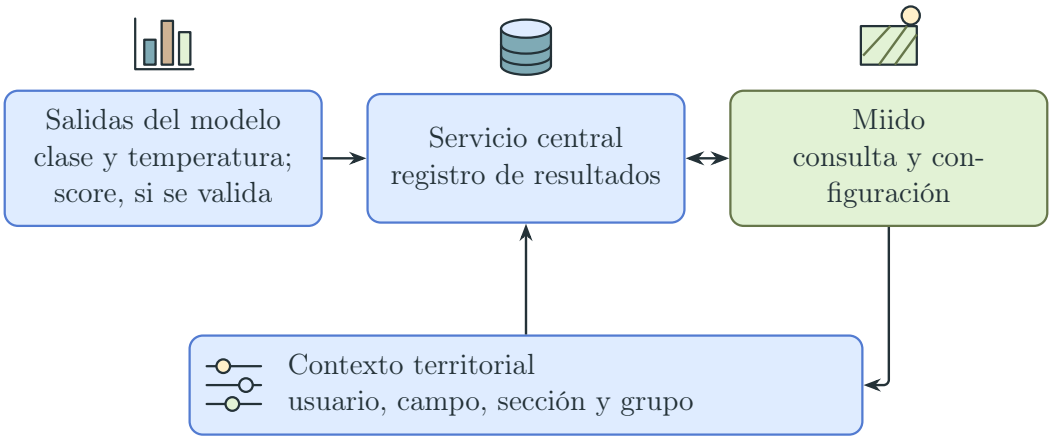


Figura 5.7: Relación lógica entre resultados del modelo y configuración territorial en Miido.

La Figura 5.7 distingue la estimación meteorológica de la información que permite situarla para un usuario. La asociación con campos y grupos se realiza en la integración, no mediante el cambio de resolución del modelo. El diagrama sintetiza las responsabilidades descritas y no representa una prueba de entrega o actualización en producción.

Interfaz territorial en producción

La interfaz de Miido permite registrar y visualizar campos en un mapa, seleccionar cuáles se monitorean y configurar parámetros básicos. La conexión del canal WhatsApp no se deduce del funcionamiento de esta interfaz.

La versión más reciente permite crear grupos de secciones geográficas, asociar secciones a grupos y gestionar configuraciones diferenciadas por grupo.



Figura 5.8: Interfaz territorial de Miido: gestión de grupos de secciones geográficas y configuración por grupo.

La Figura 5.8 permite identificar el mapa y los grupos asociados a las secciones. La captura no demuestra por sí sola la entrega de alertas ni una predicción validada por polígono.

5.7.2. Integración con WhatsApp

WhatsApp se eligió como canal de comunicación para reducir la necesidad de aprender otra interfaz al recibir una consulta o alerta. Además del acceso móvil descrito en la introducción, la encuesta de la Subsecretaría de Telecomunicaciones y CADEM de 2023 registra que el 80,1 % de quienes respondieron la pregunta sobre redes sociales declaró usar WhatsApp habitualmente [12]. Es una medida de uso en la población encuestada, no entre agricultores ni usuarios de este sistema. En el ámbito agrícola, [9] describe la coordinación de labores mediante mensajería y su posible uso en extensión rural. La Figura 3.6 distingue la alerta dirigida al productor de la publicación de avisos institucionales: el canal elegido permite llevar información asociada a un predio y a un momento de decisión sin exigir que el agricultor consulte varios portales.

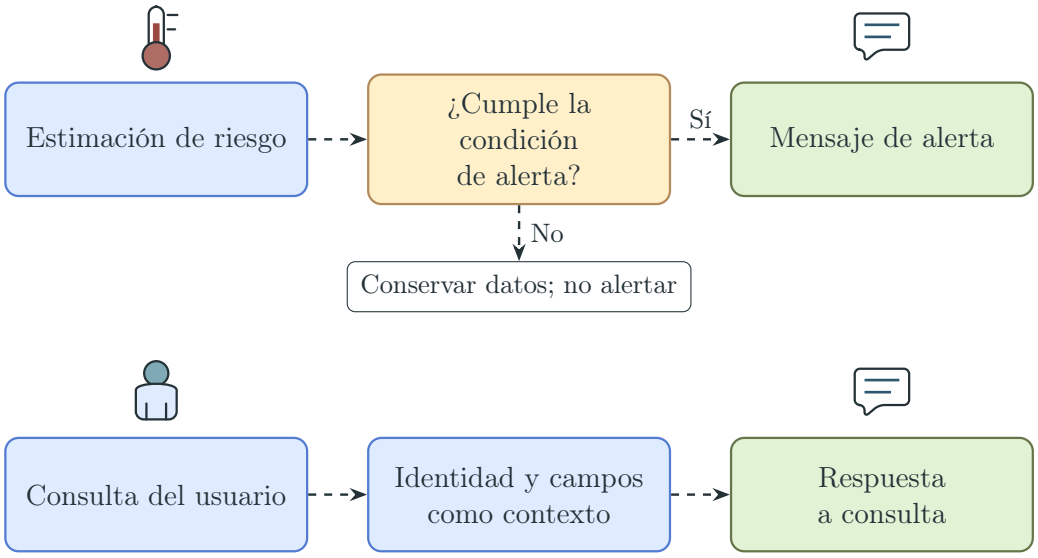
La integración mediante la interfaz de programación de aplicaciones (API) de Me-

ta Business requiere identificar destinatarios, respetar cuotas y plantillas y confirmar entregas. El entorno también contempla herramientas de OpenAI para agentes de lenguaje; esa posibilidad no demuestra que sus servicios estén desplegados ni intervengan en cada mensaje.

WhatsApp está integrado como canal de mensajería del sistema; por él se organizan el registro del campo (RF1), las consultas (RF2) y las alertas (RF6), mientras Miido administra los campos y sus configuraciones. No se presentan aquí registros que permitan cuantificar la entrega o evaluar la operación de cada modo con agricultores.

