



Exploring the Effects of Climate on Plants, Animals and Diseases

Darren J. Kriticos
Gunter F. Maywald
Tania Yonow
Eric J. Zurcher
Neville I. Herrmann
Robert W. Sutherst

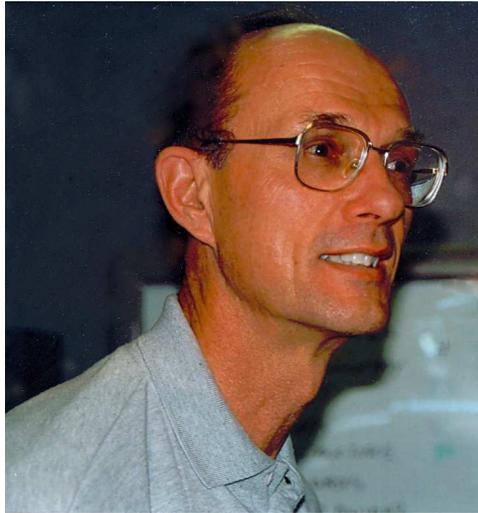


CLIMEX Version 4

Exploring the Effects of Climate on Plants, Animals and Diseases

**Darren J. Kriticos
Gunter F. Maywald
Tania Yonow
Eric J. Zurcher
Neville I. Herrmann
Robert W. Sutherst**





Dedication: This book is dedicated to the memory of Dr. Robert (Bob) W. Sutherst, the inventor of CLIMEX, and a leading light in ecological modelling. Sadly, Bob passed away in May 2013.

Please cite this book as follows:

Kriticos, D.J., Maywald, G.F., Yonow, T., Zurcher, E.J., Herrmann, N.I. and Sutherst, R.W. (2015) CLIMEX Version 4: Exploring the effects of climate on plants, animals and diseases. CSIRO, Canberra. 184 pp.

© Copyright CSIRO 2015, All rights reserved.

Hearne Scientific Software End-User Licence Agreement

This is a legal agreement between you (either an individual or an entity) and Hearne Scientific Software Pty Ltd. By clicking "I accept the terms of this Licence Agreement" and/or by using the Hearne Scientific Software package (herein SOFTWARE), you agree to be bound by the terms of this agreement. If you do not agree to the terms of this agreement, promptly return the SOFTWARE and accompanying items (including materials or other containers) to the place you obtained them for a full refund within thirty (30) days of the purchase date.

1. **Grant of Licence:** Hearne Scientific Software grants you a non-exclusive licence to use the SOFTWARE in accordance with the following terms.
 - a. **Single-User Licence:** If you purchased a single-user licence, Hearne Scientific Software allows one (1) designated individual, and only that individual, the right to install the software on one (1) home, one (1) work, and one (1) portable computer. The designated individual agrees that no more than one (1) installation will be in use at any given time.
 - b. **Group Licence-Limited Installation:** If you purchased a multi-user licence, you may install and use the SOFTWARE on individual computer workstations within your organisation up to the number of users specified in the purchase order. This type of licence is not portable outside your organisation (i.e. The SOFTWARE cannot be installed on a computer at home or on a portable computer unless your organisation purchases a maintenance plan at the time of purchasing the group licence).
 - c. **Special Restrictions:** If you obtained a licence denoted on the purchase documents as 'Academic', 'Student', 'Government' or for 'Developing Country Use', you must only use the SOFTWARE in accordance with the following restrictions:
 - i. **Academic** licences may only be used for teaching and/or non-commercial academic research within an accredited educational institution.
 - ii. **Student** licences may only be used by degree-seeking students enrolled and participating in a course of study at an accredited educational institution.
 - iii. **Government** licences may only be used within a government controlled or funded organisation.
 - iv. **Developing Country Use** licences may only be installed and used exclusively within the list of countries specified by Hearne Scientific Software at the time of purchase.
2. **Transfer:** Hearne Scientific Software further grants you the right to transfer this licence and the SOFTWARE to another party provided the following transfer conditions are met.
 - a. The other party accepts all terms of this agreement.
 - b. All copies of the SOFTWARE are transferred and you discontinue use of the SOFTWARE after transferring.
 - c. Hearne Scientific Software is promptly notified of the name and address of the other party and the serial number of the SOFTWARE.
 - d. Hearne Scientific Software is not required to supply new media.
3. **Term and Termination:** Failure to comply with any of these terms will terminate this agreement and your right to use the SOFTWARE. You may also choose to terminate the agreement at any time. Upon termination of this agreement, you must immediately destroy the SOFTWARE and all copies of it.
4. **Copyright:** The SOFTWARE (including any images, applets, photographs, animations, video, audio, music and text incorporated into the SOFTWARE) is owned by CSIRO and is protected by Australian copyright laws and international treaty provisions. You agree not to modify, adapt, translate, reverse engineer, decompile, or disassemble the Software. You must treat the SOFTWARE like any other copyrighted material (e.g. a book or musical recording) except that you may either:
 - a. Make one copy of the SOFTWARE solely for backup or archival purposes; or

- b. Transfer the SOFTWARE to a single hard disk provided you keep the original solely for backup or archival purposes.
- c. You may only copy the printed materials accompanying the SOFTWARE for your own internal use.
5. Without prejudice to Clause 4 above, rights (including copyright) in all enhancements, improvements, modifications, adaptations or translations of the SOFTWARE made pursuant to this licence shall belong to CSIRO. You hereby assign (and agree to ensure that any other creator of such rights assigns) to CSIRO all such rights, and you agree to perform all actions necessary to vest such rights in CSIRO.
6. **Limited Warranty:** Hearne Scientific Software warrants that the SOFTWARE will perform substantially in accordance with the accompanying materials for a period of thirty (30) days from the date of purchase. Any implied warranties on the SOFTWARE are limited to thirty (30) days. Some states/provinces do not allow limitations on duration of an implied warranty, so the above limitation may not apply to you.
7. **Customer Remedies:** Hearne Scientific Software's entire liability and your exclusive remedy shall be, at Hearne Scientific Software's option:
 - a. Return of the price paid; or
 - b. Repair or replacement of the SOFTWARE that does not meet Hearne Scientific Software's limited warranty and that is returned to Hearne Scientific Software with a copy of your receipt. This limited warranty is void if failure of the SOFTWARE or hardware has resulted from accident, abuse, or improper application. Any replacement of SOFTWARE will be warranted for a further thirty (30) days.
8. **No Other Warranties:** To the maximum extent permitted by applicable law, Hearne Scientific Software disclaims all other warranties, expressed or implied, including but not limited to implied warranties of merchantability and fitness for a particular purpose, with respect to the SOFTWARE and the accompanying written materials. You may have others, which vary from state to province to province.
9. **NO LIABILITY FOR CONSEQUENTIAL DAMAGES:** To the maximum extent permitted by applicable law, in no event shall Hearne Scientific Software or its suppliers be liable for any special, incidental, indirect, or consequential damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or any other pecuniary loss) arising out of the use of or inability to use the SOFTWARE product, even if Hearne Scientific Software has been advised of the possibility of such damages.
10. **Governing Law.** This Agreement will be governed by the laws in force in the State of Victoria, Australia.

Should you have any questions concerning this agreement please contact Hearne Scientific Software.

Hearne Scientific Software Pty Ltd
3/200 Toorak Rd
South Yarra VIC 3141, Australia
Ph +61 (03) 9020 5730
www.hearne.com.au

Table of Contents

Table of Contents	v
Notice	1
1 Introduction	1
1.1 WHAT IS CLIMEX?	3
1.2 MINIMUM REQUIREMENTS TO RUN CLIMEX.....	4
1.3 INSTALLING AND RUNNING CLIMEX	5
<i>To install CLIMEX on your computer</i>	5
<i>To load a CLIMEX application</i>	5
1.4 CHANGES FROM VERSION 3	6
2 An Overview of CLIMEX Functions	8
3 The CLIMEX Model	9
3.1 COMPARE LOCATIONS.....	9
3.2 COMPARE LOCATIONS (1 SPECIES)	10
3.3 COMPARE LOCATIONS (2 SPECIES)	11
3.4 COMPARE LOCATIONS (1 SPECIES, EXTENDED)	11
3.5 COMPARE LOCATIONS (SPECIES INTERACTION).....	12
3.6 COMPARE LOCATIONS (GENETIC ALGORITHM)	13
3.7 COMPARE YEARS	14
3.8 COMPARE LOCATIONS/YEARS	14
3.8.1 <i>Outputting the Map Sequence as a video</i>	21
4 Climate Matching	23
4.1 MATCH CLIMATES	23
4.2 MATCH CLIMATES (REGIONAL).....	25
5 The Conceptual Basis of the CLIMEX Model	26
5.1 SPECIES DISTRIBUTION AND ABUNDANCE	26
5.2 THE CLIMEX INDICES	28
5.3 GROWTH RELATED INDICES	29
5.3.1 <i>Growth Index (GI_A and GI_W)</i>	29
5.3.2 <i>Temperature Index (TI)</i>	30
5.3.3 <i>Moisture Index (MI)</i>	33
5.3.4 <i>Radiation Index (RI)</i>	35
5.3.5 <i>Substrate Index (SVI)</i>	37

5.3.6	<i>Diapause Index (DI)</i>	38
5.3.7	<i>Light Index (LI)</i>	39
5.3.8	<i>Biotic Index (BI)</i>	40
5.4	STRESS INDICES.....	40
5.4.1	<i>Cold Stress (CS)</i>	42
5.4.2	<i>Heat Stress (HS)</i>	44
5.4.3	<i>Dry Stress (DS)</i>	45
5.4.4	<i>Wet Stress (WS)</i>	46
5.5	INTERACTION STRESS INDICES.....	46
5.5.1	<i>Cold-Dry Stress (CDX)</i>	47
5.5.2	<i>Hot-Dry Stress (HDX)</i>	47
5.5.3	<i>Cold-Wet Stress (CWX)</i>	47
5.5.4	<i>Hot-Wet Stress (HWX)</i>	47
5.6	LIMITING CONDITIONS.....	48
5.6.1	<i>Length of growing season (PDD)</i>	48
5.6.2	<i>Obligate Diapause</i>	48
5.7	ECOCLIMATIC INDEX (EI).....	49
5.8	CORE DISTRIBUTION.....	50
6	Fitting CLIMEX Models	51
6.1	PARAMETER FITTING VIA INFERENCE FROM DISTRIBUTION.....	51
6.1.1	<i>Temperature parameters (DV0 - DV3)</i>	57
6.1.2	<i>Moisture Parameters (SM0 - SM3)</i>	57
6.1.3	<i>General Parameter Fitting Example</i>	58
6.1.4	<i>Diapause Parameter Fitting Examples</i>	66
6.2	COMPARE LOCATIONS MODEL-FITTING TIPS.....	70
6.3	DISTRIBUTION DATA.....	70
6.3.1	<i>Epistemic uncertainty</i>	70
6.3.2	<i>Confounded uncertainty</i>	70
6.4	GROWTH INDICES.....	71
6.4.1	<i>Soil Moisture</i>	72
6.5	STRESSES.....	72
6.5.1	<i>Cold Stress</i>	72
6.5.2	<i>Heat Stress</i>	73
6.5.3	<i>Wet Stress</i>	73
6.5.4	<i>Dry Stress</i>	73
6.5.5	<i>Hot-Wet Stress</i>	73
6.5.6	<i>Cold-Wet Stress</i>	73

6.5.7	<i>Hot-Dry Stress</i>	74
6.5.8	<i>Cold-Dry Stress</i>	74
6.6	MODEL COMPLEXITY AND OVER-FITTING.....	74
6.7	SEMI-AUTOMATIC PARAMETER-FITTING USING THE GENETIC ALGORITHM.....	74
6.8	PARAMETER SENSITIVITY ANALYSIS	88
6.9	PARAMETRIC MODEL UNCERTAINTY	91
6.10	ASSESSING THE GOODNESS OF FIT OF A MODEL.....	92
7	Matching Climates	95
7.1	MATCHING CLIMATES USING REGION-FITTING	97
7.2	PREPARING THE LOCATION SET.....	97
7.3	SETTING THE REGION MATCH OPTIONS.....	99
7.4	RUNNING AND DISPLAYING THE RESULTS.....	100
8	Tasks	101
8.1	LOADING A CLIMEX APPLICATION	101
8.2	OPERATIONS ON SPECIES FILES.....	102
8.3	OPERATIONS ON SPECIES PARAMETERS.....	103
8.4	USING CLIMATE CHANGE SCENARIOS.....	107
8.5	USING IRRIGATION SCENARIOS	110
8.6	SPECIFYING LOCATION SELECTIONS	110
8.7	ADDING NEW LOCATIONS TO THE DATABASE.....	112
8.8	USING THE GRID-BASED CLIMATE DATA.....	112
8.9	CREATING AND MODIFYING MAP BACKGROUNDS.....	112
8.10	EXTRACTING RUN INFORMATION FOR USE IN A SPREADSHEET OR GIS.....	116
8.11	CONVERTING SPECIES FILES FROM VERSION 1 TO VERSION 2.....	118
9	Advanced Tasks	120
9.1	SAVING OUT DETAILED RESULTS (OUTPUT STREAMING).....	120
9.1.1	<i>Output Stream in Use</i>	120
9.1.2	<i>Detail Stream Properties</i>	121
9.1.3	<i>Output File</i>	121
9.1.4	<i>Summary Stream Properties</i>	122
9.2	CREATING AUTOMATED RUN SEQUENCES.....	123
9.2.1	<i>Parameter sequence</i>	124
9.2.2	<i>Species sequence</i>	125
9.2.3	<i>Compound sequence</i>	128
9.3	CREATING COMPOSITE RISK MODELS	129
10	The Meteorological Data Manager facility (MetManager)	131

10.1	CREATING A NEW DATABASE	132
10.2	EDIT LOCATION SELECTIONS	134
10.3	IMPORT METEOROLOGICAL DATA	136
11	Useful modelling resources	138
11.1	SPECIES DISTRIBUTION DATA	138
11.1.1	<i>Global Biodiversity Information Facility (GBIF)</i>	138
11.1.2	<i>CABI</i>	138
11.2	CLIMOND.....	138
11.3	IRRIGATION	139
11.4	LAND USE	139
11.5	LAND COVER	139
11.6	CROPPING AREAS.....	139
11.7	CROP CALENDAR	139
11.8	POTENTIAL VEGETATION	140
11.9	GLOBAL ADMINISTRATIVE AREAS (GADM).....	140
11.10	PLACENAMES (GAZETEER).....	140
11.11	DIGITAL ELEVATION DATA	140
12	Frequently Asked Questions	140
13	References.....	143
14	Appendix 2. The LOC/MET file format.....	155
15	Appendix 3. The Species File format.....	158
16	Version history.....	162
16.1	MAJOR CHANGES FROM VERSION 2	162
16.2	MAJOR CHANGES FROM VERSION 1.1 TO 2.....	162
17	Index.....	164

Notice

IMPORTANT - PLEASE READ THIS

If you are a typical software user you will not read much of this guide.

If nothing else please read this page before you start.

1. If you have an older version of CLIMEX, please read the section outlining the changes that have been made since the previous version. A lot has changed since version 3!
2. *This User's Guide contains the theory behind CLIMEX and an overview of the major functions and uses of CLIMEX. Instructions on how to use functions within CLIMEX can also be found in the CLIMEX and DYMEX Help files. The Help file is accessed from within the program and from context sensitive Help Buttons within specific functions. If you have not used CLIMEX before, we strongly recommend you read this guide from cover to cover before you try any serious modelling. Understanding the theory is essential to the effective use of the program. Experienced CLIMEX users will also benefit from reading this guide as it contains expanded explanation of the theory behind CLIMEX and re-interpretations of functions within CLIMEX.*
3. New material has been added, including detailed instructions on model-fitting procedures and modelling resources.
4. Specific instructions on the operation of each feature of the software are included in the Tutorials that accompany the software. More detailed instructions on the software are contained in the DYMEX User Guide.
5. We welcome any suggestions on how to improve CLIMEX. If you have any comments or ideas, please contact:

Prof. Darren J. Kriticos
CSIRO Ecosystem Sciences, Black Mountain Laboratories
Clunies Ross Street, Acton, Australian Capital Territory, Australia 2600
Tel 61 (0)2 6246 4252
Fax 61 (0)2 6246 4000
Email Darren.Kriticos@csiro.au

1 Introduction

CLIMEX Version 4 is an upgraded and enhanced version of the CLIMEX 3 program. In the same way as its predecessor, it has been built using the DYMEX generic and modular population modelling software (<http://hearne.com.au/products/dymex/>). The major changes to the software are outlined in Section 1.4. Changes to the look and feel of the program from Version 3 have been kept to a minimum, and species

parameter and metadata files are compatible between Versions 3 and 4. This should allow users who are familiar with CLIMEX Version 3 to take immediate advantage of the new features, with minimal need to learn new operating techniques. The goals of this revision are to:

1. Make the programme easier to use, automating many tedious, repetitive tasks
2. Address the emerging concerns in the pest risk literature about exposing model uncertainty
3. Calculating and reporting the goodness of fit of models
4. Improving the semi-automated fitting of stress parameters
5. Open up new techniques for exploring the dynamic nature of climate suitability using time series models and maps

Along the way we have improved the speed of the model tenfold, eliminated many memory leaks, made the program more stable, and improved many of the dialogues.

