


UKCP Chatbot Early Prototype - User Testing Survey

(Copy) 

* Required

Declaration

The final date for completion of this survey is 18/10/2024



Have you read and understood the testing consent form?

*

☐ I have read and understood the testing consent form.

Your experience in climate science and UKCP18

During part one, you will be asked to review a selection of answers to questions relating to UKCP18. These questions will have three different answers, each from a separate configuration of the chatbot.

How long have you been working in the field of climate science?

- ☐ Up to 2 years
- ☐ Up to 5 years
- ☐ Up to 10 years
- ☐ More than 10 years

What is your understanding of UKCP18 *

- ☐ None at all
- ☐ Very little understanding
- ☐ Some understanding
- ☐ Very well understood
- ☐ I would consider myself an expert

Part 1 - Answer Evaluation

The next section will present you with examples of user interactions with the chatbot. Each interaction has been produced by a different configuration of the UKCP Chatbot. Under each interaction you will be requested to grade the chatbot response provided.

Entry 1 - Chatbot response A

Question - I am currently writing a climate change impact assessment chapter for a solar development in the East Midlands. As such, I need to include cloud cover projections in my report. Please could you advise if this information is available, and if so where I can locate it?

Answer - The UKCP18 provides information on cloud cover projections for various regions, including the East Midlands. Cloud cover projections are important for solar development as they impact the amount of sunlight reaching solar panels. In the UKCP18, cloud cover projections can be accessed through the User Interface, which is a platform where various datasets, including probabilistic regionally-averaged data, are available. These datasets provide projected changes for different regions, including the East Midlands. Additionally, the Centre for Environmental Data Analysis (CEDA) Data Catalogue is mentioned as

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was authoritative (It sounded sure of itself).	<div></div>		

Entry 1 - Chatbot response B

Entry - I am currently writing a climate change impact assessment chapter for a solar development in the East Midlands. As such, I need to include cloud cover projections in my report. Please could you advise if this information is available, and if so where I can locate it?

Chatbot response - To find information on cloud cover projections for a solar development in the East Midlands, you can refer to the UKCP18 data. Specifically, for detailed projections related to cloud cover, you should utilize the UKCP Local data. The UKCP Local model is recommended for analyzing convection that generates thunderstorms, which implies it provides detailed information on various weather parameters, including cloud cover. Additionally, the UKCP Local data includes hourly wind data, rainfall data, and temperature data, making it a comprehensive source for analyzing weather patterns at a local scale. This level of detail is crucial for understanding the specific conditions that may impact solar development in the East Midlands. For more specific details on

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was authoritative (It sounded sure of itself).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please explain why you gave your scores for each of the five criteria *

Entry 2 - Chatbot response A

Entry - I work for United Utilities and as per Ofwat's guidance we are looking at some climate change scenario testing. To help us build a case for RCP 2.6 we are trying to pin point the time period on the graph on the document I cited where the temperature projections diverge between RCP 2.6 and RCP 8.5. So what year are RCP 2.6 and RCP 8.5 stop following the same path? I am trying to get hold of the data set behind the graph on the document you produced to help me with this work.

Chatbot response - The temperature projections for RCP 2.6 and RCP 8.5 stop following the same path according to the data set produced in the year 2040. Up to around 2040, both emissions scenarios show similar increases in the global mean surface air temperatures. However, after this point, the projections

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was why you gave your scores for each of the authoritative (It sounded sure of itself).	<div></div>		

Entry 2 - Chatbot response B

Entry - I work for United Utilities and as per Ofwat's guidance we are looking at some climate change scenario testing. To help us build a case for RCP 2.6 we are trying to pin point the time period on the graph on the document I cited where the temperature projections diverge between RCP 2.6 and RCP 8.5. So what year are RCP 2.6 and RCP 8.5 stop following the same path? I am trying to get hold of the data set behind the graph on the document you produced to help me with this work.

Chatbot response - The temperature projections for RCP 2.6 and RCP 8.5 di-

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was authoritative (It sounded sure of itself).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Why you gave your scores for each of the	<div></div>		

Entry 3 - Chatbot response A

Entry - I have a few questions regarding the correct choice of Member ID (1 - 15), more specifically what the differences are between the perturbed physics IDs. Specifically, I am looking at future heatwaves across the UK. I'm aware of the Sexton et al (2019) paper, however, I am still unsure as to which member ID is most suitable or how best to produce an ensemble of the model members. Averaging seems to remove the extreme events we are interested in. Perhaps a weighted average removing values lower than 2/3 standard deviations of the average for each given day across the 12 members? Would you be able to advise who I might be best contacting regarding the best model member to use for our modelling? Any help is much appreciated!

Chatbot response - The perturbed physics IDs refer to the ensemble members within the UKCP18 climate projections. In the context of the provided excerpts, the perturbed physics ensemble (PPE) consists of different realizations of the climate model with variations in certain parameters to account for uncertainties. In UKCP18, the GC3.05-PPE ensemble consists of 15 realizations of the UK Met Office's HadGEM3-GC3.05 global model, while the CMIP5-13 ensemble includes 13 realizations of selected CMIP5 coupled ocean-atmosphere models. These ensemble members are used to capture a range of possible future climate scenarios based on different parameter settings. For studying future heatwaves across the UK, the most suitable perturbed physics member ID from the GC3.05-PPE ensemble would need to be identified based on the specific characteristics of

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

the question).