- Modo proactivo: el agente prepara una alerta cuando el criterio configurado indica riesgo (RF6).
- Modo reactivo: el agricultor consulta el pronóstico o riesgo de un campo y recibe una respuesta contextualizada (RF2).

La producción de una salida por el modelo y la entrega de un mensaje son etapas distintas; el esquema siguiente representa ambos modos sin cuantificar las entregas realizadas.



Conexiones segmentadas: rutas de consulta y alerta; el esquema no representa métricas de entrega.

Figura 5.9: Separación de los modos proactivo y reactivo de interacción por WhatsApp.

La Figura 5.9 reserva la evaluación del umbral para la ruta de alertas y separa la respuesta a consultas. La rama negativa conserva información sin enviar un aviso. Verificar destinatarios, confirmaciones de entrega y utilidad para agricultores es distinto de integrar el canal y de evaluar el modelo.

5.7.3. Prototipos de interacción del caso de estudio

Los prototipos se conservan como ejemplos del caso concreto. WhatsApp es el canal de consulta y alerta. Miido es la plataforma agtech: mapa, campos, grupos y configuración. Sus valores ilustrativos no constituyen mediciones ni resultados de validación. Los archivos de imagen originales se conservan; las adaptaciones de texto se realizan en LaTeX.

Prototipo 1 – Interfaz Conversacional vía WhatsApp

La conversación propone comunicar ubicación, momento y condiciones previstas sin exigir la consulta directa de los servicios meteorológicos.



Figura 5.10: Prototipo 1 – Interfaz conversacional vía WhatsApp; texto del control de entrada adaptado al español.

La Figura 5.10 contrapone una alerta asociada a un campo y una ventana temporal con una respuesta a consulta. El 70 % y los horizontes de tres días o una semana son ejemplos de presentación, no resultados que acrediten calibración probabilística o alcance validado del modelo.

Evento	Interacción	Acción	Objeto afectado
Consulta de pronóstico	El usuario pregunta por las temperaturas de la semana.	Respuesta con el pronóstico en texto estructurado.	Chat de WhatsApp del usuario
Alerta de helada	El criterio de alerta configurado se cumple (voto o umbral).	Mensaje ilustrativo con imagen, fecha, hora y temperaturas estimadas.	Chat de WhatsApp del usuario

Tabla 5.5: Eventos e interacciones del prototipo 1: interfaz conversacional vía WhatsApp

La Tabla 5.5 diferencia los disparadores: una petición del agricultor inicia la consulta y la evaluación de un umbral inicia la alerta. Compartir el chat no hace equivalentes ambas operaciones.

Prototipo 2 – Interfaz de Alertas en la Plataforma Miido

El panel permite seleccionar cultivos y ajustar umbrales de temperatura y humedad. A diferencia de la conversación, ofrece una vista persistente de la configuración activa.



Figura 5.11: Prototipo 2 – Interfaz de alertas en Miido; controles adaptados al español.

La Figura 5.11 distingue la selección de valores predefinidos de la modificación manual mediante controles. El resumen lateral permite revisar qué parámetros quedarían activos. Las temperaturas y humedades mostradas son ilustrativas: no establecen umbrales agronómicos validados para las especies listadas ni deben confundirse con una probabilidad de ocurrencia.

Evento	Interacción	Acción	Objeto afectado
Selección de cultivo	Selección de un cultivo de la lista.	Visualización de sus parámetros.	Panel de configuración actual
Ajuste de umbrales	Modificación de los controles de temperatura.	Actualización de límites para el monitoreo.	Configuración de alertas
Activación de alertas	Activación manual para un cultivo.	Evaluación de datos según los umbrales definidos.	Estado de alertas y backend

Tabla 5.6: Eventos e interacciones del prototipo 2: configuración y estado de alertas en Miido

La Tabla 5.6 separa elegir un cultivo, modificar sus límites y activar las alertas. Esa distinción evita interpretar un cambio visual como confirmación de monitoreo operativo: la interfaz expresa una configuración solicitada cuya aplicación depende del backend y de la disponibilidad de datos.

Prototipo 3 – Visualización Geográfica de Campos

El mapa organiza la información mediante polígonos asociados a cultivos y datos térmicos. El panel lateral permite seleccionar un campo y acceder a acciones de registro o modificación.

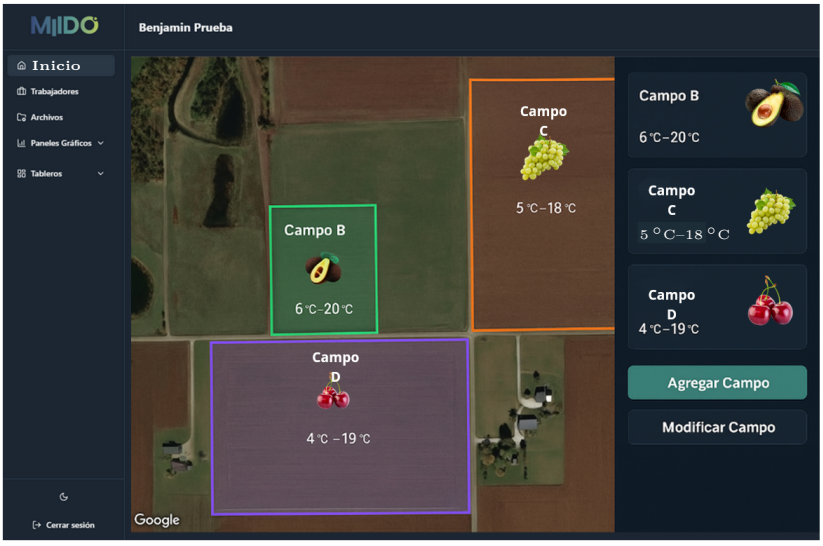


Figura 5.12: Prototipo 3 – Visualización geográfica de campos; navegación adaptada al español y rango ilustrativo de Campo C unificado.

La Figura 5.12 permite relacionar los polígonos y las tarjetas laterales mediante los nombres de los campos y los iconos de cultivos. Los colores diferencian los polígonos,

pero no se reproducen como identificadores en las tarjetas. Esta organización visual no aumenta la resolución del pronóstico ni demuestra que exista una estimación validada por polígono.