Some conventions used throughout this guide are as follows:



The exclamation symbols throughout the manual represent important notes.



The light bulb represents examples and ideas that can be used when constructing a model.

1.1 What is CLIMEX?

The CLIMEX software package contains two quite different climate modelling tools. There is the CLIMEX model (referred to as 'CLIMEX' or as the 'CLIMEX model'), and the CLIMEX 'Match Climates' function. The latter is a tool for comparing the meteorological data of different places without reference to any particular species, and we will return to it later. The CLIMEX simulation model was first described by Sutherst and Maywald (1985) and a number of enhancements and further caveats and insights into using the model have been described in a series of publications listed at the end of this Guide, particularly Sutherst *et al.* (1995), Sutherst *et al.* (1998), and Kriticos *et al.* (2005). The CLIMEX model is based *partly* on the assumption that if you know where a species lives you can infer what climatic conditions it can tolerate. This is the cardinal assumption underpinning most species distribution models. However, where other models such as Bioclim, Domain, GARP and MaxEnt attempt to characterise the environment occupied by the species, CLIMEX simulates the mechanisms that limit species' geographical distributions and determine their seasonal phenology, and to some extent their relative abundance. Where most models focus on describing the relationship between the occurrences of the species with respect to static environmental covariates, CLIMEX describes how the species responds to climatic variables at appropriate temporal scales (daily or weekly). In this respect, CLIMEX shares a similarity of approach with Nichemapper (Kearney and Porter, 2009), but is less complex.

In CLIMEX the species functional responses are formulated in accordance with a set of theory supported by more than a century of scientific study of the response of species to climatic variables. The Growth Indices and Stress Indices accord with the *Law of Tolerance* (reviewed in Shelford, 1963), and the Growth Indices are combined in accordance with the *Law of the Minimum* (reviewed in van der Ploeg *et al.*, 1999). These deep-seated ecological roots distinguish CLIMEX, and provide the basis for modelling the potential distribution of species successfully in novel climates. CLIMEX enables the user to estimate the potential geographical distribution and seasonal abundance of a species in relation to climate. It does not try to match the patterns of climate and species' distribution in the same way that a correlative statistical model would seek to achieve. More on this later.

CLIMEX is applied to a species by selecting values for a set of parameters that describe its response to temperature, soil moisture and if applicable, light. In this guide the term 'population' is used as the target entity, representing an average population of an animal or plant species or biotype for example. An Annual Growth Index (GI_A) describes the potential for growth of a population during the favourable season. Four stress indices (Cold, Hot, Wet and Dry), and in some cases their interactions, describe the extent to which the population is reduced during the unfavourable season(s). The Growth and Stress Indices are combined into an Ecoclimatic Index (EI), to give an overall measure of favourableness of the location or year for permanent occupation by the target species. Two limiting conditions, i.e., the minimum length of the growing season and obligate diapause, act as overall constraints to the EI value where relevant. Results are presented as tables, graphs, maps or movies.

CLIMEX can combine inductive and deductive modelling techniques. A species' climatic requirements can be inferred from its known geographical distribution (either in its native range or in another region where it has been established for a long time), relative abundance or seasonal phenology. Some laboratory data, such as developmental threshold temperatures, can be used to fit some CLIMEX parameter

values. Initial estimates of parameter values are fine-tuned by comparing the resulting indices with the known presence or absence, seasonal phenology and, preferably, relative abundance of the species in each location.

Once the parameter values have been estimated, and where possible validated against geographically independent data, CLIMEX can be used to make projections for other, independent locations or climate scenarios. Independent data means that there is no connection between these data, and those used for fitting the model; hence it is not appropriate to sub-sample a geographical distribution randomly, and then use the remaining data to test the model. This form of resampling may only be suitable for modelling species native distributions, in which case it provides no information about the likely ability of the model to project accurately into novel climate situations.

Despite its mechanistic structure, the CLIMEX model is usually parameterised using correlation of field observations on species occurrence and seasonal phenology with environmental variables, (e.g., Vera *et al.*, 2002; Kriticos *et al.*, 2005; Sutherst and Maywald, 2005; de Villiers *et al.*, 2012). The dynamic nature of the model, with its temporal resolution of 1-week and its ability to account for interactions between environmental variables, enables it to function reliably across all latitudes, climate zones and seasons as required for biosecurity and global change applications. It can also be informative for tracking the suitability of migration routes (Zalucki *et al.*, 2004).

A further strength of the CLIMEX method is the cross-validation that is inherent in the requirement for consistency between the estimated potential geographical distribution and the seasonal phenology of a target species in different parts of its range (e.g., de Villiers *et al.*, 2012). The two measures have to be independently consistent with field observations to obtain a model that is internally consistent on both spatial and temporal scales. In addition, the parameters used to generate the model should be biologically reasonable, according with what is known about the species from experimental or theoretical domains.

These features of CLIMEX have repeatedly provided new insights into the ecology of modelled species, and users are encouraged to 'listen to CLIMEX' when fitting parameter values. The process helps to avoid the tendency to force-fit parameters based on empirical data or incorrect pre-conceptions of how a system operates. New regional occurrence records for species have been captured based on guidance from the results of the CLIMEX model e.g. (Hall and Wall, 1995; Kriticos *et al.*, 2007; Sutherst and Bourne, 2009) as well as new biological control opportunities that were successful (Dhileepan *et al.*, 2006). Pardey *et al.* (2013) used CLIMEX to identify global areas with different risks of damage from *Puccinia graminis* (wheat stem-rust) in order to improve estimates of the necessary investment in research to keep pace with the emergence of new strains of resistance.

1.2 Minimum Requirements to Run CLIMEX

The minimum requirements to run CLIMEX V4 are:

32 bit Windows Vista

- 2 Gb memory
- 3 Gb of disk space

The large disk space requirement is due to the inclusion of some sample meteorological data.

1.3 Installing and Running CLIMEX

To install CLIMEX on your computer

1. Run CLIMEXsetup.exe.
2. Follow the installation guide.

If Version 3 of the DYMEX program has been installed previously, the CLIMEX components can be added by copying the CLIMEX model files (in directory \Dymex\Models\Climex\Species) to the Models directory on your hard drive.

To load a CLIMEX application

See Section 8.1 of this document for instructions on loading a CLIMEX application.

1.4 Changes from Version 3

- Speed increased massively!
 - The model now runs approximately ten times faster than Version 3.
- Improved automatic parameter-fitting
 - The semi-automatic fitting of model parameters has been greatly simplified and modified to remove the need to specify pseudo-absences. The fitting algorithm now requires simply a set of location points in shapefile format. Diapause parameters can now be included in the fitting process.
- The Compare Locations model can now create weekly output maps for the stress and growth indices.
- A new model has been added, the CLIMEX Compare Locations/Years, which allows inter-annual variation in climate suitability to be explored.
- You can now make movies of seasonal and interannual patterns of climate suitability.
 - Time series outputs from Compare Locations and Compare Locations/Years can be viewed as movies. Weekly or annual summary variables can be viewed as a movie in order to assess the effects of seasonal or interannual variability in climate.
- Automated parameter sensitivity and model uncertainty analyses are available.
- The set of points in a reference dataset containing climate locations matching the set of location records for the species being fitted can now be mapped as a variable called "Reference Data".
- The regional climate matching function can now handle much larger datasets. We have tested it with hundreds of reference points (home locations) running with the CliMond 10' world grid.
- Batch parameter runs can now be sequenced, automating some highly repetitive tasks.
 - This has required an extension to the user interface (Fig 1-1). The sequencing allows multiple species models to be run in sequence, or multiple parameter options can be run for the same species. This system is powerful, but quite complex, and is capable of generating extremely large amounts of data. You should read section 9.1 of the manual very carefully before attempting to use these functions.

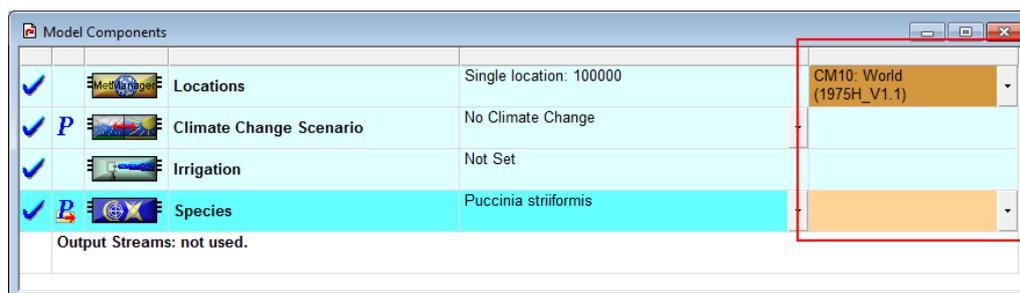


Fig 1-1 Model Components window highlighting the sequencing components added to the control window.

- CLIMEX can now be installed on computers running the 64-bit version of Windows. However, the present version of CLIMEX is still compiled as a 32 bit application and does not access more than 3.2 Gb of memory.
- This User Guide includes a section on modelling tips and resources.

2 An Overview of CLIMEX Functions

The CLIMEX software package now contains ten ‘applications’ that are optimised for different tasks. Each of the applications is in fact a separate “DYMEX model”. The term ‘application’ is used in order to avoid confusion between these and the **CLIMEX model** in this manual. The applications can be grouped into two types. Eight of them are based around the “CLIMEX Model” (described in detail in Section 5) while two others carry out a simple climate matching functions. Each of these applications is described later in this Section. The “CLIMEX Model” based functions are further grouped into those that carry out the “Compare Locations” task, one that does the “Compare Years” task, and a hybrid that does both Compare Years and Compare Locations simultaneously. The applications marked with an asterisk are those that were added in this Version of CLIMEX, adding new functionality as described in the corresponding section below. The Compare Locations (Genetic Algorithm) model uses a genetic algorithm to undertake semi-automated parameter fitting. This application has been extensively enhanced to make it easier to use, and dispensing with the need to include pseudo-absence data.

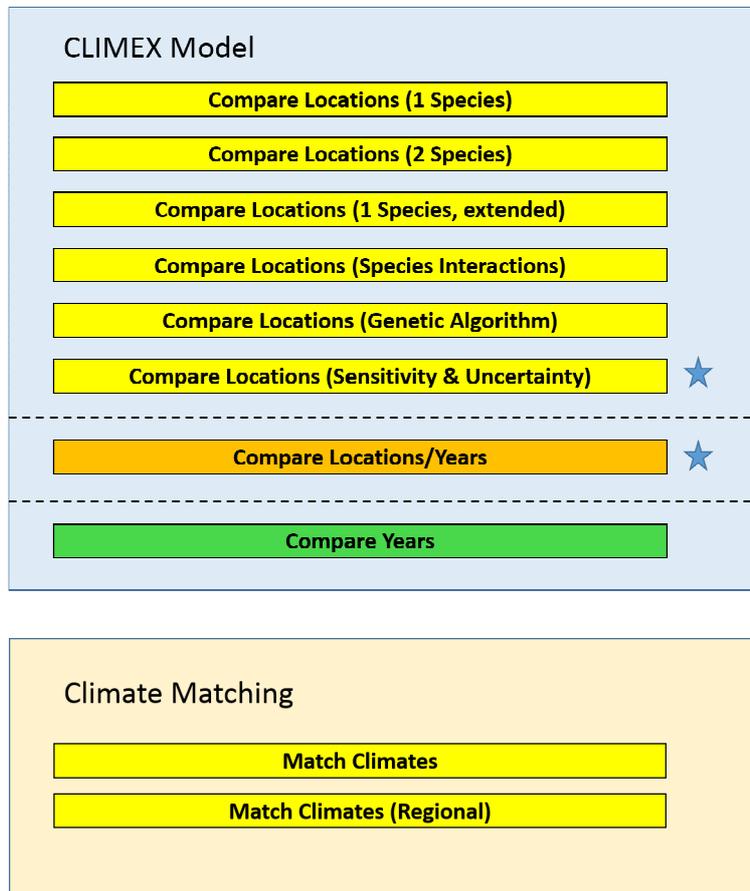


Fig. 2-1. The ‘applications’ included in the CLIMEX package, grouped into their functional types. The applications marked with a star are new in Version 4.

3 The CLIMEX Model

The CLIMEX model can be run in three different modes: Compare Locations, Compare Years and Compare Locations/Years. With Compare Locations, the response of a species to the long-term average climates of different locations is compared. In the Compare Years mode the species response to climates in different years at a single location is compared. In the Compare Locations/Years mode climate suitability is compared across both time and space for selected sets of Climate Stations. This means that there are three different types of meteorological data involved in CLIMEX: Long-term climate averages, weather sequences for individual stations, and spatio-temporal time series data. Section 5 provides a thorough conceptual treatment of the CLIMEX model and should be read before using CLIMEX.

3.1 Compare Locations

The Compare Locations model enables the user to estimate the potential geographical distribution of a species or a pair of species in relation to climate only, based on its climatic preferences. This function calls upon a meteorological database consisting of monthly long-term average climatic variables for any number of locations. A set of parameters describes the species' response to the average climate of a given location, with the implicit assumptions that (i) the species is able to survive the annual variation in climate that is associated with that location, and (ii) that the annual variation at all locations is equal (to comply with the statistical requirement that comparison of means assumes equal variances). This latter assumption is not true and needs to be considered when comparing results from regions with different annual variation such as occurs when comparing areas with continental climates, like the southern USA or southern China, with other areas that do not experience the occasional very extreme seasons, like much of Africa for example. Currently, there is no global database available to allow CLIMEX to take account of the differences in variation between regions.

The model parameters are broken into three groups: one group defines seasonal population growth, one defines stresses that limit the geographical distribution, and the third describes limiting conditions that may exclude species from some areas. An iterative process of comparing the known and modelled distributions for the same region is used to derive values for the parameters. Once the parameter values have been defined, they can be used to estimate the potential distribution of the species in other independent locations or novel climates.

The Compare Locations function is used primarily for describing the potential distribution of a species as determined by climate. Being able to describe and map the potential distribution of species enables policy makers at both national and international levels to make rational decisions about managing pest threats or redistributing natural enemies. For example, if the Compare Locations function indicates that areas are climatically unsuitable for a potential biocontrol agent, there is little point in introducing such a species into that area. From a pest risk analysis perspective, a Compare Locations run may indicate that an area is highly suitable for an exotic pest species and so could cause losses of production or damage to the natural environment.

Once the species' parameters have been defined adequately, the Compare Locations function also enables the user to examine the relative seasonal potential for growth of the species in any given location. Such data can be used to train and verify the

performance of a model (e.g., de Villiers *et al.*, 2012; Kriticos *et al.*, 2015). Having a relative measure of the species' likely performance in a given location can also inform the local management of pest species.



Since defining a species' parameter file involves estimating the stresses that limit its distribution, the Compare Locations function contributes towards understanding the biology of a species for which the user may have little biological information. The process of fitting the parameter values may help a user to identify areas where further research on some aspect of the biology (e.g., cold tolerance, or adaptations to heat stress) might prove useful. The parameter-fitting process frequently highlights inconsistencies between distribution data, meteorological data, reported behaviour of the species in relation to meteorological variables, and theoretical expectations of species tolerance limits. In approaching these issues, we have found it useful to adopt a philosophical framework of multiple working hypotheses (Chamberlin, 1965), exploring the veracity of all elements.

The Compare Locations function is also useful in selecting field sites for a research project. Ecological research often demands that field sites represent the range of conditions experienced by the species, from marginal to optimal. A Compare Locations run will help an ecologist readily identify the most representative field sites. Furthermore, it is usually the marginal sites that are the most interesting to the researcher. This function of CLIMEX enables the identification of sites that are marginal for different reasons (e.g., because they are too hot, too dry, too cold, or too wet), and thus helps the researcher to select sites that will yield interesting results.