The answer was authoritative (It sounded sure of itself). why you gave your scores for each of the

Entry 3 - Chatbot response B

Entry - I have a few questions regarding the correct choice of Member ID (1 - 15), more specifically what the differences are between the perturbed physics IDs. Specifically, I am looking at future heatwaves across the UK. I'm aware of the Sexton et al (2019) paper, however, I am still unsure as to which member ID is most suitable or how best to produce an ensemble of the model members. Averaging seems to remove the extreme events we are interested in. Perhaps a weighted average removing values lower than 2/3 standard deviations of the average for each given day across the 12 members? Would you be able to advise who I might be best contacting regarding the best model member to use for our modelling? Any help is much appreciated!

Chatbot response - The perturbed physics IDs refer to the different ensemble members within the UKCP18 CPM-12 ensemble. These members are created by perturbing the physics of the model to account for uncertainties in the driving data and internal climate variability. The CPM-12 ensemble consists of 12 members, each representing a slightly different version of the model due to these perturbations. For studying future heatwaves across the UK, the member ID that would be most suitable is not explicitly mentioned in the provided excerpts. However, when dealing with extreme events like heatwaves, it is generally rec-

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was			

authoritative (It
sounded sure
of itself).

why you gave your scores for each of the
five criteria *

Entry 4 - Chatbot reponse A

Entry - What data products should be used for the analysis of storms according to the Met Office?

Chatbot response - For the analysis of storms, the Met Office recommends using multiple strands of UKCP data depending on the specific application. It is suggested to consider using the UKCP Global 60km, Regional 12km, and Local 2.2km data products. Additionally, incorporating information from other modeling centers could be beneficial. The UKCP Global projections are suitable for exploring the broader scale in which windstorms are situated, analyzing storm tracks, and weather patterns. For a more detailed analysis at daily and regional scales, the UKCP Global projections that include daily wind data are recommended. Furthermore, for a focus on thunderstorms, the UKCP Local data is ad-

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was why you gave your scores for each of the authoritative (It sounded sure of itself).	<div></div>		

Entry 4 - Chatbot response B

Entry - What data products should be used for the analysis of storms according to the Met Office?

Chatbot response - The data products that should be used for the analysis of storms according to the Met Office include the UKCP Global projections, UKCP Local data, and UKCP Probabilistic Projections. 1. ****UKCP Global Projections****: These projections are recommended for analyzing storms at daily and regional scales. They provide daily wind data and are suitable for exploring the broader scale in which windstorms are situated, analyzing storm tracks, and weather patterns. The UKCP Global projections cover the North Atlantic where windstorms develop. 2. ****UKCP Local Data****: The UKCP Local data is specifically recommended for the analysis of thunderstorms. This model permits the convection that generates thunderstorms and provides details on lightning data ('flashrate'). It includes hourly wind data, hourly rainfall, and temperature data, making it suitable for analyzing storms at a more localized level. 3. ****UKCP Probabilistic**

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
The answer was relevant to the question.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was authoritative (It sounded sure of itself).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please explain why you gave your scores for each of the five criteria *

For the previous questions, which set answers did you prefer? *

☐ Answer A

☐ Answer B

Part 2 - Interacting with the chatbot

For part 2 of the test, you will be given access to the UKCP Chatbot. As part of this interaction you will be asked to complete the following task, finding the relevant information using the chatbot. Please do not exceed 40 minutes interacting with the chatbot.

Task - The MoD needs to construct 30 large buildings by 2030 in various locations around the UK coastline. The MoD would like to ensure that the buildings are suitably prepared to stay cool in the future. Use the chatbot to find relevant information and try to achieve this task.

****All interactions with the chatbot will be recorded for testing purposes. Please do not enter any personal or identifying information that you may not want shared.****

To access the chatbot, go to the following link and select **Pipeline 3**.

.....

Part 2 - Interacting with the chatbot

For part 2 of the test, you will be given access to the UKCP Chatbot. As part of this interaction you will be asked to complete the following task, finding the relevant information using the chatbot. Please do not exceed 40 minutes interacting with the chatbot.

Task - The MoD needs to construct 30 large buildings by 2030 in various locations around the UK coastline. The MoD would like to ensure that the buildings are suitably prepared to stay cool in the future. Use the chatbot to find relevant information and try to achieve this task.

****All interactions with the chatbot will be recorded for testing purposes. Please do not enter any personal or identifying information that you may not want shared.****

To access the chatbot, go to the following link and select **Pipeline 4**.

.....

Part 2 - Chatbot feedback

In this section you will be asked to review your experience of interacting with the chatbot.

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
It was simple to get the information I needed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was authoritative (It sounded sure of itself).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

why you gave your scores for each of

It felt like a conversation.

Is there any additional feedback that you have about the chatbot?

Part 2 - Chatbot feedback

In this section you will be asked to review your experience of interacting with the chatbot.

How much do you agree with the following statements?

*

	Strongly Disagree	Disagree	Neutral
It was simple to get the information I needed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was easy to understand.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was specific (It got right to the point).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was comprehensive (It covered every aspect of the question).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The answer was authoritative (It sounded sure of itself).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

why you gave your scores for each of

It felt like a conversation.

Is there any additional feedback that you have about the chatbot?

Feedback

Thankyou for completing the UKCP Chatbot User Testing Survey.

If you have the time we would appreciate you taking a moment to fill in this feedback checklist.

How much do you agree with the following statements

	Strongly Disagree	Disagree	Neutral
I was asked relevant questions about the UKCP Chatbot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I was able to give all of the feedback I wanted about the Chatbot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Do you have any additional comments about the testing and the survey?

This content is neither created nor endorsed by Microsoft. The data you submit will be sent to the form owner.



Microsoft Forms