Acción del usuario	**Respuesta de la interfaz**	**Elemento afectado**
Seleccionar un polígono	Resaltar el campo y actualizar su información lateral.	Mapa y panel lateral
Agregar un campo	Abrir el formulario o la herramienta de delimitación.	Mapa y backend de campos
Modificar un campo	Habilitar la edición de nombre, cultivo y ubicación.	Campo seleccionado

Tabla 5.7: Eventos e interacciones del prototipo 3: mapa interactivo de campos en la plataforma Miido

La Tabla 5.7 distingue seleccionar un campo de crear o modificar su registro. La selección cambia el foco de visualización; las otras acciones actualizan el contexto territorial que utilizará el sistema.

5.8. Resultados preliminares

5.8.1. Verificación y Validación

Verificación

La verificación contrasta el avance documentado de los componentes implementados con los requisitos de diseño. No afirma que todos los RF estén cumplidos. En esta fase se evaluó principalmente:

- Consistencia con el diseño: la documentación del caso describe trabajo en M1–M3 (recolección, limpieza y modelo predictivo) y la interfaz Miido, en línea con el diseño. La lectura de los prototipos de extracción no equivale a una auditoría integral ni acredita un pipeline operativo de extremo a extremo.
- Integración de datos: se documenta la consolidación y estandarización de registros históricos. Sin los registros de transformación no se puede certificar ausencia de errores ni cuantificar posibles pérdidas de información.
- Funcionamiento del modelo: se reportan resultados de clasificación binaria. La calibración de probabilidades y su adecuación para disparar alertas no se deducen únicamente de las métricas de clase ni del uso de un ensamble.
- Interfaz: la plataforma Miido en producción permite seleccionar campos y grupos, lo cual verifica la capa territorial del caso.

En esta etapa se verifica avance en M1–M3 y en la interfaz Miido. La integración

de WhatsApp está realizada según el estado informado del proyecto, pero verificar el cumplimiento de RF1, RF2 y RF6 exige pruebas funcionales de cada interacción.

Validación

La validación busca determinar si el sistema satisface las necesidades planteadas por los usuarios y los objetivos generales del proyecto. Se divide en dos niveles: teórico y práctico.

Validación Teórica La validación teórica se realizó mediante comparación de métricas y revisión conceptual:

- Métricas de desempeño: se comparan varias métricas por modelo. Ninguna de ellas se adopta como umbral de aceptación ni como resumen suficiente del prototipo; el desempeño sigue siendo un aspecto de evaluación y mejora.
- Coherencia metodológica: los comités de votación cuentan con antecedentes en predicción de heladas [4]; su desempeño en este caso se examina mediante comparación con los clasificadores individuales, no por transferencia automática de resultados bibliográficos.
- Comparación con el estado del arte: el enfoque (aprendizaje automático / ensambles para heladas) comparte el tipo de problema con trabajos recientes en terrenos complejos [5]; no se afirma equivalencia de métricas ni transferencia geográfica.

Para la validación de los modelos se planteó el siguiente problema: con datos históricos hasta hoy, ¿ocurrirá mañana un evento bajo el umbral operativo? y ¿cómo estará la temperatura a través del día?. Como ejemplo reportado se utilizaron datos de la comuna de Pudahuel, definiendo el evento cuando la temperatura es menor a $3°C$. El corpus de más de 750.000 registros es la base amplia del trabajo; esta tabla ilustra el desempeño en ese ejemplo.

Tabla 5.8: Resultados de clasificación para Pudahuel; umbrales de varios clasificadores ajustados maximizando CSI en entrenamiento.

Modelo	CSI	Sensibilidad	Precisión	F1	TP	TN	FN	FP
ENSAMBLE_9	0.5873	0.8605	0.6491	0.7400	74	364	12	40
SVM	0.5826	0.7791	0.6979	0.7363	67	375	19	29
CatBoost	0.5536	0.7209	0.7045	0.7126	62	378	24	26
XGBoost	0.5455	0.8372	0.6102	0.7059	72	358	14	46
RandomForest	0.5364	0.6860	0.7108	0.6982	59	380	27	24
LightGBM	0.5086	0.6860	0.6629	0.6743	59	374	27	30
MLP	0.4724	0.6977	0.5941	0.6417	60	363	26	41
LSTM_Clf	0.4379	0.8605	0.4713	0.6091	74	321	12	83
LSTM	0.3270	1.0000	0.3270	0.4928	86	227	0	177
GRU	0.3094	1.0000	0.3094	0.4725	86	212	0	192

En varios clasificadores, el umbral de decisión se eligió maximizando el CSI sobre el conjunto de entrenamiento antes de evaluar en prueba. La Tabla 5.8 presenta esos resultados para el ejemplo de Pudahuel, sin extrapolarlos a una validación operativa.

Las siglas de la tabla corresponden a máquina de vectores de soporte (SVM, del inglés *Support Vector Machine*), perceptrón multicapa (MLP, del inglés *Multilayer Perceptron*), memoria a largo y corto plazo (LSTM, del inglés *Long Short-Term Memory*) y unidad recurrente con compuertas (GRU, del inglés *Gated Recurrent Unit*). Sus arquitecturas se explican en el Anexo.

Las métricas de prueba están condicionadas a umbrales elegidos en entrenamiento y no sustituyen una validación operativa. Las filas LSTM y GRU corresponden a regresores convertidos en votos mediante una regla basada en 1,96 veces la raíz del error cuadrático medio (RMSE, del inglés *Root Mean Squared Error*); LSTM_Clf es un clasificador. Los regresores detectan los 86 eventos, pero generan 177 y 192 falsas alarmas. El ensamble obtiene el mayor CSI y F1, no la mayor precisión ni la mayor sensibilidad. Elegir un modelo depende del costo de la falsa alarma, no de una sola métrica.



Figura 5.13: Matriz de confusión del ensamble de nueve modelos, reconstruida a partir de los conteos reportados.