By running CLIMEX with average data from different years, some indication of the relative changes in the geographical range can be obtained. However, note that averaging smooths data, and so such results cannot be compared directly with those from simulations using long-term average data (see below).

3.2 Compare Locations (1 species)

This application applies the CLIMEX model to a single species. It is normally the most frequently used of the CLIMEX applications and its main window is illustrated in Fig. 3-1.

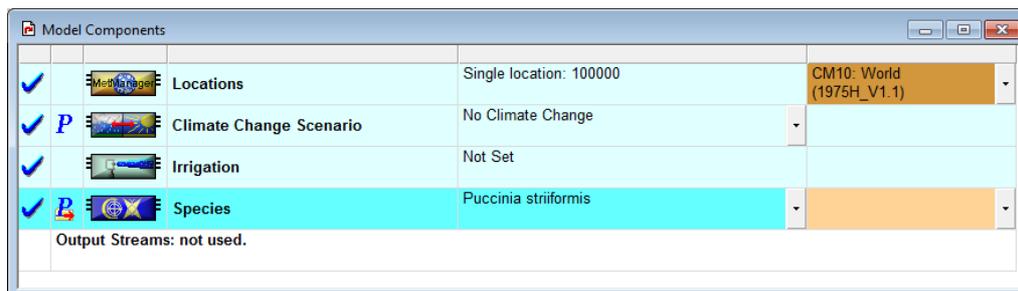


Fig. 3-1 The Compare Locations (1-species) application.

The set of locations for which the simulation is to be run is specified in the top (**Locations**) panel (see Section 8.6 for details on how to do this). Optionally, a climate change scenario (Section 8.4) and/or irrigation regime (Section 8.5) can be set from the corresponding panels. The species is selected in the **Species** panel (Section 8.2).

3.3 Compare Locations (2 species)

The ‘Compare Locations (2 species)’ just adds a second species to the facilities offered by the ‘Compare Locations (1 species)’ application (Fig. 3-2). The two species are completely independent; there is no interaction between them. A common reason for using this application is to plot both the species of interest and its host on the same map for convenient comparison.

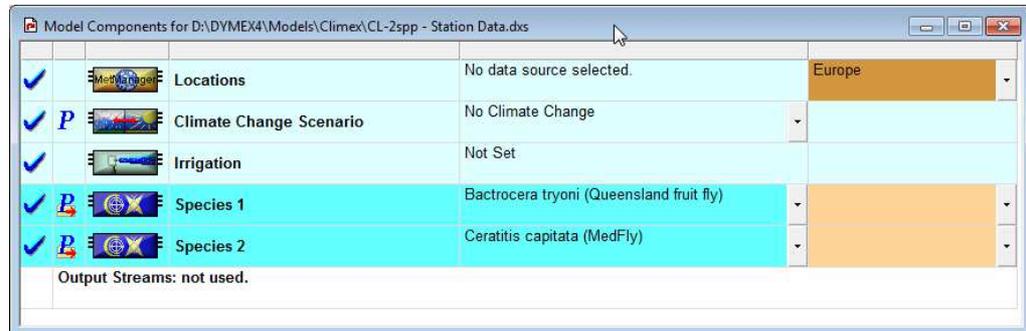


Fig. 3-2. The Compare Locations (2-species) application.

The setup for this application is the same as that for the 1-species application (Section 3.2), with only the addition of the second species that needs to be selected.

3.4 Compare Locations (1 species, extended)

This application models a single species, but adds several components to the Growth Index (GI). These are a Physical Substrate Index and a Biotic Substrate Index (discussed in Section 5.3.5) and a Radiation Index (Section 5.3.4). The radiation value that is used to calculate the Radiation Index can be modified by slope properties (aspect and tilt), whose values can be constant or read from the data file for each location.

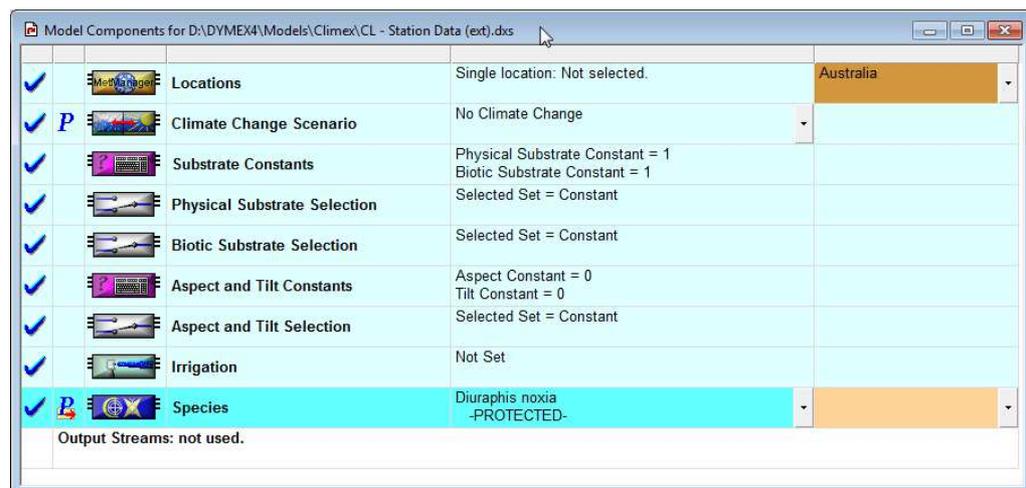


Fig. 3-3. The Compare Locations (1-species, extended) application.

The Component Window for this application is shown in Fig. 3-3. Each of the “substrate” variables can be either the same value for all locations or be different for each location. This is specified in the **Physical Substrate Selection** and **Biotic**

Substrate Selection panels, respectively. If “Constant” is selected for either substrate variable, the value that is used is that set in the **Substrate Constants** panel. If one of these substrate variables is set to be “Location-specific”, its value must be available from the MetManager data file specified in the Location Selection that is being used (as set in the **Locations** panel).

The radiation value for calculating the Radiation Index (RI) is affected by the settings for the aspect and tilt of the slope onto which the radiation is incident. Aspect and tilt can be either constant (i.e., have a common value for all locations) or set separately for each location. If selected to be constant, the values set in the **Aspect and Tilt Constants** panel. If these variables are set to be “Location-specific”, their values must be available from the MetManager data file specified in the Location Selection that is being used (as set in the **Locations** panel).

3.5 Compare Locations (Species Interaction)

The ‘Compare Locations (Species Interaction)’ application (Fig. 3-4) models two species simultaneously, but in contrast to the ‘Compare Locations (2 species)’ application, competition or synergy between the species can be included explicitly.

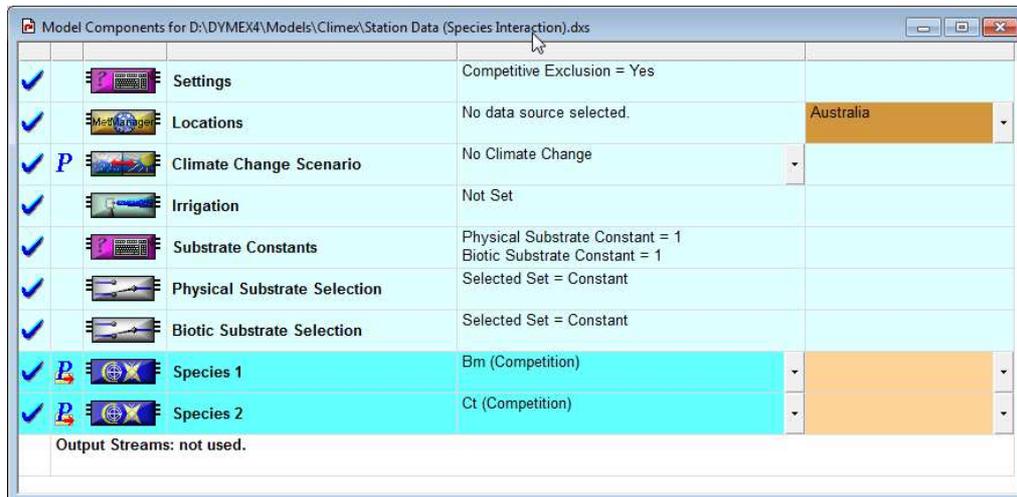


Fig. 3-4. The Compare Locations (Species Interaction) application.

Note that this application allows use of Physical and Biotic Substrate Indices, as well as the Radiation Index (if required). The effect of competition or synergy is modelled via a Biotic Index that affects the value of the Growth Index, and its calculation is detailed in Section 0. In some cases of competitive interactions, the species that “wins” will exclude the other species completely. This can be modelled by selecting “Competitive Exclusion” in the **Settings** panel. The substrate variables can be set to be the same for all locations, or they can vary between locations, as described for the ‘Compare Locations (1-species, extended)’ application (Section 3.4).

3.6 Compare Locations (Genetic Algorithm)

3.7 This special application is used only for the semi-automatic fitting of parameters using the built-in genetic algorithm [see Model complexity and over-fitting]

Deciding on an appropriate level of model complexity is often a difficult challenge, especially for inexperienced modellers. If you include too many range-defining factors in a model, the model is able to increase its specificity, you may find opportunities to “explain” additional facets of the species range. However, there is a risk that you are “explaining” sampling artefacts such as biases, biotic effects such as competition or predation, or teleological artefacts (e.g. an invasion lag). These could all lead to model projections that contain a conservative error. This phenomena is called model over-fitting. Fortunately, CLIMEX limits the number of range-defining factors that can be used, so there is less danger of over-fitting a model than when using a correlative species distribution model. Generally, in CLIMEX models interaction stresses are not required, and when present only one is likely to be needed.

Semi-automatic parameter-fitting using the genetic algorithm]. It is a simplified CLIMEX model that does not include any Growth Index components or the ability to set climate change scenarios or apply irrigation (Fig. 3-5). In the previous version of CLIMEX this model was called the Compare Locations (Stress) model. It was renamed to make it more intuitively obvious to users which model it was.

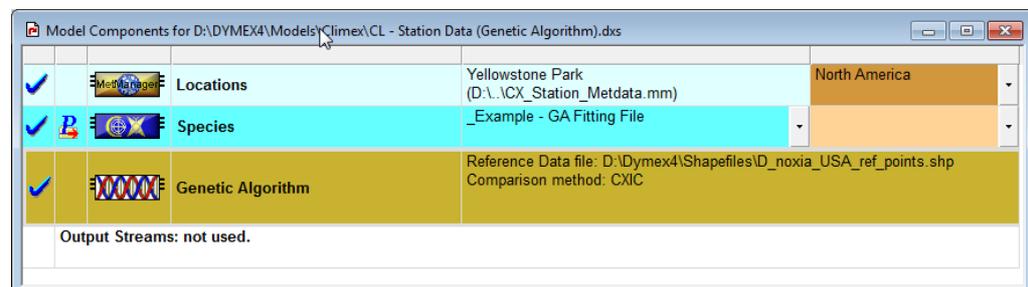


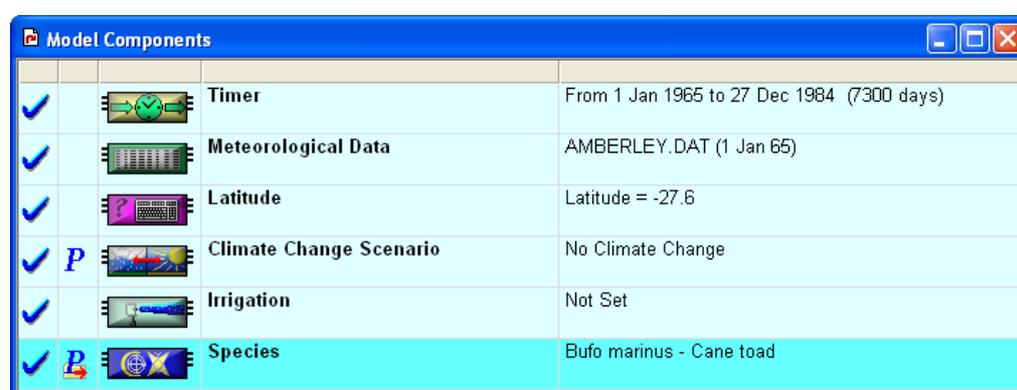
Fig. 3-5. The Compare Locations (Genetic Algorithm) application.

We stress here that the genetic algorithm model should be considered a supervised process. There may be several sets of parameters that can be used to define models that accord similarly with any given set of distribution data in the training environment. A more reliable, and generalizable model will result if the modeller applies knowledge of the species biology or ecology to select stresses for inclusion in the model and to constrain parameter ranges to biologically reasonable values. Candidate models should be critically evaluated in relation to whether the stresses are likely to act on the organism, whether the fitted parameters are reasonable and whether the resulting climate suitability patterns agree with independent data and expectations based on an appreciation of the range limits of similar organisms. You should be prepared to challenge the distribution data used to fit the model by specifying what you believe are biologically reasonable values and critically scrutinising any points that the model then indicates are unsuitable. Similarly, you should be prepared to challenge your pre-conceptions about the species and which factors limits its range in different areas.

3.8 Compare Years

The Compare Years application uses a database consisting of consecutive years of monthly data from the same location, rather than long-term average data from many different places as in the Compare Locations function. Compare Years lets the user examine the effect of climatic variation on the potential population growth and survival of a species over consecutive years. The user can enter weekly or monthly climatic data for any number of years and CLIMEX will display the species' response for each year.

Another use for the Compare Years facility is to compare the times in different years when the climate of representative locations on a migration path become suitable for the establishment of migrant insects. This enables the user to differentiate between climate and arrival times as factors that limit the appearance of migrants in different years.



Component	Value
✓ Timer	From 1 Jan 1965 to 27 Dec 1984 (7300 days)
✓ Meteorological Data	AMBERLEY.DAT (1 Jan 65)
✓ Latitude	Latitude = -27.6
✓ P Climate Change Scenario	No Climate Change
✓ Irrigation	Not Set
✓ Species	Bufo marinus - Cane toad

Fig. 3-6 The Compare Years application.

It is important to note that daily data averaged over weeks or even months will still have greater variability than the long-term average data with which the original parameter values were (usually) estimated by inference from the geographical distribution. In order to minimise the incompatibility between the different data sets it is desirable first to average daily data to monthly and allow CLIMEX to interpolate it back to weekly averages. Secondly, it is inevitable that even monthly data for a particular year will generate more extreme values than occur in long-term average data, so the results from Compare Years and Compare Locations should not be compared. The values of each index in the Compare Years output can be compared with each other as they are relative measures of growth or stress. Often, stress values at locations within the normal species' distribution will exceed 100 for short periods or years but this should not be interpreted as precluding the existence of the species. The value of 100 is estimated from smoothed, long-term average meteorological data and long-term average geographical distributions and so represents an average value.

3.9 Compare Locations/Years

The Compare Locations/Years application allows a modeller for the first time to visualise the spatio-temporal dynamics of climate suitability. Where the Compare Locations model provides a map of average conditions, and the Compare Years model gives a temporal picture of conditions at a single location, Compare Years/Locations creates map sequences that allow you to see how the suitability changes in both space and time.

The Compare Locations/Years application uses a database consisting of a time series of monthly station data, rather than long-term average data from many different places as in the Compare Locations function. Where the Compare Years function can compare the effect of inter-annual variation on a single location, Compare Locations/Years allows users to explore the impact of inter-annual variation in climate on the suitability of a selection of locations for a species. The output from this module is a series of maps, one for each week or year, depending on the selected variable.

The ability to simulate and visualise the climatic factors that influence a species range in space and time simultaneously is a powerful new tool. For the first time, the dynamic aspects of range limits can be visualised.



Figure 3-7 Model components window for Compare Locations/Years

The most obvious difference from the Compare Locations model is the inclusion of the Meteorological Data module and the Cropping Season module in the model components window (Fig 3-7).

Compare Locations/Years uses a monthly time-series of climate data. It is presently set up to run using time series data in GRIM format. The grim format is structured by a header consisting of 5 lines:

line 1, attribution and date stamp, format: a

line 2, variable, format: a4,a3,a

line 3, description, format: a

line 4, spatial dimensions, format: a6,f7.2,a1,f7.2,a8,f7.2,a1,f7.2,a12,i4,a1,i4,a1

line 5, data storage in file, format: a7,i8,a9,i4,a1,i4,a9,f10.4,a11,i4,a1

The header is followed by a grid box identifier (format: a9,i4,a1,i4) and one data line per year of 12 monthly values (format: 12i5; from January to December reading from left to right). The data are stored as integer values multiplied by ten as indicated in the header by the value behind the "Multi=" keyword. The first data line represents data for the first year, the second data line for the second year and so forth.

You will need to download time series climatic data. Presently, we know of two suitable datasets:

The Climatic Research Unit time series is available from the Centre for Environmental Data Archival

(http://badc.nerc.ac.uk/view/badc.nerc.ac.uk_ATOM_dataent_1256223773328276).

The WFDEI (WATCH Forcing Data ERA Interim) dataset (Weedon *et al.*, 2010) has been resampled and reformatted to GRIM format, and is available from the CliMond website, www.climond.org. Both datasets are provided at 30 arc minute spatial resolution.

You will need to download data for precipitation, minimum temperature, maximum temperature, and vapour pressure. Please be aware that these are large files (more than 450 Mb each for the full CRU dataset). You should place all four files in the Dymex/Metdata directory. Having downloaded the data, you need to let CLIMEX know where the data is. You do this by initialising the Meteorological Data reader (Fig. 3-8 and Fig. 3-9). In the initialisation dialogue you need to specify the path to the files. You can do this by clicking on the button with the three dots and browsing to the minimum temperature file. So long as the other files are named correctly, and are located in the same directory, the programme will know where to find them.



Fig. 3-8 Model Components window for the Compare Locations/Years model

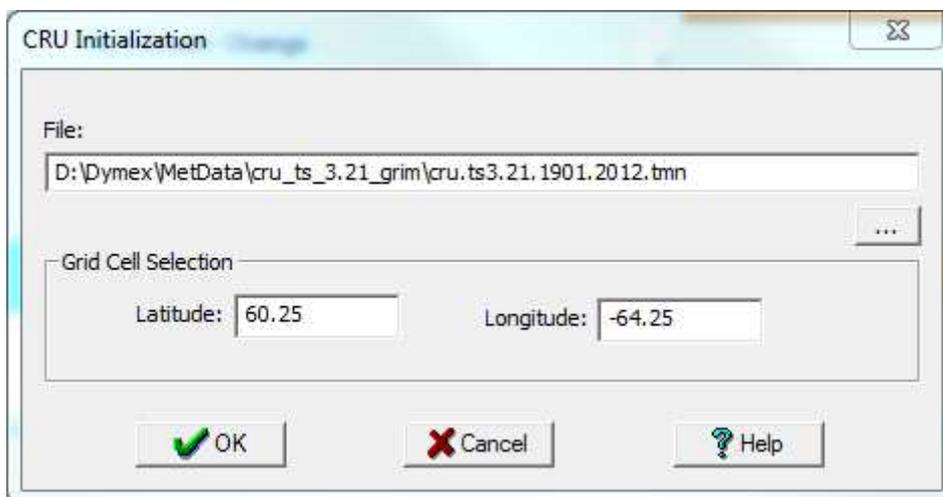


Fig. 3-9 Initialization Window for linking the Time Series data.

You now need to indicate to CLIMEX which set of locations you want to include in the run. This is done by adding a sequence. Either select an existing sequence from the dropdown list, or add a new one using the Add/ edit / remove option (Fig. 3-10). Select New from the next window and then specify a name for the run, the maximum number of weeks to include in the model run and a set of bounding coordinates in decimal degrees. The maximum number of runs is to prevent the system from inadvertently undertaking excessively long simulations.



Fig. 3-10 Model Components window illustrating the addition of a MetData sequence for the Compare Locations/Years model.

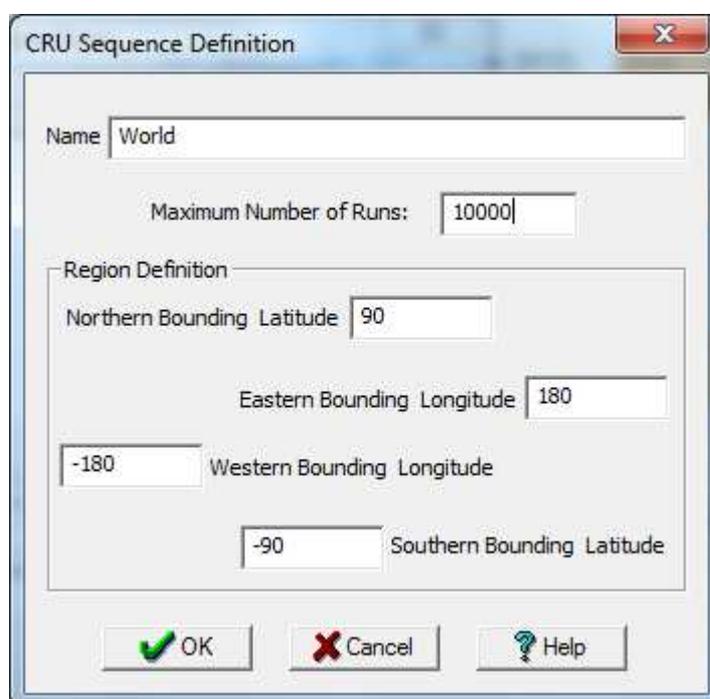


Fig. 3-11 Defining the area for a CRU timeseries run

You can then set any climate change or irrigation scenarios using the normal dialogues. If you want, you can specify a cropping season, which will set a flag that indicates the duration of the cropping season. On its own, this variable is unremarkable. Advanced DYMEX users who are comfortable with editing gmd files can use this variable to limit the Growth Index. As with the Compare Locations model, you need to select a species or template whose climate suitability you want to simulate. Finally, you need to initialise the simulation, specifying the start date for the simulation and either the end date or length of the run.

When you have fully initialised the model you can run it. The Compare Locations/Years model allows you to output the same types of data products as Compare Locations and Compare Years, but additionally allows you to output Map Sequences (videos) of climate suitability. There are two types of Map Sequences that you can map – Step or Yearly Snapshot. Step sequences provide a weekly picture

for the selected variable. To create a new Map Sequence format, click on the Movie icon button  and select New and either Step or Yearly Snapshot (Fig. 3-12).

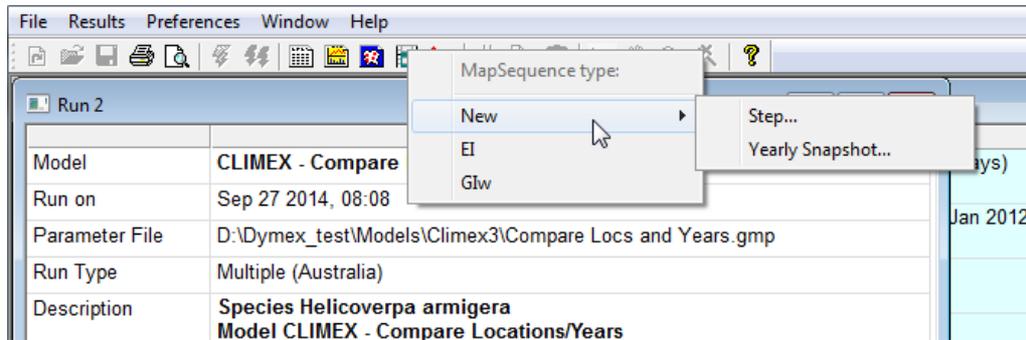


Fig. 3-12 Selecting the type of map to create: seasonal (Step) or annual summary (Yearly Snapshot).

The variables that are available for display differ depending on this choice. A much broader set of variables can be displayed with the Step option, including the underlying meteorological variables. The Yearly Snapshot option includes only variables that are integrated at the annual timestep. For example, the Ecoclimatic Index and Number of Generations variables only have meaning within the Annual Timestep as they are calculated annually. For each of these variable types you specify the video format in exactly the same way that you specify a normal map format (Fig. 3-13).

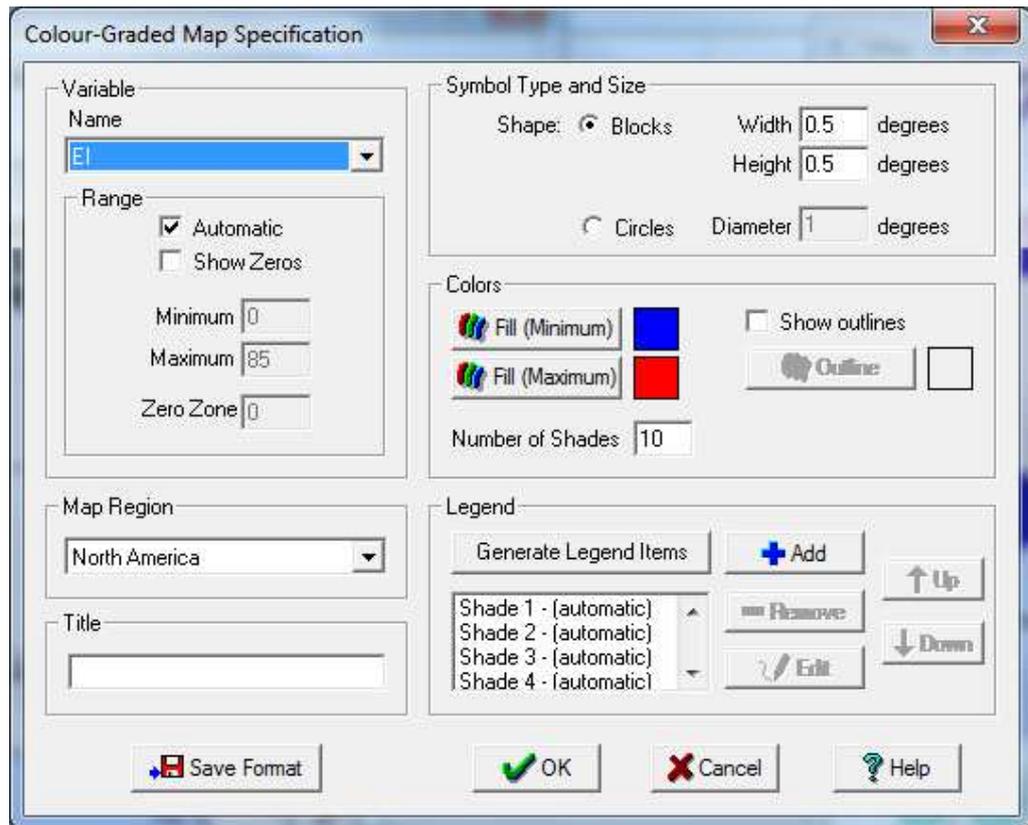


Fig. 3-13 Map Specification window for time-series maps.

The Map Sequence window is similar to the normal map output, but includes a series of controls to control the actions of the video viewer (rewind, go backward one step, play, go forward one step). You can also skip to anywhere in the time sequence using the slider bar, and you can set the time delay between frames while playing the sequence automatically (Fig. 3-14).

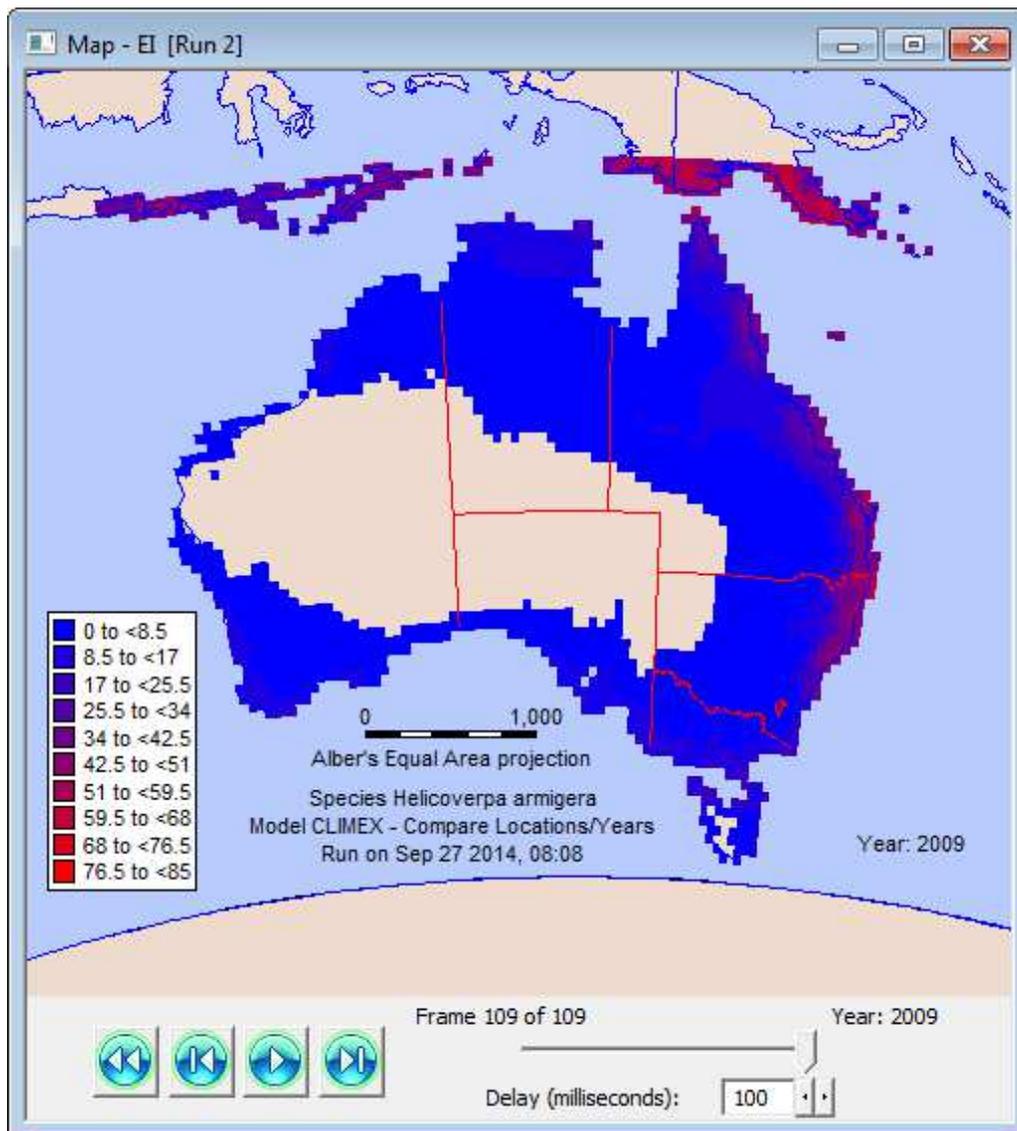


Fig. 3-14 The movie control for displaying spatio-temporal suitability maps.

! The time sequences can take a long time to create and refresh if you have large spatial extents.

In addition to the normal map controls, the Map Sequences have the option of including the time step in the video frame (Fig. 3-15). They also have the option to output the sequence as a video.

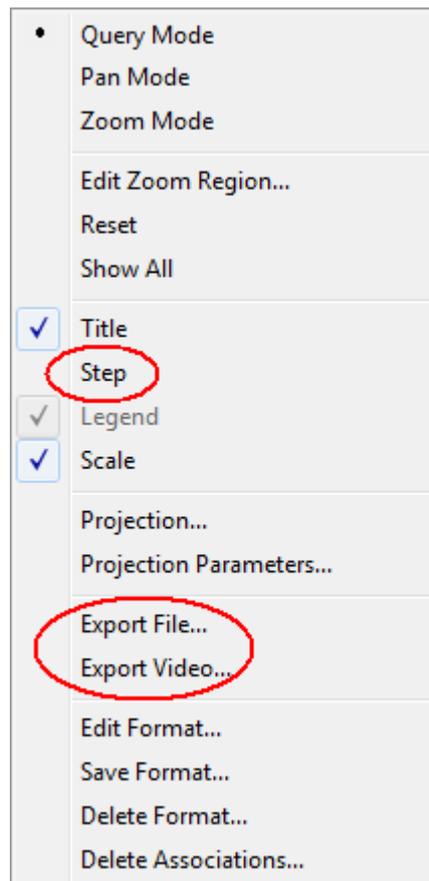


Fig. 3-15 Selecting whether to include the timestep in the video, and the controls for exporting the raw data file or the video.

3.9.1 Outputting the Map Sequence as a video

The map sequence can be saved out in avi format. Firstly, format your map as you want it to display. You can pan, zoom or resize the display to include area of the map that you want to display. You can also include or exclude information such as the timestep, scale bar etc. (Fig. 3-13). You should also decide on the time delay between refreshing the maps. When you are ready to save the file select Export Video. You first have to specify the location where you will save the file and give it a name. You then need to specify the video compression details (Fig. 3-16). The list of software compressors will depend on your specific computer (Fig. 3-16). The Cinepak Codec has been found to work well, though you may wish to experiment with your own system. The resulting avi file can be used to share your results.