La Figura 5.13 presenta los conteos reportados para el ensamble: detectó 74 de los 86 casos positivos, mientras 40 de sus 114 predicciones positivas correspondieron a falsas alarmas. Esto explica la diferencia entre sensibilidad y precisión: una detección elevada no elimina el costo de avisos incorrectos. La visualización reproduce los 490 casos de la evaluación reportada y no agrega experimentos nuevos.

Validación Práctica Las comprobaciones prácticas documentadas corresponden a:

- Simulación con datos históricos: la documentación menciona evaluación sobre series del periodo 2020–2025; la evidencia cuantificada publicada en este capítulo es el ejemplo de Pudahuel (490 casos de la Tabla 5.8).
- Pruebas en la plataforma Miido: se validó que la interfaz permita la selección de campos, lo que constituye un primer paso hacia la personalización de parámetros por agricultor.

La evaluación con datos de API en tiempo real no está documentada. Las pruebas piloto con agricultores no formaron parte del alcance del proyecto y se proponen para una etapa posterior.

Plan de testing

Para una etapa posterior, se propone el siguiente plan de pruebas funcionales y no funcionales:

1. Pruebas unitarias: verificar individualmente los módulos M1–M3 (recolección de datos, limpieza y modelo predictivo).
2. Pruebas de integración: comprobar el flujo completo de información desde los datos de entrada hasta la salida del modelo en la interfaz Miido.

3. Pruebas funcionales: evaluar que el sistema responda a consultas del usuario (modo reactivo) y que genere alertas automáticas (modo proactivo, aún pendiente).
4. Pruebas de rendimiento: medir tiempos de procesamiento y capacidad de respuesta del modelo frente a grandes volúmenes de datos.
5. Pruebas de validación de usuario: ejecutar pilotos con agricultores para evaluar la usabilidad, claridad de las alertas y utilidad de las recomendaciones.
6. Pruebas de robustez: simular fallas en las APIs externas y verificar la capacidad del sistema de mantener una respuesta estable.

En conjunto, este plan orienta las verificaciones pendientes hacia una solución confiable: no garantiza por sí solo el cumplimiento de los requisitos ni la adopción por el usuario final.

5.9. Decisiones de arquitectura

La documentación permite identificar las siguientes decisiones y sus consecuencias para el caso de estudio:

- **Conectores por fuente:** separar el acceso externo del procesamiento limita el alcance de un cambio de formato o navegación. La contrapartida es mantener y verificar cada conector por separado.
- **Persistencia y procedencia:** conservar registros originales y su origen permite revisar transformaciones. Esto requiere distinguir versiones brutas, preparadas y resultados, evitando sobrescrituras que impidan reconstruir el análisis.
- **Clasificación y regresión:** una salida responde si se cumple el criterio de evento y la otra estima la temperatura. Necesitan métricas y objetivos de evaluación distintos; ninguna debe presentarse automáticamente como probabilidad calibrada.
- **Contexto territorial en Miido:** la interfaz organiza los campos y parámetros del usuario, mientras el modelo produce estimaciones climáticas. Asociar ambos niveles no demuestra mayor resolución espacial ni validez agronómica por cultivo.
- **Mensajería desacoplada:** calcular una estimación, decidir si corresponde alertar y entregar un mensaje son responsabilidades distintas. La integración del canal no sustituye las pruebas de entrega ni la evaluación del modelo.
- **Evaluación progresiva:** los resultados históricos constituyen un primer nivel de evidencia. La generalización territorial, las fallas de servicios externos y la utilidad para agricultores requieren validaciones adicionales.

Estas decisiones explican la organización del trabajo descrito y sus dependencias. No se presentan versiones de bibliotecas, medidas de latencia, fechas exactas de partición o tasas de entrega que no estén respaldadas por el material disponible.

5.10. Limitaciones actuales

La revisión de los requisitos y del diseño distingue los componentes documentados de las capacidades que aún requieren comprobación. Las pruebas con agricultores quedaron fuera del alcance de este proyecto: su ausencia limita las conclusiones sobre utilidad en terreno, pero no debe presentarse como un hito incumplido.

5.10.1. Revisión de requerimientos

Los requisitos funcionales de monitoreo y detección (RF3 y RF4) cuentan con trabajo documentado sobre datos históricos, pero no con evidencia de monitoreo continuo de todas las fuentes. La estimación del riesgo (RF5) se evaluó de manera preliminar; personalización por cultivo, resolución territorial y probabilidades calibradas requieren verificación adicional. Para el registro, la consulta y las alertas (RF1, RF2 y RF6), la integración del canal no demuestra aún el funcionamiento de cada interacción ni las entregas.

La calidad predictiva (RNF6) sigue en evaluación. CSI, sensibilidad, precisión y F1 de la Tabla 5.8 responden a preguntas distintas; los porcentajes preliminares de planificación no se adoptan como umbral universal de aceptación. La evaluación adicional debe examinar la transferencia a otros predios y el error de temperatura de los modelos de regresión bajo un protocolo explícito. Tampoco se dispone aquí de medidas de disponibilidad, escalabilidad o seguridad que permitan dar por cumplidos los demás requisitos no funcionales.

5.10.2. Revisión del diseño

Se conservó la separación entre recolección, limpieza, predicción, interacción y validación (M1–M6). La preparación contempla formatos heterogéneos y procedencia por fuente; Miido aporta el contexto territorial y WhatsApp el canal de comunicación. Esta correspondencia entre diseño e implementación no implica una medición de latencia ni un despliegue integral de todos los módulos.