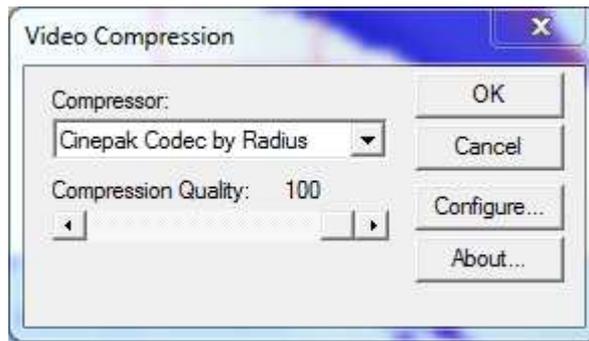


Fig. 3-16 Selecting the video compression method

! The time series of monthly climate patterns are far less coherent than long-term (e.g., 30-year) averages. Hence, for a given level of accuracy, a denser network of stations is required for monthly estimates than climatologies (New and McSweeney, 2010). This variability is particularly pronounced for precipitation. Hence, we should expect that gridded monthly time series data will be less accurate than climatological grids.

! The output from the Compare Locations/Years runs can be saved as a NetCDF or csv file. In order to generate the required files to save out in NetCDF format it is necessary to turn on the appropriate output streaming. This option is accessed using the Preferences|Model Options window. To enable the saving of the detailed time sequence data it is necessary to select the “Retain results for map sequences” option. Selecting this option increases the run time considerably as results are written into a NetCDF file. This option should generally **not** be selected until you have fitted your model fairly well as it slows down the model considerably.

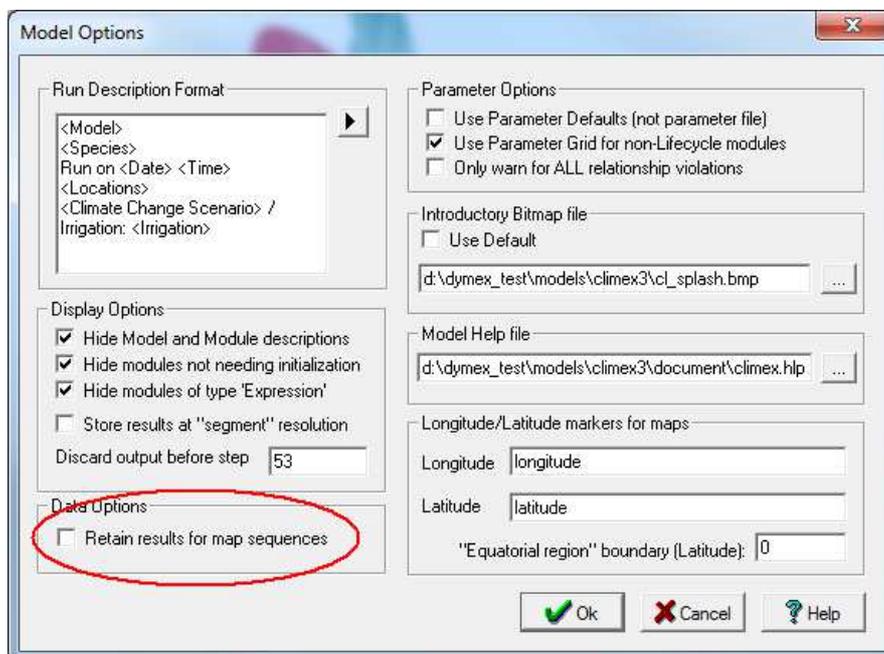


Figure 3-17 Retaining intermediate results for building map sequences

4 Climate Matching

Climate matching consists of selecting a location and then looking for locations elsewhere that have a climate that is similar. This is a much less ‘intelligent’ function than the CLIMEX model. It simply enables the user to compare the meteorological data from different places, with no reference to the preferences of a given species. It answers questions such as: Are the extreme minimum and maximum temperatures similar? Do the two locations have similar amounts of rainfall and similar seasonal rainfall patterns?

A comparison of climates in different locations does however have implications for species. If the user knows that a given species occurs in given locations in Mexico, and is looking to import that species to Australia as a biological control agent, a Match Climates analysis can show whether the climate of the destination is similar to that of the area of collection of the species. If the analysis indicates that there is only a small area in Australia that has a similar climate to the Mexican locations that support populations of the species, there are high risks of failure of establishment by the species in Australia. Conversely, the Match Climates function can be used to indicate where to focus the search for a biological control agent (Dhileepan *et al.*, 2006; Robertson *et al.*, 2008).

Another use of the Match Climates function is to provide a rough assessment of the risk of a pest establishing in a new location (country, continent), with implications for quarantine and pest risk analysis. The Match Climates analysis can be conducted with no knowledge of the species, except that it does occur in certain locations. Thus, if there are no biological data or distribution map of the species, the Match Climates analysis nonetheless enables the user to assess the risk qualitatively.

4.1 Match Climates

The Match Climates application (Fig. 4-1) allows the user to select a location (the “Home” location) and find locations within a specified Location Selection (the “Away” locations) whose climate is similar to that of the “Home” location. The measure that is used to describe the similarity between a pair of locations is the **Composite Match Index**. Its formulation is described in Section 7.



Fig. 4-1 The Match Climates application.

The “Home” location is specified by clicking on the Location (Home) icon and selecting “Initialize (Single Location)” from the resulting menu. This opens the

MetManager Location Selection dialog (Fig. 4-2), showing all the locations in the currently loaded Meteorological Database. The desired “Home” location can be selected from this list.

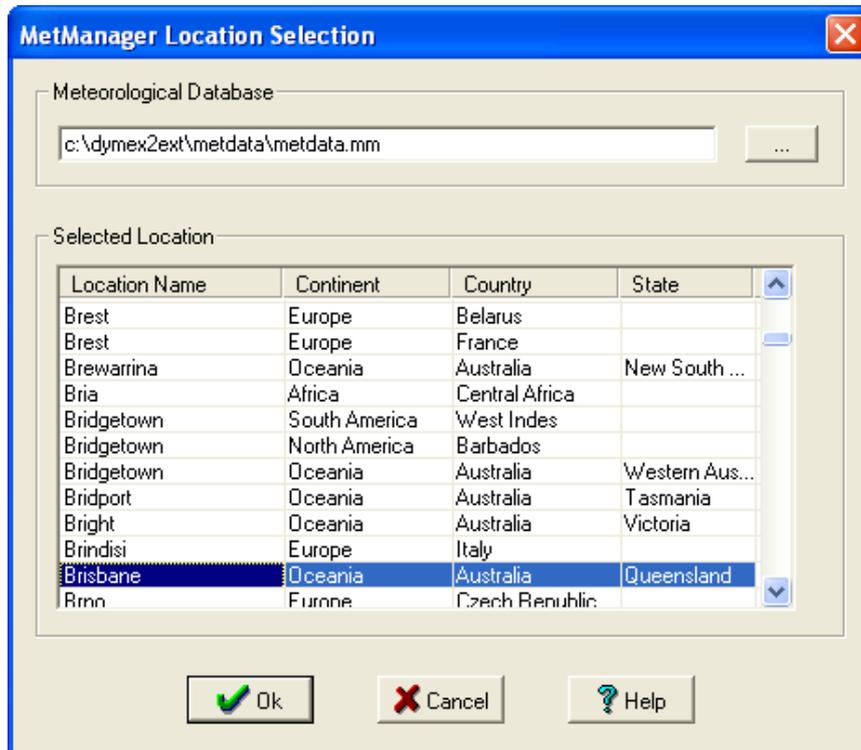


Fig. 4-2 The MetManager Location Selection dialog, showing Brisbane selected as the ‘Home’ location.

The “Away” Location Selection for which the application is to be run is specified in the **Locations (Away)** panel (see Section 8.6 for details on how to do this). Optionally, a climate change scenario (Section 8.4) and/or irrigation regime (Section 8.5) can be set for either the “Home” or “Away” locations from the corresponding panels.

The Match Climates panel allows some match properties to be set via the **Match Climates Setting** dialog (Fig. 4-3). In particular, users can set different weights for the components of the **Composite Match Index** (see Section 7) and the period for which the comparison is made.

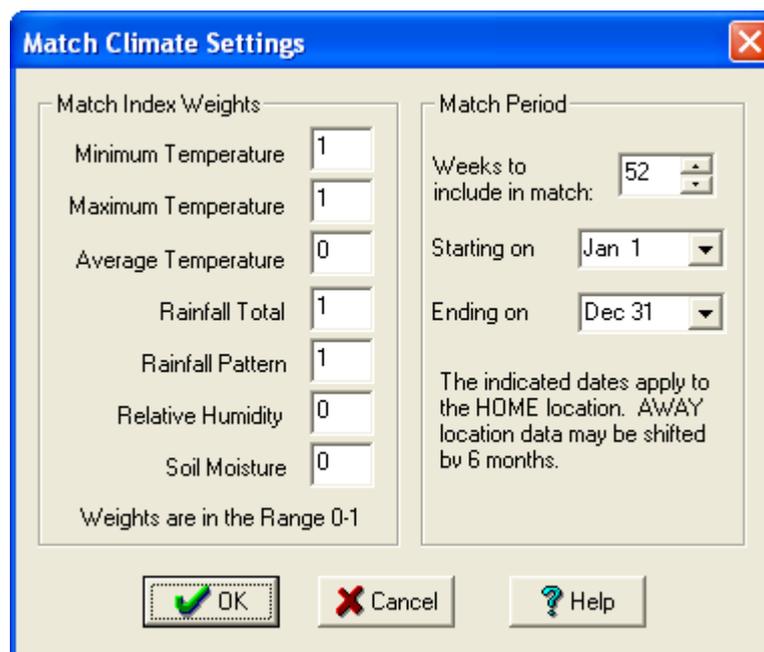


Fig. 4-3 The Match Climate Settings dialog.

4.2 Match Climates (Regional)

The “Match Climates (Regional)” application is exactly the same as the Match Climates application, with the exception that a Location Selection (i.e., a set of locations) rather than a single location is specified as the “Home” location. The application’s appearance in the **Model Components** window is exactly the same as that of the Match Climates application (Fig. 4-1), and match options, climate change scenarios and irrigation can be set in the same way. Section 7.1 gives a detailed description of the use of this application.

5 The Conceptual Basis of the CLIMEX Model

CLIMEX is a simplified dynamic model that enables the user to simulate the climate-related processes that determine the relative abundance, seasonal variation in numbers and geographical distributions of species. It is therefore designed to infer the mechanisms affecting population processes, rather than simply provide an empirical description of the field observations based on matching patterns of meteorological data.

5.1 Species Distribution and Abundance

It has long been recognised that the ultimate limits of the distribution of most plants and poikilothermic animals (those that have a body temperature that fluctuates with the temperature of the environment) are determined by climate (Andrewartha and Birch, 1954; Krebs, 1978; Woodward, 1987). Of course, climate is not the only factor limiting species geographical distribution (Fig. 3 1, Brown *et al.*, 1996). Competition, including predation, and habitats (such as host availability, soil type and more recently the impact of humans) are often also important. However, climate is the only factor that is readily tractable to provide an initial and conservative estimate of the potential range of the species. All of these modifiers must be considered when making assessments of CLIMEX results and a **Report Template** is provided below to prompt CLIMEX users to address each of those issues when fitting CLIMEX parameter values. Detailed studies will be needed to define the roles of the non-climatic factors.

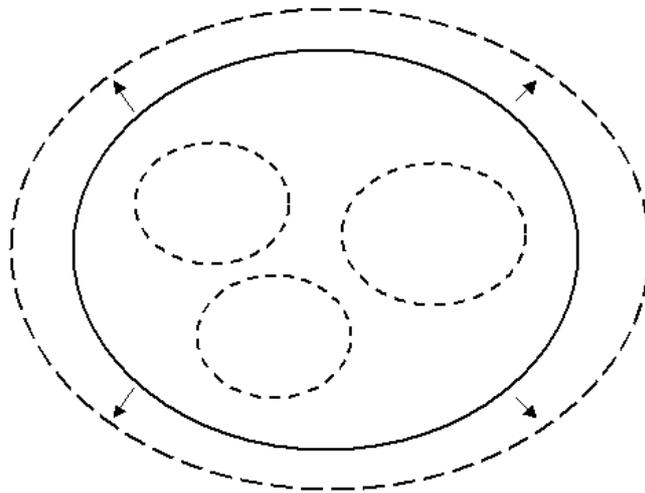


Fig. 5-1 The potential range of a species in relation to climate as estimated from the observed geographical distribution (—) using CLIMEX. This equates to the 'realized' or 'ecological' niche of a species. The effect of removal of competition is to expand the range (---) to fill the 'fundamental' or 'physiological' niche. Within these zones, the availability of suitable habitats (---) will further restrict the range.

CLIMEX Model Theory

Template for CLIMEX Species Analysis

Name:
Biology:
Distribution: CABI CPC Reference, etc.
Comments and literature review related to fitting CLIMEX parameters
Distribution records used to estimate CLIMEX parameter values
Evidence for non-climatic limits to distribution Physical barriers Hosts Vectors Other species (competition, predation) Artificial environments (e.g. irrigation or glasshouses)
Stress indices Hot Cold Dry Wet
Climatic Constraints Obligate diapause Length of growing season
Growth Temperature Moisture
Results Goodness of Fit Independent Validation Source Risks Geographical Seasonal Destination Risks Geographical Seasonal
Discussion
References
Species Parameters Table

In the absence of evidence to the contrary, CLIMEX works on the assumption that the distribution of a species in its native range is limited by climate in the context of a background of biotic resistance i.e., it uses the ecological (realised niche) rather than the physiological (fundamental niche) limits of a species to infer its responses to climate. When fitted to species native range data, CLIMEX assumes that similar intensities of biotic resistance (not the same species) will occur in new environments. This is a gross assumption, but specific information is rarely obtainable before a species introduction, unless a specific species interaction has occurred elsewhere. When fitted to distribution data that includes a species invaded range there is an opportunity that the species has been released from the effects of its natural enemies and it may be occupying climates that it would not be able to when in the presence of its natural enemies. It is possible to use CLIMEX to explore these issues by comparing models fitted to the native range only data and to the combined native and invaded ranges (Wharton and Kriticos, 2004). As the realised range will almost always be less than the range set by physiological limits, the results are conservative. Thus CLIMEX constitutes a pragmatic tool that needs to be handled accordingly to provide guidance without any pretence of unreachable precision.

CLIMEX also assumes that most animal and plant populations experience one or more periods of the year that are favourable for population growth, and some periods that are unfavourable and may prevent the species' persistence at a given location. In areas with bimodal rainfall for example the limiting season is less evident, and CLIMEX results will need to be examined in more detail to interpret the situation.

Note that if physiological data are available for a species, the CLIMEX parameters can be estimated to describe the fundamental or physiological niche of the species. Such data are rarely available, however. It may also be possible to infer some of these parameters from observations of the growth phenology of the species at sites where for at least one climatic factor (temperature or soil moisture), conditions are sub- or supra-optimal for at least part of the year (de Villiers *et al.*, 2012). By examining the phenological patterns and relating these to the climatic conditions at the time of the onset and cessation of growth, or optimal growth rates, the appropriate parameter values can be set. This method should be undertaken cautiously, bearing in mind the potential for inter-annual variation in weather patterns to affect a species phenology strongly.

CLIMEX contains a meteorological database of ~ 3 000 station locations worldwide and quickly interfaces the climatic preferences of a species with the database. Version 4 also comes loaded with a 10' and a 30' gridded global terrestrial climatology centred on 1975, and a future climate scenario for 2080. To estimate the climatic suitability of locations for a species, CLIMEX uses a meteorological database to calculate several indices from its climatic preferences, which are defined in the species parameter file. A description of the theory behind the CLIMEX indices is provided in this guide so that the user can gain a better understanding of the mechanics and reasoning of CLIMEX calculations. Understanding these indices is essential to the use and interpretation of CLIMEX.

A principal aim of the model is to provide a single number (the Ecoclimatic Index, EI) to describe how favourable the climate of a location is for a particular species. This figure can be broken down into component parts for a more detailed examination of the species' response to climate at any given location (Sutherst and Maywald, 1985). In acknowledgement of the semi-quantitative nature of the results, given the type of observations on which the relationships were derived, CLIMEX presents the results on maps in the form of circles that are proportional to the value of the index.