5.10.3. Análisis crítico

El desarrollo del prototipo permitió constatar avances relevantes, pero también dejó en evidencia limitaciones técnicas y prácticas que deben considerarse en las siguientes etapas:

- Fortalezas:
  - Se cuenta con un dataset consolidado y estructurado de amplio volumen, según la documentación del caso; la homogeneidad y la ausencia de errores no quedan certificadas sin los registros de transformación (véase verificación).
  - Se implementó un modelo predictivo preliminar cuya evaluación ilustrativa se reporta con varias métricas en la Tabla 5.8, sin adoptar un porcentaje único de precisión como resultado de aceptación.
  - La plataforma Miido opera en producción para mapa, campos y grupos; la Figura 5.8 ilustra su interfaz territorial.
- Debilidades:
  - El desempeño predictivo requiere evaluación adicional y mejora. Los ajustes de hiperparámetros, variables o modelos deben comprobarse con datos de evaluación apropiados, sin suponer que una mayor complejidad mejora necesariamente las predicciones.
  - No se presentan métricas de entrega de mensajes; la utilidad de las alertas en terreno todavía no puede determinarse con la evidencia disponible.
  - La dependencia de APIs externas puede afectar la disponibilidad y robustez del sistema.
- Oportunidades de mejora:
  - Explorar técnicas de optimización de modelos de ensemble (p. ej., *stacking* o *boosting*) para mejorar CSI, sensibilidad y F1 bajo un protocolo de umbral explícito.
  - Planificar pruebas piloto con agricultores que permitan retroalimentar directamente la usabilidad del sistema.

5.11. Gestión de riesgos

Se reúnen las condiciones necesarias para operar y los escenarios que pueden comprometer el funcionamiento del sistema.

5.11.1. Supuestos

Estos supuestos orientan el diseño; no se presentan como hechos verificados en este capítulo.

- Los agricultores cuentan con smartphones y acceso básico a WhatsApp.
- Las APIs de sensores remotos y meteorológicas permanecen disponibles con una frecuencia suficiente para el caso (su estabilidad no está demostrada aquí).
- Los usuarios pueden comprender y utilizar la información entregada. Su confianza y adopción deben evaluarse con usuarios, no presumirse por la existencia de la interfaz.

5.11.2. Dependencias

- Proveedores de datos meteorológicos (INIA, servicios comerciales).
- Disponibilidad y cuotas de Meta Business API (WhatsApp).
- Si se utiliza un agente de lenguaje basado en OpenAI, su operación depende de la disponibilidad del servicio y de sus límites de tasa de llamadas.
- Calidad y orden de los datos entregados por la empresa.

5.11.3. Restricciones

- Usuarios sin conocimientos técnicos: las interacciones deben ser claras y amigables.
- La cobertura de sensores es limitada: las zonas sin observaciones representativas restringen la extrapolación de resultados y requieren explicitar la incertidumbre.

5.11.4. Evaluación de riesgos

Los escenarios siguientes pueden comprometer el desempeño. Las mitigaciones se formulan como orientación; no se presentan como ya implementadas:

- Fallas en APIs externas: la interrupción de servicios de INIA o de predicciones meteorológicas podría retrasar alertas.
  Mitigación prevista: redundancia de fuentes y almacenamiento en caché.
- Errores de predicción del modelo: falsos positivos o falsos negativos ante la variabilidad climática.
  Mitigación prevista: ensambles, ajuste de umbrales y monitoreo continuo de métricas.
- Límites de cuota de APIs: exceso de llamadas puede interrumpir el servicio.
- Interpretación de lenguaje natural: posibles malentendidos en comandos; mitigación con confirmaciones.

Los límites de la evaluación predictiva se precisan en la Sección 5.10.1.

El Capítulo 6 retoma los objetivos a partir de estos resultados y delimita el trabajo futuro. Los prototipos y simulaciones no sustituyen la validación en terreno.

Capítulo 6

Conclusiones

Este proyecto abordó el desarrollo de una herramienta para obtener, preparar y utilizar datos climáticos en apoyo a la predicción de heladas. El trabajo reunió fuentes externas, procesos de extracción y estandarización, resultados preliminares de modelos predictivos y componentes de interacción con el agricultor. La integración con Miido y WhatsApp vinculó la gestión territorial con el canal de comunicación, como partes de una solución orientada al uso agrícola.

6.1. Cumplimiento de los objetivos

Los resultados del proyecto permiten relacionar el trabajo realizado con los objetivos específicos:

- Se identificaron fuentes climáticas, como INIA, MeteoChile y el Atlas de Riesgo Climático, y se describieron sus mecanismos de acceso y la información pertinente para la predicción de heladas.
- Se diseñó un flujo de obtención mediante consultas a servicios, descarga de archivos y programas de extracción, según la forma de publicación de cada fuente.
- Se organizaron las tareas de estandarización y limpieza para reunir los registros en una base utilizable. El volumen consolidado reportado supera los 750.000 registros; esta cifra corresponde al corpus de trabajo y no al conjunto de evaluación.
- Se examinaron los resultados de clasificación reportados para Pudahuel, mediante métricas de detección y falsas alarmas, y se describió la estimación horaria mediante regresión. Los 490 casos de la Tabla 5.8 corresponden al ejemplo de clasificación, no a una evaluación del corpus completo ni del error de temperatura por hora.

- Se integró la herramienta con la gestión de campos y grupos de Miido y con WhatsApp como canal de comunicación. La Figura 5.8 muestra la interfaz territorial que aporta el contexto de los predios.

La arquitectura distribuyó el trabajo entre conectores por fuente, preparación y modelamiento de datos e integración con las interfaces. La separación de los conectores permitió organizar la obtención según los formatos de cada proveedor; la conservación de registros y procedencia aporta una base para revisar las transformaciones; y la mensajería desacoplada distingue la generación de estimaciones de su comunicación. Estas decisiones sirvieron para abordar los objetivos de obtención, limpieza e integración como partes de un mismo flujo. El trabajo realizado permitió desarrollar y documentar esos componentes; la comprobación de su operación conjunta y de su utilidad en terreno requiere una evaluación posterior.

6.2. Limitaciones y trabajo futuro

La integración de las interfaces y del canal de comunicación no equivale a verificar cada registro, consulta, alerta y entrega de mensajes. Esa comprobación funcional debe distinguirse de la evaluación predictiva. Las pruebas piloto con agricultores no formaron parte del alcance del proyecto y se proponen para una etapa posterior, en la que pueda evaluarse la claridad y utilidad de la información entregada.