5.2 The CLIMEX Indices

CLIMEX uses several indices and two constraints to estimate the potential growth and survival of a population at a given location. These indices are grouped into growth-related indices and stress-related indices. All of the indices are calculated weekly, and the weekly values are then either used to calculate a derived index (eg Growth) or combined into an annual value of the index. The subscripts 'w' and 'A' are used to distinguish the weekly from the annual values where the distinction is necessary.

Stress Indices relate to factors limiting the geographical distribution. The component indices are:

Cold Stress (CS)

Heat Stress (HS)

CLIMEX Model Theory

Dry Stress (DS)

Wet Stress (WS)

Stress Interactions (Cold/Wet, Cold/Dry, Hot/Wet, Hot/Dry)

CLIMEX combines the growth and stress indices into an Ecoclimatic Index, which describes the climatic suitability of a location for a species as a single number between 1 and 100.

The philosophy in CLIMEX is that 'population growth' occurs within the range of climatic parameters defined by the Growth Indices (eg TI, MI). Stress, leading to negative population growth, is assumed to accumulate outside the range of values suitable for growth and is described by the Stress Indices.

The value of EI is reduced to zero if either of two constraints is not met. These are firstly, a minimum amount of thermal accumulation during the growing season (PDD) and secondly, existence of conditions that enable species with obligate diapause/dormancy (DI) to complete diapause development.

5.3 Growth Related Indices

5.3.1 Growth Index (GI_A and GI_W)

CLIMEX uses a weekly thermo-hydrological Growth Index (TGI_W) (enhanced by radiation, light and/or a user-defined 'substrate' if desired) scaled from 0 – 1 to describe conditions that favour growth of a population. The target organism can be any biological entity such as a species, variety or biotype, but not an aggregate entity such as a community because the components of the latter will each have a unique response. For brevity, the terms 'species' and 'population growth' are used throughout this document to represent these concepts.

The Weekly (Thermo-hydrological) Growth Index,

$$TGI_W = TI_W \times MI_W \times RI_W \times SVI_W \times LI_W \times DI_W$$

where TI_W , MI_W , RI_W , SVI_W , LI_W and DI_W are the weekly Temperature, Moisture, Radiation, Substrate, Light and Diapause Indices respectively. MI_W , TI_W , RI_W and LI_W model the organism's weekly response to moisture, temperature, radiation and light, while DI_W takes into account the effects of periods of dormancy in plants, or diapause in insects, on the ability of an organism to grow or propagate. SVI_W models the weekly response to a specified substrate (such as soil type).

The final weekly Growth Index includes a component, the Biotic Index BI_W , which reflects the effect on population growth of a competing (or synergistic) second species.

The Weekly Growth Index, $GI_W = TGI_W \times BI_W$

BI_W reflects the presence of the second species, either lowering or increasing the potential growth rate of the species under consideration. BI_W is always 1 (and thus has no effect on the Growth Index) when a single species is being modelled. The

indices are explained in more detail in the following sections. The GI_W values are averaged to give an Annual Growth Index (GI_A) scaled between 0 - 100.

$$\text{The Annual Growth Index, } GI_A = 100 \sum_{i=1}^{52} GI_{W_i} / 52$$

The GI_W describes the weekly suitability of the climate for the growth of, for example, an individual plant or a population of insects (e.g. development, reproduction, survival). GI_W is used to infer the period of seasonal growth of the species without reference to life cycle stages or generations, although the latter can be computed if the thermal requirements of the species are known in terms of degree-days per generation. It therefore effectively uses the average response of different stages. The annual GI_A describes the overall potential for growth, and gives an indication of the potential size or relative abundance of a species across its range as determined by climate alone.

When applied to a population and averaged over one generation, the GI_W is analogous to a weekly-scaled 'finite rate of increase' that combines fecundity and mortality rates (Andrewartha and Birch, 1954), as explained by (Sutherst *et al.*, 1995). Both measures suffer from a failure to incorporate the effects of temperature on development rates of multivoltine species (which have more than one generation per year). This deficiency becomes evident when attempting to integrate the values over periods much longer than one generation. Then, multiplicative effects become apparent, and the relationship between the annual total population and GI_A (or finite rate of increase) is not linear for a variety of reasons such as cumulative growth and negative feedback (density dependent) effects.

5.3.2 Temperature Index (TI)

The weekly Temperature Index (TI_W) describes the response of the species to the daily temperature cycle. It varies between 0 and 1. Population growth is maximised when $TI_W = 1$, and is zero when $TI_W = 0$.

Four parameters define the range of suitability for temperature, which are used to calculate TI_W . The four parameters used to define the range of temperatures suitable for population growth (see Fig. 5-2) are:

DV0 = the lower temperature threshold

DV1= the lower optimum temperature

DV2= the upper optimum temperature

DV3=the upper temperature threshold

CLIMEX Model Theory

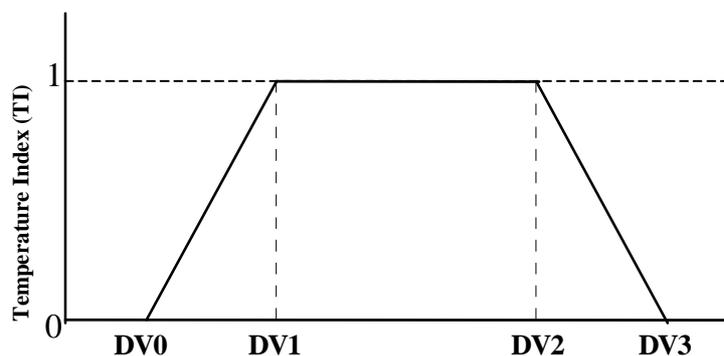


Fig. 5-2 The Temperature Index as a function of population growth. $TI_w = 1$ when the range of weekly temperatures is between the parameters DV1 and DV2, and $TI_w = 0$ when the temperature is below DV0 or above DV3.

A fifth parameter, PDD, denotes the number of degree-days above DV0 (the lower temperature threshold for population growth) required to complete an entire generation. It is used to calculate the potential number of generations per year and also acts as a limiting condition when a minimum of one generation has to be completed for the species to exist in a given location. If the number of degree-days set for PDD is not reached, the species is unable to complete a generation. In this case, the EI is set to 0 and the species is deemed not to be able to persist at that location.

CLIMEX uses the degree-day concept when calculating the Temperature Index (TI) and the number of generations per year (from PDD). The value of the temperature parameters DV0 - DV3 for a species describe the temperature range in which it can live. In a 24-hour period, the total degree-days for population growth is equivalent to Q, the shaded area in Fig. 5-3, i.e. the area under the temperature curve above DV0 but below DV3. To calculate this area, CLIMEX uses the algorithm published by Baskerville and Emin (1969)

Annual degree-day totals above DV0 are used to calculate the number of generations per year if a value greater than zero has been entered for PDD in the parameter file. For example, if a species requires 660 degree-days above DV0 but below DV3 to complete its life cycle ($PDD=660$) and this total number at a particular location is 4300, the species can complete $4300/660 = 6.5$ generations in a year.

The weekly Temperature Index (TI_w) is calculated from two components. The first of these (I_q) describes the effect of temperature on population growth rates at optimal temperatures and below, while the second (I_h) describes the reduction in population growth rate of a species at high temperatures. The value of I_q is calculated from the average daily maximum and minimum temperatures.

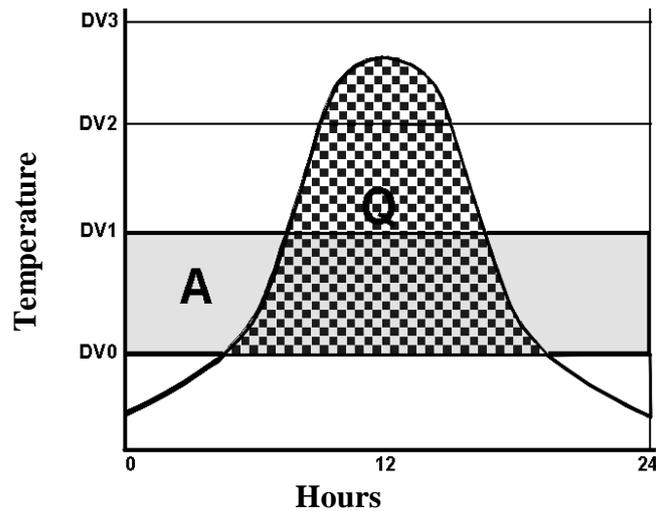


Fig. 5-3 The Temperature Index in relation to the daily temperature cycle. Q = shaded area under curve between DV0 and DV3, and A = shaded area between DV0 and DV1.

The Weekly Temperature Index, $TI_w = I_q \times I_h$

if $Q \leq A$, then $I_q = Q/A$ (fig. 3, left)

if $Q > A$, then $I_q = 1$ (fig. 3, left)

where A is the accumulated degree-days above DV0 assuming a constant temperature of DV1 (see Fig. 5-4).

The second temperature component, I_h , is defined as

if $T_{max} < DV2$, then $I_h = 1$,

if $T_{max} < DV3$ and $> DV2$, then $I_h = 1 - (T_{max} - DV2) / (DV3 - DV2)$,

if $T_{max} \geq DV3$, then $I_h = 0$,

Note that I_h reduces with increasing maximum temperature above DV2 (Fig. 5-4, right).

CLIMEX Model Theory

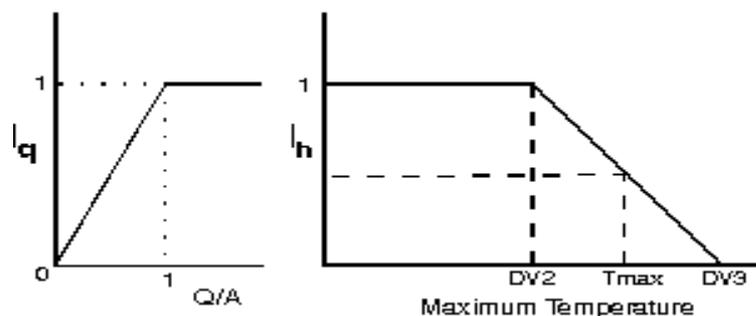


Fig. 5-4 The Temperature Index as a function of mean, maximum and minimum daily temperatures.

The Annual Temperature Index, $TI_A = 100 \sum_{i=1}^{52} TI_{w_i} / 52$ is quoted when querying results on Maps.

5.3.3 Moisture Index (MI)

The Moisture Index is based on the assumption that soil moisture is the dominant factor determining the moisture content of vegetation and hence the microclimatic conditions. The weekly soil moisture balance is calculated with a hydrological model from the stored soil moisture from the previous week and the current week's rainfall and evapotranspiration (moisture loss from plants and soil surface).

Evapotranspiration is calculated using a scheme based on Fitzpatrick (1963) and Sands and Hughes (1976), based in turn on the Tetens equation (Tetens, 1930).

SM = 0 indicates no soil moisture; SM = 0.5 indicates soil moisture content is 50% of capacity; SM = 1 indicates that the soil moisture content is 100% of capacity; SM > 1 indicates a water content greater than the soil holding capacity of (S_{max}) ie run-off.

The user can adjust two parameters to affect soil moisture calculations in CLIMEX. The first of these is soil moisture storage capacity. The default capacity is set to 100 mm (which represents a nominal global average value), but in reality this will vary with soil types. The 100 mm moisture storage value represents a soil profile with a depth of approximately 500 mm and a moderate water holding capacity. Sandy soils may only hold 50 mm while clays can hold 150 mm within reach of the rooting zone of typical crop plants. The second parameter is the evapotranspiration coefficient, which is the amount of moisture lost from a vegetated soil relative to rates of loss from a standard evaporation pan. The default value is 0.8 (Fitzpatrick and Nix, 1969), which is typical for mature pasture/crop areas. The value will reduce with a lower leaf area. Both parameters can be adjusted for individual species, to reflect rooting depth or morphology.

$$\text{Soil Moisture (SM)} = \frac{S_{i-1} + P_i - E_i}{S_{max}}$$

Where, S_{i-1} = the estimated available water from prior week,

P_i = the rainfall from the current week,

E_i = the evapotranspiration from the current week, and

S_{max} (100 by default) = the assumed maximum available soil water storage (Fitzpatrick and Nix, 1969).

The Moisture Index (MI) describes the response of the species to the soil moisture values (Fig. 5-5). The parameters SM0, SM1, SM2 and SM3 have the same form as the temperature parameters, except that moisture is assumed to be constant over the 24-hour period.

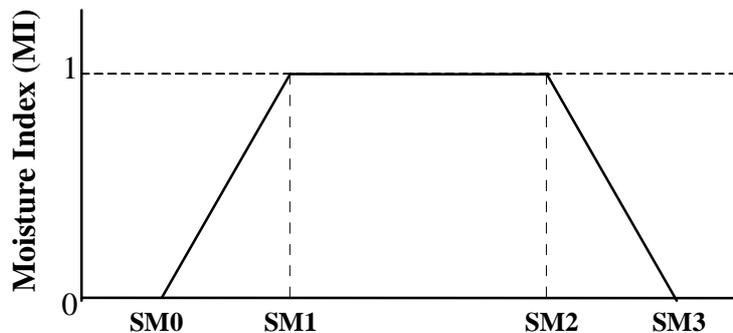


Fig. 5-5 The Moisture Index as a function of available soil moisture.

The four parameters used in calculating Soil Moisture to determine population growth include:

SM0 = lower soil moisture threshold

SM1 = lower optimal soil moisture

SM2 = upper optimal soil moisture

SM3 = upper soil moisture threshold

The Moisture Index ranges between 0 and 1. When MI = 1, population growth is maximised. This occurs when the range of soil moisture is between SM1 and SM2. At MI = 0, population growth is zero, and this occurs when the soil moisture is below SM0 or above SM3.

For two competing species sharing the same space, such as a set of competing plant species, the presence of the other species can reduce the available moisture for the first species. This is modelled by a reducing the moisture available to each species as a function of the Growth Index of the competing species (Fig. 5-6).

CLIMEX Model Theory

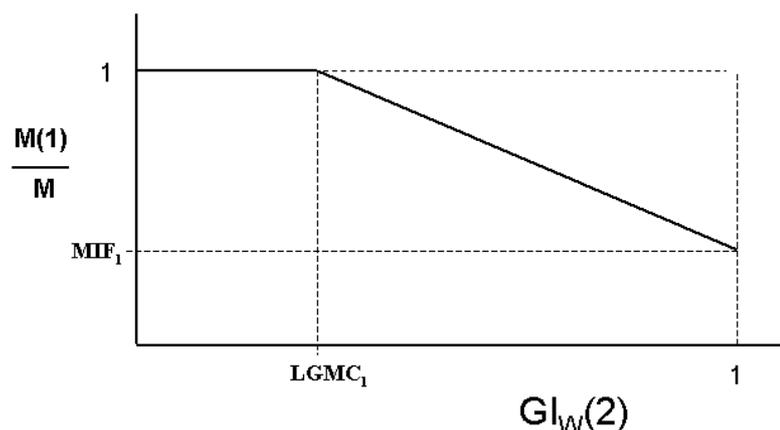


Fig. 5-6 The Moisture Index as reduced by the presence of a competing species.

Two parameters are used to specify the interaction. If the two species are designated as Species 1 and Species 2, Figure 3.6 illustrates the relationship. As the Growth Index of Species 2, $G_{i_w}(2)$, increases above the parameter $LGMC_1$, the available moisture for Species 1, $M(1)$, is reduced from the amount (M) simulated by the Soil Moisture Model. The second parameter (MIF_1) specifies the maximum reduction in available moisture when Species 2 is achieving its maximum growth rate. The subscript (i) indicates that these are parameters of species 1. Note that setting $LGMC_1$ to 1 specifies that Species 2 does not affect the available moisture for Species 1. A second set of equivalent parameters ($LGMC_2$ and MIF_2) specifies the effect of Species 1 on the available moisture for Species 2.

5.3.4 Radiation Index (RI)

The Radiation Index is useful for defining the growth index for plants, as the amount of available radiation is an important determinant of their potential growth. Clear sky global radiation reaching the surface of the globe at any point is calculated from orbital parameters. The actual radiation on a horizontal surface (in MJ per sq. metre per day) at any point is then calculated in one of two ways. If sunshine hours are available in the meteorological data, this is used; otherwise a calculation is made using the temperature range at a location.

The Radiation Index (RI) describes the response of the species to the soil moisture values (Fig. 5-7). The parameters RV_0 , RV_1 , RV_2 and RV_3 have the same form as the moisture parameters.

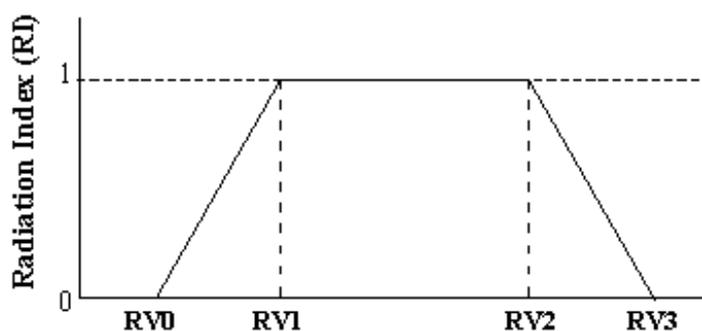


Fig. 5-7 The Radiation Index as a function of terrestrial radiation.

The four parameters used in calculating Soil Moisture to determine population growth include:

RV0 = lower radiation threshold

RV1 = lower optimal radiation

RV2 = upper optimal radiation

RV3 = upper radiation threshold

The Radiation Index ranges between 0 to 1. When RI = 1, population growth is maximised. This occurs when the range of radiation is between RV1 and RV2. At RI = 0, population growth is zero, and this occurs when the radiation is below RV0 or above RV3.

When two species are present in the same place, one may shade the other species and reduce the radiation available to it. This is modelled by a reducing the radiation available to each species as a function of the Growth Index of the competing species (Fig. 5-6).

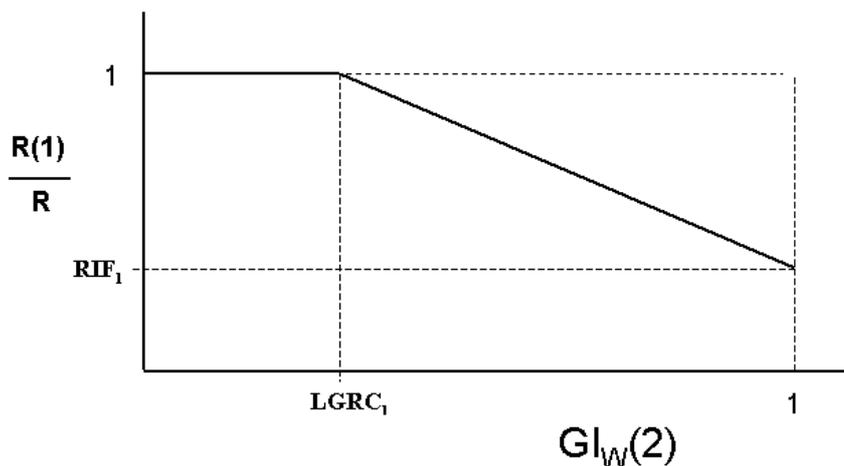


Fig. 5-8 The Radiation Index as reduced by the presence of a competing species.

Two parameters are used to specify the interaction. If the two species are designated as Species 1 and Species 2, Figure 3.8 illustrates the relationship. As the Growth Index of Species 2, $G_{iw}(2)$, increases above the parameter $LGRC_1$, the available radiation for Species 1, $R(1)$, is reduced from the amount (R) calculated by the Radiation Model. The second parameter (RIF_1) specifies the maximum reduction in available radiation when Species 2 is achieving its maximum growth rate. The subscript (i) indicates that these are parameters of species 1. Note that setting $LGRC_1$ to 1 specifies that Species 2 does not affect the available radiation for Species 1. A second set of equivalent parameters ($LGRC_2$ and RIF_2) specifies the effect of Species 1 on the available radiation for Species 2.

CLIMEX Model Theory

5.3.5 Substrate Index (SVI)

The Substrate Index assumes that there is a non-climatic variable (the substrate) that can affect population numbers. Examples of substrates are soil type, topography or host availability or suitability for a parasitic organism.

SVI = 0 indicates that the substrate is completely limiting (i.e., no population growth is possible during that week), while SVI = 1 indicates a substrate that is completely non-limiting to population growth for that week.

The Substrate Index (SVI) describes the response of the species to the substrate variable (Fig. 5-9). The parameters SV0, SV1, SV2 and SV3 have the same form as the temperature parameters, except that the substrate variable is assumed to be constant over the 24-hour period.

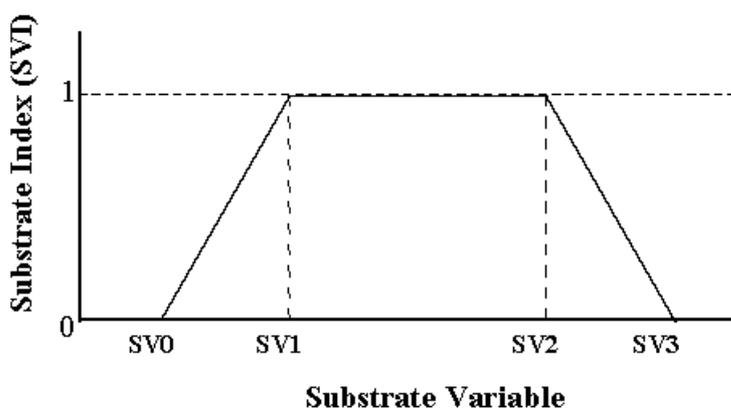


Fig. 5-9 The Substrate Index as a function of the substrate variable.

The four parameters used in calculating Substrate index to determine population growth include:

SV0 = lower substrate variable threshold

SV1 = lower optimal substrate variable value

SV2 = upper optimal substrate variable value

SV3 = upper substrate variable threshold

The Substrate Index ranges between 0 and 1. When SVI = 1, population growth is maximised. This occurs when the range of the substrate variable is between SV1 and SV2. At SVI = 0, population growth is zero, and this occurs when the value of the substrate variable is below SV0 or above SV3.

Note that CLIMEX allows two separate substrate variables to be used. These are referred to as the Physical and Biotic Substrates, and result in a Physical Substrate Index (SVP1) and a Biotic Substrate Index (SVBI). The parameter mnemonics for the former are SVP0, SVP1, SVP2 and SVP3, while those for the latter are SVB0, SVB1, SVB2 and SVB3.

5.3.6 Diapause Index (DI)

Diapause (or dormancy) is an adaptation that enables a species to withstand prolonged periods of unfavourable temperature or moisture conditions. During diapause, a species does not experience stress.

If the diapause is obligate, a species must be exposed to specified conditions for a set period of time in order to complete diapause, just as a certain number of degree-days of temperature are needed to complete normal development. In the plant kingdom, the analogous process to diapause is referred to as 'vernalisation', and is needed to stimulate flowering.

CLIMEX reduces the Weekly Growth Index (GI_w) to zero during the period of diapause development and it sets the annual GI_A to zero if the obligate diapause conditions are not satisfied. This has the effect of setting $EI=0$ and the species is excluded from that location. While a species is not experiencing diapause, the weekly Diapause Index (DI_w) =1. When the species is in diapause, $DI_w = 0$.

In CLIMEX, diapause can be triggered by short daylengths and low temperatures (winter diapause), or long daylengths and high temperatures (summer diapause). Five parameters are used in the diapause calculations.

1. **Diapause Induction Daylength (DPD0)** is the average weekly daylength that induces diapause. Setting DPD0 to 0 disables the diapause function.
2. **Diapause Induction Temperature (DPT0)** is the average weekly minimum temperature (for winter diapause) or maximum temperature (for summer diapause) that induces diapause.
3. **Diapause Termination Temperature (DPT1)** is the average weekly minimum temperature (for winter diapause) or maximum temperature (for summer diapause) that terminates diapause.
4. **Diapause Development Days (DPD)** is the minimum number of days required for diapause development to be complete. If the required number of days in diapause (DPD) is not achieved, the species is assumed not to have survived diapause and the Annual Growth Index (GI) (and hence the EI) becomes zero. The default value for DPD is zero, which corresponds to no diapause.
5. **Diapause Summer or Winter indicator (DPSW)** is set to zero for winter diapause, and to one for summer diapause.

The diapause function works in the following way.

For **Winter diapause (DPSW = 0)**

Diapause is switched **ON** when

- Daylength < DPD0 and
- Daily minimum temperature is < DPT0 and
- Daylength is decreasing (automatically calculated by CLIMEX)

Diapause is switched **OFF** when

CLIMEX Model Theory

- Daily minimum temperature > DPT1 and
- Daylength is increasing (automatically calculated by CLIMEX) and
- Diapause development is complete (at least DPD days have been spent in diapause).

Diapause development is complete when DPD days with a minimum temperature < DPT0 have passed since diapause was triggered. DPD should be set to zero to disable the diapause development condition.

For **Summer diapause (DPSW = 1)**

Diapause is switched **ON** when

- Daylength > DPD0 **and**
- Daily maximum temperature is > DPT0 **and**
- Daylength is increasing (automatically calculated by CLIMEX).

Diapause is switched **OFF** when

- Daily maximum temperature < DPT1 **and**
- Daylength is decreasing (automatically calculated by CLIMEX) **and**
- Diapause development is complete (at least DPD days have been spent in diapause).

Diapause development is considered to be complete when DPD days have occurred with an average maximum temperature > DPT0 since diapause started.

The annual summary variable for Diapause Index indicates the proportion of the year that the organism spends in diapause, expressed as a percentage.

5.3.7 Light Index (LI)

The Light Index enables the growth and development of plants to be described by taking into account seasonal variation in daylength, which affects photosynthesis. The weekly LI (LI_w) has a value between 0 and 1. A value of 0 indicates maximum limiting conditions with inadequate hours of daylight; a value of 1 indicates no limitation to growth from light.

Two parameters determine the value of the weekly Light Index (LI). LT0 is the daylength (in hours) above which growth is at a maximum, while LT1 is the daylength below which growth is 0 (Fig. 5-10). LI is used in the calculation of the Weekly Growth Index (GI_w). The default value for the LI is one, as this has no effect on the GI_w calculations.

if D (daylength) ≥ LT0, then LI = 1,

if LT1 < D < LT0, then LI = 1 - (D - LT1)/(LT0-LT1),

if $D \leq LT_1$, then $LI = 0$.

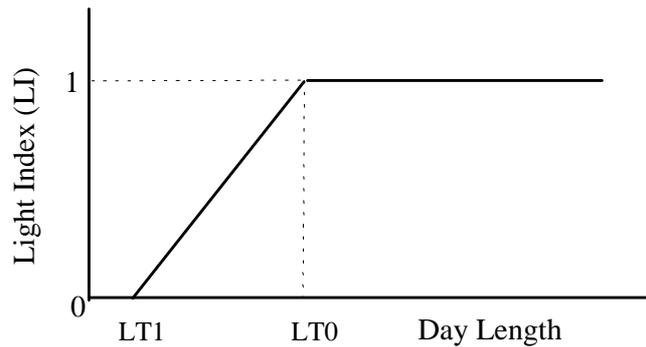


Fig. 5-10 The Light Index function calculates the weekly LI from the day length and the value of two parameters LT_0 and LT_1 .

5.3.8 Biotic Index (BI)

If two species are being modelled simultaneously, CLIMEX allows the effects of competition or synergy between the species to be modelled using the Biotic Index (BI). The Biotic Index is calculated weekly for each species according to the following equation:

$$BI_w(1) = 1 + SIP_{01} * TGI_w(2)$$

$$BI_w(2) = 1 + SIP_{02} * TGI_w(1)$$

where, BI_w are the Biotic Indices for the first (1) and second (2) species, respectively. SIP_{01} is a parameter that specifies the intensity and direction of the effect of species 2 on species 1. The subscript (1) indicates that it is a parameter of species 1. When the parameter is 0, species 2 does not have any effect on species 1, when it is less than 0, species 2 **reduces** the weekly growth rate of species 1 while when it is greater than 0, species 2 **enhances** the weekly growth rate of species 2. Note that in the latter case, BI_w has a value that is greater than 1. In that way, this index is different to the other Growth Index components. SIP_{02} works in the same way as SIP_{01} , but this time giving the effect of species 1 on species 2 (it is a species 2 parameter).

5.4 Stress Indices

The stress indices in CLIMEX are set to limit the species' ability to survive during adverse seasonal conditions, and so determine its geographical distribution. Four of these indices represent the extremes of temperature and moisture. The philosophy used in fitting stress indices is that population growth occurs between the temperature parameters DV_0 and DV_3 , and between the moisture parameters SM_0 and SM_3 . Stress, leading to negative population growth, is assumed to accumulate outside these values. So for example, Dry Stress can only begin to accumulate once soil moisture drops below SM_0 . An annual stress value of 100 is lethal, and precludes a species from persisting in a given location.

While, we assume that there is only one period of each type of stress each year, in practice, exceptions do occur. For example, annual plants are frequently not limited by an unfavourable growth period because they are in the seed stage. Note that there are small areas of the world - in East Africa for example - where there are two

CLIMEX Model Theory

seasons of favourable and unfavourable conditions each year associated with bimodal rainfall distributions. CLIMEX can describe the favourable conditions in each of the two seasons. However, as stresses are accumulated annually, care is needed in interpreting each period of stress, and the annual stress index, which is reset to zero in the favourable period between the two separate periods of stress.

There are four primary and commonly used stress indices in CLIMEX: Heat Stress, Cold Stress, Wet Stress and Dry Stress. There are also four less commonly used stress indices, which are interactions of the four primary indices: Hot-Wet Stress, Hot-Dry Stress, Cold-Wet Stress and Cold-Dry Stress. A threshold value and an accumulation rate define all stresses. In all cases, accumulation of stress during the year is exponential: each week's stress is multiplied by the number of weeks since the stress first exceeded zero.

Each stress index in CLIMEX is associated with a stress accumulation rate. The rate parameter determines how quickly the species accumulates stress when climatic conditions exceed the stress threshold. When the accumulated weekly stress index equals 1.0, the species is unable to persist at the location. For example, the Cold Stress Temperature Threshold (TTCS) defines a temperature below which Cold Stress begins to accumulate. The rate at which this stress accumulates is determined by the Cold Stress Temperature Rate (THCS). If we were modelling a frost-susceptible plant that dies when exposed to 2 °C daily for a week, stress has to accumulate to equal 1.0 or more in that week to simulate this effect. For another species that is more tolerant of cold, it may require two months before the accumulated Cold Stress is fatal. In this case a lower rate of stress accumulation would be more appropriate.

In CLIMEX, it is assumed that stress accumulates exponentially over time, to reflect the progressive increase in severity of the same amount of stress stimulus as time goes by and the population exhausts its resources. Let us assume that Cold Stress begins to accumulate below 10 °C (TTCS) at a rate of 0.1 (THCS), and that the average weekly minimum temperature is consistently 9°C after week 0. The weekly Cold Stress for consecutive weeks is therefore:

Weekly Stress = Rate × (Threshold – average weekly minimum temperature) °C × week number,

OR (for Cold Stress), THCS × (TTCS - average weekly minimum temperature) × week number

Week	Weekly Stress	Total Accumulated Stress
1	$0.1 \times 1 \times 1 = 0.1$	0.1
2	$0.1 \times 1 \times 2 = 0.2$	0.3 (0.1 + 0.2)
3	$0.1 \times 1 \times 3 = 0.3$	0.6 (0.3 + 0.3)
4	$0.1 \times 1 \times 4 = 0.4$	1.0 (0.6 + 0.4)
5	$0.1 \times 1 \times 5 = 0.5$	1.5 (1.0 + 0.5)

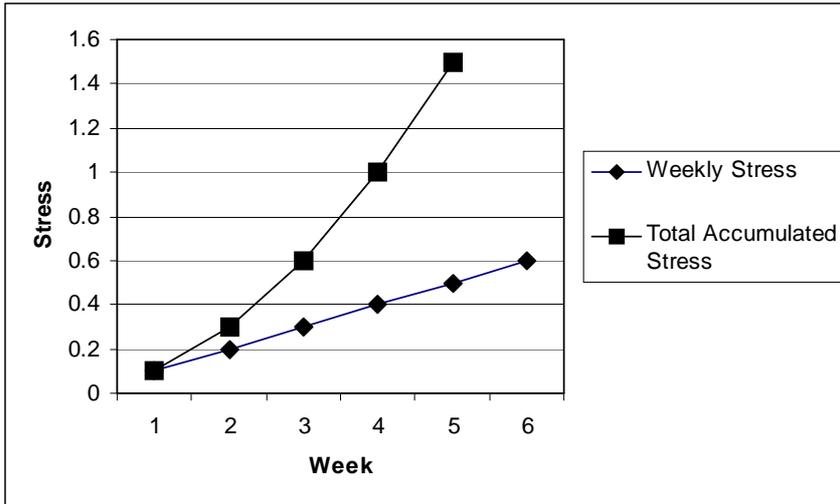


Fig. 5-11 Comparison of weekly and accumulated stress.