En predicción, detectar más eventos también puede implicar generar más falsas alarmas. La Figura 5.13 muestra este equilibrio para el ensamble: detectó 74 de los 86 eventos reportados, pero produjo 40 falsas alarmas. Estos resultados permiten interpretar conjuntamente sensibilidad y precisión; no constituyen una demostración de utilidad agronómica ni una medición del error de temperatura de los modelos de regresión.

Una línea de continuidad consiste en evaluar la aplicación de esta metodología a otras condiciones de interés agrícola, como olas de calor, episodios de humedad extrema y otros eventos climáticos. Los mecanismos de obtención, preparación e integración de datos ofrecen una base reutilizable, pero cada aplicación requiere definir el evento y su duración, seleccionar las variables pertinentes y establecer el horizonte de predicción. Los mismos registros podrían aprovecharse únicamente si contienen las variables, cobertura espacial y temporal, continuidad y resolución necesarias para identificar y evaluar ese evento; cuando sean insuficientes, se necesitarían datos adicionales. La adaptación debe considerar los umbrales relevantes para los cultivos y comprobar el desempeño de los modelos; reutilizar el flujo de trabajo no implica que el modelo de heladas pueda emplearse sin modificaciones.

En síntesis, este proyecto aporta una base técnica que relaciona datos climáticos, modelamiento y comunicación agrícola. Su continuidad requiere comprobar el funcionamiento del sistema y evaluar la adaptación de la metodología a otras amenazas.

6.3. Comentarios del autor

Este proyecto aborda una oportunidad relevante para la agricultura chilena. El país dispone de diversas fuentes de información climática cuyo valor no siempre se traduce en herramientas accesibles para quienes trabajan en terreno. Desde esta perspectiva, el aporte no consiste únicamente en obtener datos o generar predicciones, sino en presentar una síntesis comprensible y vinculada con las necesidades del agricultor.

Acercar estas capacidades a los productores puede contribuir a reducir la brecha tecnológica entre los espacios rurales y urbanos. Facilitar el acceso a información útil podría fortalecer las actividades agrícolas y favorecer el desarrollo de las zonas rurales. Sin embargo, estos beneficios deben entenderse como posibilidades que requieren evaluación, no como efectos ya demostrados por este proyecto.

La tecnología debe complementar el conocimiento y la experiencia del agricultor, manteniendo en sus manos la decisión final. Su utilidad depende tanto de la calidad de las predicciones como de la oportunidad y claridad de su comunicación. Por ello, avanzar hacia una aplicación efectiva exige considerar la conectividad, la alfabetización digital y la confianza del usuario, además de comprobar el funcionamiento técnico. La complejidad del sistema no debería convertirse en una nueva barrera de acceso.

Anexo

Este anexo explica los modelos considerados en el capítulo 5. Separa los clasificadores, los regresores de temperatura mínima que aportan un voto y los modelos que estiman la temperatura hora a hora. anexo a

Votos del ensamble de helada

La Figura 5.6 muestra cómo el ensamble obtiene una clase a partir de nueve votos: siete clasificadores y dos regresores cuya temperatura estimada se convierte en clase. La Tabla 5.8 compara esas decisiones con los eventos observados en Pudahuel. Sus diferencias en detección y falsas alarmas muestran por qué interesa evaluar cada voto y el resultado combinado.

Clasificadores

- **Bosque aleatorio (Random Forest):** combina árboles entrenados con distintas muestras y características; la clase más votada por esos árboles determina su predicción.
- **XGBoost (eXtreme Gradient Boosting):** agrega árboles de forma secuencial para corregir errores de los anteriores.
- **LightGBM (Light Gradient Boosting Machine):** también construye árboles secuencialmente, pero prioriza la hoja cuya división reduce más el error.
- **CatBoost (Categorical Boosting):** utiliza árboles secuenciales y un procedimiento de entrenamiento ordenado; puede tratar variables categóricas, aunque el caso descrito no incorpora categorías explícitas.
- **Máquina de vectores de soporte (SVM, del inglés *Support Vector Machine*):** busca una frontera que separe ambas clases y utiliza una transformación para distinguir relaciones no lineales.
- **Perceptrón multicapa (MLP, del inglés *Multilayer Perceptron*):** red de capas conectadas que aprende relaciones no lineales entre las variables meteorológicas y la clase de helada.

- **Memoria a largo y corto plazo (LSTM, del inglés *Long Short-Term Memory*):** red recurrente que procesa una secuencia de siete días y conserva información temporal mediante compuertas para emitir una clase.

En varios clasificadores, la salida se convierte en clase mediante un umbral elegido en entrenamiento para maximizar el índice de éxito crítico (CSI, del inglés *Critical Success Index*). Este índice considera aciertos de helada, eventos omitidos y falsas alarmas; no debe confundirse con la proporción de aciertos sobre todos los días. anexo b

Regresores que aportan votos

La LSTM de regresión y la unidad recurrente con compuertas (GRU, del inglés *Gated Recurrent Unit*) procesan siete días para estimar la temperatura mínima del día siguiente. La GRU utiliza menos compuertas que la LSTM para mantener información temporal. Ninguna de estas dos salidas es por sí misma una clase: se convierten en votos mediante una regla común.

Esa regla marca helada cuando la temperatura predicha menos 1,96 veces la raíz del error cuadrático medio (RMSE, del inglés *Root Mean Squared Error*) calculada en entrenamiento queda en o bajo 3 °C. Se denomina voto pesimista porque anticipa el evento incluso ante una temperatura estimada algo mayor que el umbral. En la Tabla 5.8, los votos LSTM y GRU detectan los 86 casos positivos, pero también producen 177 y 192 falsas alarmas, respectivamente.

Regresión horaria

De forma independiente, se describen tres modelos que reciben las últimas 24 horas de datos y estiman las 24 horas siguientes:

- La red neuronal convolucional unidimensional (CNN, del inglés *Convolutional Neural Network*) identifica patrones locales en la evolución de la temperatura mediante filtros.
- La LSTM procesa la secuencia hora a hora y conserva información de pasos anteriores.
- La GRU también procesa la secuencia temporal mediante compuertas, con una estructura más simple que la LSTM.

El esquema promedia las tres temperaturas horarias; la Tabla 5.8 compara votos de helada, no errores por hora.

Bibliografía

[1] Food and Agriculture Organization of the United Nations, "The impact of disasters and crises on agriculture and food security: 2021," tech. rep., FAO, 2021. Disponible en: https://www.fao.org/documents/card/en/c/cb3673en.

[2] IPCC, *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Cambridge University Press, 2022. DOI: https://doi.org/10.1017/9781009325844.

[3] R. L. Snyder and J. P. de Melo-Abreu, *Frost Protection: Fundamentals, Practice, and Economics. Volume 1*. No. 10 in Environment and Natural Resources Series, Food and Agriculture Organization of the United Nations, 2005. Disponible en: https://www.fao.org/4/y7223e/y7223e00.htm.

[4] V. A. de Almeida, J. A. Anochi, J. R. Rozante, and H. F. de Campos Velho, "Machine learning with voting committee for frost prediction," *Meteorology*, vol. 4, no. 1, 2025. Artículo 6. DOI: https://doi.org/10.3390/meteorology4010006.

[5] S. Bhakare, M. Matiu, A. Crespi, and D. Zardi, "Spatial downscaling of daily temperature minima using machine learning methods and application to frost forecasting in two alpine valleys," *Atmosphere*, vol. 16, no. 1, 2025. Artículo 38. DOI: https://doi.org/10.3390/atmos16010038.

[6] Instituto de Investigaciones Agropecuarias, "Red de agrometeorología inia." Disponible en: https://agrometeorologia.cl/, 2025.

[7] OECD, "Bridging the rural digital divide," Tech. Rep. 265, Organisation for Economic Co-operation and Development, 2018. DOI: https://doi.org/10.1787/852bd3b9-en.

[8] V. Grimblatt, "IoT for agribusiness: An overview," in *2020 IEEE Latin American Symposium on Circuits and Systems*, pp. 1–4, 2020. DOI: https://doi.org/10.1109/LASCAS45839.2020.9068986.

[9] O. Sotomayor, M. Namdar, H. Martínez, C. Saa, F. Barrera, M. Aedo, C. Kobrich, and I. Cano, "Evolución de las políticas públicas aplicadas en el área de

la agricultura digital en chile," in *Transición digital en agricultura y políticas públicas en América Latina*, pp. 159–201, E-papers Serviços Editoriais, 2024. DOI: https://doi.org/10.48207/9786587065878-5.

[10] C. A. Agostini and M. Willington, "Radiografía de la brecha digital en chile: ¿se justifica la intervención del estado?," *Estudios Públicos*, no. 119, pp. 5–32, 2010. DOI: https://doi.org/10.38178/cep.vi119.390.

[11] L. Castrillon, P. Villalobos, and L. Klerkx, "Análisis funcional-estructural del sistema digital de innovación agroalimentaria de la región del maule-chile," in *Transición digital en agricultura y políticas públicas en América Latina*, pp. 337–365, E-papers Serviços Editoriais, 2024. DOI: https://doi.org/10.48207/9786587065878-11.

[12] Subsecretaría de Telecomunicaciones and CADEM, "Presentación: Acceso y uso de internet 2023," tech. rep., Subsecretaría de Telecomunicaciones de Chile, 2023. Disponible en: documento oficial de SUBTEL. Consultado el 26 de agosto de 2026.

[13] Instituto de Investigaciones Agropecuarias, "Heladas. factores, tendencias y efectos en frutales y vides," Tech. Rep. Boletín INIA N.º 417, Instituto de Investigaciones Agropecuarias, 2020. Disponible en: https://hdl.handle.net/20.500.14001/6847.

[14] J.-F. Le Coq, F. Goulet, F. Bert, J. Van Loon, and D. Martínez-Baron, eds., *Transición digital en agricultura y políticas públicas en América Latina*. E-papers Serviços Editoriais, 2024. DOI: https://doi.org/10.48207/9786587065878.

[15] Instituto Nacional de Tecnología Agropecuaria, "Inta digital geo." Disponible en: https://geo.inta.gob.ar/, 2026. Consultado el 4 de agosto de 2026.

[16] Empresa Brasileira de Pesquisa Agropecuária, "Agroapi." Disponible en: https://www.portal.agroapi.cnptia.embrapa.br/, 2026. Consultado el 4 de agosto de 2026.

[17] Unidad de Planificación Rural Agropecuaria, "Observatorio de información rural y agropecuaria – agronet." Disponible en: https://www.agronet.gov.co/, 2026. Consultado el 4 de agosto de 2026.

[18] Ministerio de Ganadería, Agricultura y Pesca, "Snia – sistema nacional de información agropecuaria." Disponible en: https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/snia, 2026. Consultado el 4 de agosto de 2026.

[19] Dirección Meteorológica de Chile, "Portal de la dirección meteorológica de chile." Disponible en: https://www.meteochile.gob.cl/PortalDMC-web/index.xhtml, 2025.

[20] AFIPA, "Pérdidas del agro totalizan casi us$ 2.000 millones por eventos climáticos y desastres." Disponible en: https://www.afipa.cl/perdidas-del-agro-totalizan-casi-us-2-000-millones-por-eventos-climaticos-y-desastres/, 2024.

[21] R. Lam, A. Sanchez-Gonzalez, M. Willson, P. Wirnsberger, M. Fortunato, F. Alet, S. Ravuri, T. Ewalds, Z. Eaton-Rosen, W. Hu, A. Merose, S. Hoyer, G. Holland, O. Vinyals, J. Stott, A. Pritzel, S. Mohamed, and P. Battaglia, "Learning skillful medium-range global weather forecasting," *Science*, vol. 382, no. 6677, pp. 1416–1421, 2023. DOI: https://doi.org/10.1126/science.adi2336.

[22] S. Ravuri, K. Lenc, M. Willson, D. Kangin, R. Lam, P. Mirowski, M. Fitzsimons, M. Athanassiadou, S. Kashem, S. Madge, R. Prudden, A. Mandhane, A. Clark, A. Brock, K. Simonyan, R. Hadsell, N. Robinson, E. Clancy, A. Arribas, and S. Mohamed, "Skilful precipitation nowcasting using deep generative models of radar," *Nature*, vol. 597, no. 7878, pp. 672–677, 2021. DOI: https://doi.org/10.1038/s41586-021-03854-z.

[23] M. Fuentes, C. Campos, and S. García-Loyola, "Application of artificial neural networks to frost detection in central chile using the next day minimum air temperature forecast," *Chilean Journal of Agricultural Research*, vol. 78, no. 3, pp. 327–338, 2018. DOI: https://doi.org/10.4067/S0718-58392018000300327.

[24] Dirección Meteorológica de Chile, "Datos Modelo Regional Weather Research Forecast (WRF)." Disponible en: https://climatologia.meteochile.gob.cl/application/documentacion/getDocumento/24, 2026. Consultado el 4 de agosto de 2026.

[25] Instituto de Investigaciones Agropecuarias, "Heladas y Pronóstico." Disponible en: https://agrometeorologia.cl/heladas/, 2026. Consultado el 4 de agosto de 2026.

[26] National Weather Service, "Alerts web service." Disponible en: https://www.weather.gov/documentation/services-web-alerts, 2026. Consultado el 5 de agosto de 2026.

[27] World Meteorological Organization, "WMO and the early warnings for all initiative." Disponible en: https://wmo.int/activities/early-warnings-all/wmo-and-early-warnings-all-initiative, 2024. Consultado el 26 de agosto de 2026.

[28] M. Negishi and H. Kusaka, "Development of statistical and machine learning models to predict the occurrence of radiation fog in japan," *Meteorological Applications*, vol. 29, no. 2, p. e2048, 2022. DOI: https://doi.org/10.1002/met.2048.

[29] J. R. Rozante, E. R. Gutierrez, P. L. da Silva Dias, A. de Almeida Fernandes, D. S. Alvim, and V. M. Silva, "Development of an index for frost prediction: Technique

and validation," *Meteorological Applications*, vol. 27, no. 1, p. e1807, 2020. DOI: https://doi.org/10.1002/met.1807.

[30] C. J. Talsma, K. C. Solander, M. K. Mudunuru, B. Crawford, and M. R. Powell, "Frost prediction using machine learning and deep neural network models," *Frontiers in Artificial Intelligence*, vol. 5, p. 963781, 2023. DOI: https://doi.org/10.3389/frai.2022.963781.

[31] L. Alfieri, P. Salamon, F. Pappenberger, F. Wetterhall, and J. Thielen, "Operational early warning systems for water-related hazards in Europe," *Environmental Science & Policy*, vol. 21, pp. 35–49, 2012. DOI: https://doi.org/10.1016/j.envsci.2012.01.008.

[32] European Union/Copernicus, "European flood awareness system." Disponible en: https://www.copernicus.eu/en/european-flood-awareness-system, 2026. Consultado el 5 de agosto de 2026.

[33] National Weather Service, "Heatrisk overview." Disponible en: https://www.wpc.ncep.noaa.gov/heatrisk/overview.html, 2026. Consultado el 5 de agosto de 2026.

[34] S. W. Taylor and M. E. Alexander, "Science, technology, and human factors in fire danger rating: the Canadian experience," *International Journal of Wildland Fire*, vol. 15, no. 1, pp. 121–135, 2006. DOI: https://doi.org/10.1071/WF05021.

[35] Natural Resources Canada, "Canadian wildland fire information system." Disponible en: https://cwfis.cfs.nrcan.gc.ca/, 2026. Consultado el 5 de agosto de 2026.

[36] National Weather Service, "Glossary: Frost Advisory and Freeze Warning." Disponible en: https://forecast.weather.gov/glossary.php, 2026. Consultado el 5 de agosto de 2026.

[37] University of Florida Institute of Food and Agricultural Sciences, "FAWN text message freeze alert system." Disponible en: https://fawn.ifas.ufl.edu/tools/freeze_alert/start.php, 2020. Consultado el 5 de agosto de 2026.

[38] OASIS, "Common alerting protocol version 1.2." Disponible en: https://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2.html, 2010. Consultado el 5 de agosto de 2026.

[39] Dirección Meteorológica de Chile, "Imágenes Monitor de Heladas." Disponible en: https://climatologia.meteochile.gob.cl/application/diarioe/vistaMonitordeHeladas, 2026. Consultado el 4 de agosto de 2026.

[40] Dirección Meteorológica de Chile, "Sistema de Alerta Meteorológica." Disponible en: `https://archivos.meteochile.gob.cl/portaldmc/AAA/aaa_mapa.php`, 2026. Consultado el 2 de septiembre de 2026.

[41] Dirección Meteorológica de Chile, "Pronóstico de horas con temperatura a 2 m menor o igual a 0 °c para la región del maule." Disponible en: `https://climatologia.meteochile.gob.cl/application/diarioe/vistaMonitordeHeladas`, 2026. Imagen mostrada descargada desde el monitor el 26 de agosto de 2026; corresponde a pronóstico válido para el 26 de agosto de 2026. La página no se afirma como archivo de la imagen.

[42] National Center for Atmospheric Research, "Weather research and forecasting model (wrf)." Disponible en: `https://www.mmm.ucar.edu/models/wrf`. Consultado el 23 de septiembre de 2026.

[43] Node.js Contributors, "Overview of blocking vs non-blocking." Disponible en: `https://nodejs.org/en/learn/asynchronous-work/overview-of-blocking-vs-non-blocking`. Consultado el 23 de septiembre de 2026.

[44] Express.js Contributors, "Routing." Disponible en: `https://expressjs.com/en/guide/routing.html`. Consultado el 23 de septiembre de 2026.

[45] React Contributors, "Thinking in react." Disponible en: `https://react.dev/learn/thinking-in-react`. Consultado el 23 de septiembre de 2026.