The species will therefore accumulate a lethal stress of 1.0 during the 4th week (Fig. 5-11). Note that CLIMEX scales stress indices, so that an accumulated stress of 1.0 or greater becomes an annual stress index value of 100. In calculating the EI value, all Stress Index values are limited to a maximum of 100. However, to help with estimating the parameter values, the actual value of the Stress Index is shown in the tables. For example, if the Stress Index in a marginal zone is 300 it indicates that the parameters are too severe.

The yearly stress index for a variable V (eg Cold Stress) = $100 \sum_{i=1}^{52} V_i / 52$

5.4.1 Cold Stress (CS)

Cold Stress can be defined in three ways. Firstly, poikilotherms may die because the daily thermal accumulation is too low to maintain metabolism. In reality, this is because the plant cannot photosynthesise (or the animal forage) sufficiently to accumulate the amount of energy required to offset respiration losses. In modelling terms this occurs when a threshold number of degree-days above the developmental temperature threshold (DVCS) are not reached. This threshold parameter is called the Cold Stress Degree-day Threshold (DTCS) and is in units of degree-days. Cold Stress accumulates at a given rate (DHCS) if the weekly number of degree-days above DVCS (DD) does not reach the threshold value (DTCS), as in the following equation:

if $DD < DTCS$, then $CS = (DTCS - DD) \times DHCS$,

if $DD \geq DTCS$, then $CS = 0$,

The fewer the number of degree-days experienced at a location, the more Cold Stress will accumulate. If the weekly number of degree-days exceeds the designated threshold (DTCS), a zero Cold Stress value is returned (Fig. 5-12a).

CLIMEX Model Theory

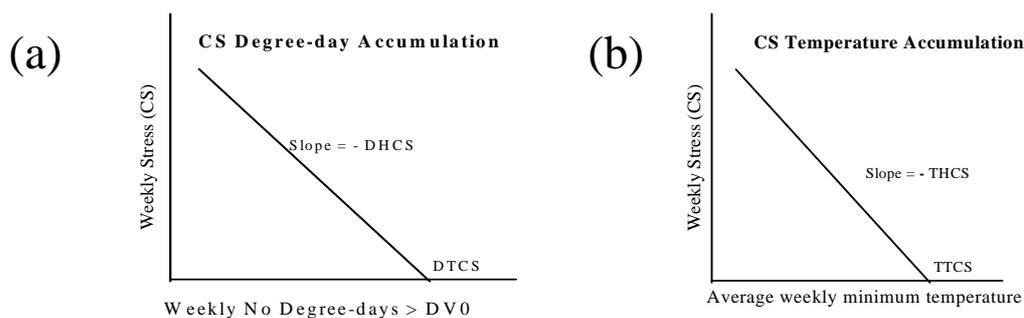


Fig. 5-12 The Cold Stress function. a) Degree-day Threshold (DTCS) below which Cold Stress accumulates at the rate DHCS; b) Temperature Threshold (TTCS) below which CS accumulates at the rate THCS.

Refer to Fig. 5-3, Section 5.3.2, for an explanation of how degree-days are calculated. Cold stress will accumulate if area Q in Fig. 5-3, accumulated over a period of a week, is less than the threshold defined by DTCS. In practical terms, the higher the DTCS value, the more cold stress will accumulate. This is because more warmth (either longer duration or greater temperatures) will be required to prevent cold stress from accumulating. The following graph (Fig. 5-13) indicates the weekly number of degree-days above DVCS (in this case, 14 °C) accumulated at two locations in Australia. With DTCS set at 30, cold stress will accumulate at both locations, and will accumulate for most of the year at Bairnsdale, probably precluding a species from occurring there. If DTCS were reduced to 10, less cold stress would accumulate at both locations. The first reason is that the total number of weeks experiencing fewer than 10 degree-days above DVCS is greatly reduced. The second reason is that the difference between the actual number of degree-days above DVCS experienced and a threshold value of 10 (as compared to 30) is also less, and so when multiplied by the rate of stress accumulation (DHCS), gives a lower cold stress value. Thus, a higher DTCS value results in a longer period of cold stress accumulation and a greater amount of weekly cold stress accumulation.

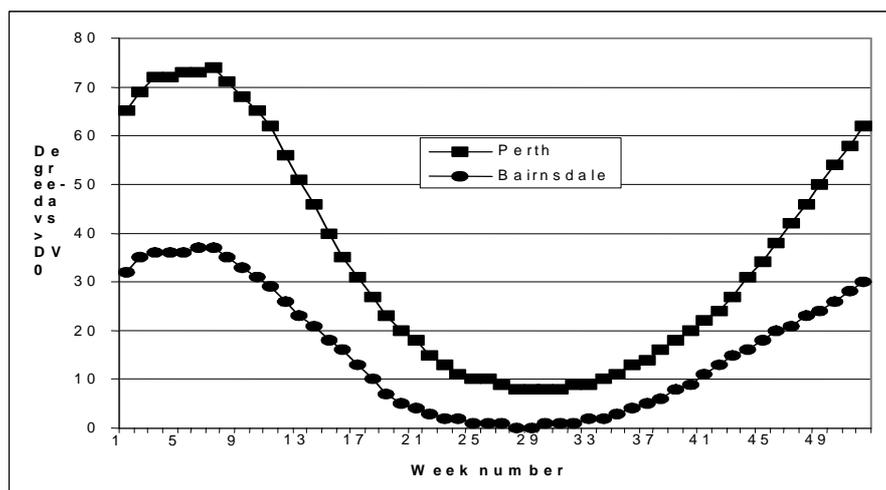


Fig. 5-13 Degree-days > DVCS for Perth and Bairnsdale in Australia each week.

Alternatively, a species may not survive the winter if exposed to excessively low temperatures. The Cold Stress Temperature Threshold (TTCS) represents the average weekly minimum temperature (°C) below which Cold Stress accumulates. Cold Stress accumulates at a given rate (THCS) when the average weekly minimum

temperature (T_{\min}) drops below a given threshold value (TTCS), as in the following equation:

if $T_{\min} < TTCS$, then $CS = (TTCS - T_{\min}) \times THCS$,

if $T_{\min} \geq TTCS$, then $CS = 0$,

If the average weekly minimum temperature is greater than the threshold TTCS, a zero CS value is returned (Fig. 5-12b). As discussed earlier, the annual cumulative Cold Stress value will also depend on the number of weeks experiencing stress.

As a third option, a species may spend the winter in locations where it is buffered from extreme minima and maxima. In such situations the Cold Stress Average Temperature Threshold (TTCSA) represents the weekly average temperature ($^{\circ}C$) below which Cold Stress accumulates. Cold Stress accumulates at a given rate (THCSA) when the weekly average temperature (T_{av}) drops below a given threshold value (TTCSA), as in the following equation:

if $T_{av} < TTCSA$, then $CS = (TTCSA - T_{av}) \times THCSA$,

if $T_{av} \geq TTCSA$, then $CS = 0$,

Again, the annual cumulative Cold Stress value will depend on the number of weeks experiencing stress.

5.4.2 Heat Stress (HS)

It is also possible to model Heat Stress in two ways. The degree-day model of Heat Stress accumulation assumes that species are unable to persist because the daily heat load is too high for essential physiological processes. This model distinguishes between areas that experience brief exposures to extremely high temperatures and areas that experience less extreme maximum temperatures for longer periods each week. The difference between the weekly degree-days above DV3 (DD_{DV3}) and the weekly degree-day threshold (DTHS) is multiplied by the stress accumulation rate (DHHS) to give Heat Stress.

Note: in versions of CLIMEX before Version 2, the degree-day model of Heat Stress calculated degree-days above DV0. Starting with Version 2, CLIMEX uses DV3 as the temperature threshold above which stress is accumulated. Species parameter files constructed using earlier versions of CLIMEX will need to be recalibrated to reflect this change.

if $DD_{DV3} > DTHS$, then $HS = (DD_{DV3} - DTHS) \times DHHS$,

if $DD_{DV3} \leq DTHS$, then $HS = 0$.

If the number of weekly degree-days (DD) above DV3 is less than the threshold (DTHS) value, this equation returns a negative value, which CLIMEX interprets as zero Heat Stress (Fig. 5-14a).

CLIMEX Model Theory

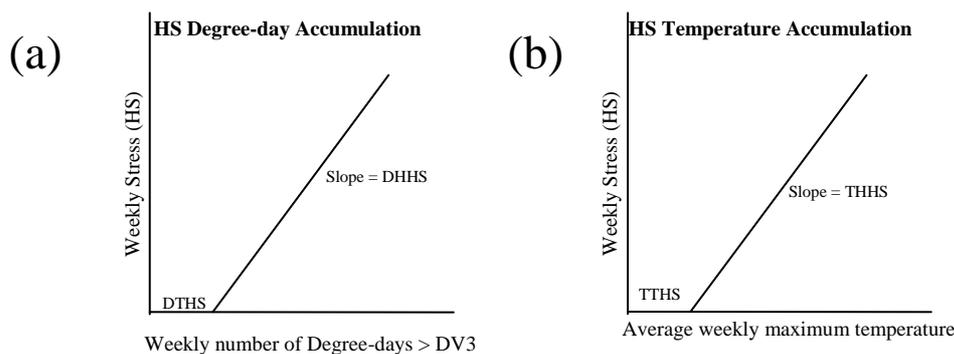


Fig. 5-14 The Heat Stress function. a) Degree-day Threshold (DTHS) above which stress accumulates at the rate DHHS; b) Temperature Threshold (TTHS) above which stress accumulates at the rate THHS.

The second option, the Heat Stress Temperature Threshold (TTHS) model, relates to species that fail to survive when exposed to excessively high temperatures. The threshold temperature model for Heat Stress has two parameters. TTHS is the threshold temperature ($^{\circ}\text{C}$) above which Heat Stress accumulates, and THHS is the rate at which stress accumulates. Weekly Heat Stress is calculated by the following equation:

$$\text{if } T_{\max} > \text{TTHS, then HS} = (T_{\max} - \text{TTHS}) \times \text{THHS,}$$

$$\text{if } T_{\max} \leq \text{TTHS, then HS} = 0.$$

If the average weekly maximum temperature is below the threshold (TTHS), this equation returns a zero Heat Stress (Fig. 5-14b). As discussed earlier, the annual cumulative Heat Stress value will also depend upon the number of consecutive weeks of Heat Stress. The higher the maximum temperature, the greater the Heat Stress in a given week.

5.4.3 Dry Stress (DS)

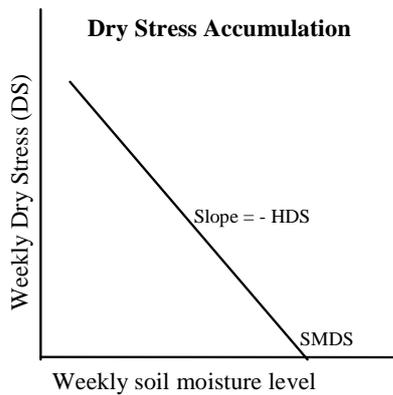
Moisture can cause stress for a species when it is either too dry or too wet. Dry Stress accumulates when the soil moisture level falls below the Dry Stress Threshold (SMDS). The difference between SMDS and the soil moisture level (SM) is multiplied by the Dry Stress Rate (HDS) to give the resultant Dry Stress for the week. Annual Dry Stress will be a cumulative value, as described earlier, and will also depend on the number of weeks experiencing stress.

$$\text{if } \text{SM} < \text{SMDS, then DS} = (\text{SMDS} - \text{SM}) \times \text{HDS,}$$

$$\text{if } \text{SM} \geq \text{SMDS, then DS} = 0.$$

If the weekly soil moisture is greater than the Dry Stress Threshold (SMDS), the equation returns zero Dry Stress (Fig. 5-15).

Fig. 5-15 The Dry Stress Function. Threshold soil moisture (SMDS) below which stress accumulates at the rate HDS.



5.4.4 Wet Stress (WS)

Wet Stress accumulates if the soil moisture level exceeds the Wet Stress Threshold (SMWS). The difference between SMWS and the soil moisture level (SM) is multiplied by the Wet Stress Rate (HWS) to give the resultant Wet Stress. Annual Wet Stress will be a cumulative value as described earlier, and will also depend on the number of consecutive weeks experiencing stress.

if $SM > SMWS$, then $WS = (SM - SMWS) \times HWS$,

if $SM \leq SMWS$, then $WS = 0$.

If the weekly Soil Moisture level is less than the Wet Stress Threshold (SMWS), the equation returns zero Wet Stress (Fig. 5-16).

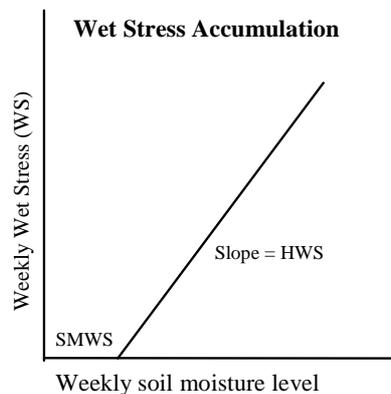


Fig. 5-16 The Wet Stress Function. Threshold soil moisture (SMWS) above which stress accumulates at the rate HWS.

5.5 Interaction Stress Indices

CLIMEX provides for four types of synergistic interactions between temperature and moisture. The interaction indices (CDX, CWX, HDX and HWX) describe the combined effects of two stress factors acting simultaneously, and apply to cold-dry, cold-wet, hot-dry and hot-wet conditions, respectively. These indices are not often needed, but provide extra capability for certain situations in which the simple stress

CLIMEX Model Theory

indices are inadequate. For example, Hot-Wet Stress may be needed to exclude Mediterranean species from tropical areas, or Cold-Wet stress may be needed to exclude tropical species (adapted to cool dry winters) from temperate regions that have cold wet winters.

All these indices are scaled between 0 and 1, with 0 indicating no stress interaction and 1 indicating maximum (lethal) stress.

5.5.1 Cold-Dry Stress (CDX)

DTCD 'Cold-Dry Degree-day Threshold'

MTC D 'Cold-Dry Moisture Threshold'

PCD 'Cold-Dry Stress Accumulation Rate'

Cold-Dry Stress is accumulated if the weekly number of degree-days above DV0 falls below DTCD and the moisture level falls below MTC D. PCD determines the rate of accumulation of stress.

5.5.2 Hot-Dry Stress (HDX)

DTHD 'Hot-Dry Temperature Threshold'

MTHD 'Hot-Dry Moisture Threshold'

PHD 'Hot-Dry Stress Accumulation Rate'

Hot-Dry Stress is accumulated if the weekly maximum temperature exceeds TTHD and the moisture level falls below MTHD. PHD determines the rate of accumulation of this stress.

5.5.3 Cold-Wet Stress (CW X)

DTCW 'Cold-Wet Degree-day Threshold'

MTCW 'Cold-Wet Moisture Threshold'

PCW 'Cold-Wet Stress Accumulation Rate'

Cold-Wet Stress is accumulated if the weekly number of degree-days above DV0 falls below DTCW and the moisture level exceeds MTCW. PCW determines the rate of stress accumulation.

5.5.4 Hot-Wet Stress (HWX)

DTHW 'Hot-Wet Degree-day Threshold'

MTHW 'Hot-Wet Moisture Threshold'

PHW 'Hot-Wet Stress Accumulation Rate'

Hot-Wet Stress is accumulated if the weekly maximum temperature exceeds TTHW and the moisture level exceeds MTHW. PHW determines the rate of accumulation of stress.

5.6 Limiting Conditions

5.6.1 Length of growing season (PDD)

The rigour of an unfavourable season is not the only way in which climate can limit a species' distribution. It may also be necessary to have enough thermal accumulation during the growth season to enable the species to complete a minimum amount of development, which is usually one generation. At high latitudes and altitudes, the length of the growing season often limits the distribution of annual species with one generation per year. The PDD parameter is set to the number of degree-days of thermal accumulation above the development threshold (DVO) required by the species to complete a generation. PDD can be estimated from field or laboratory data.

5.6.2 Obligate Diapause

Diapause (dormancy) is an adaptation that enables a species to withstand prolonged periods of unfavourable temperature or moisture conditions. A species must be exposed to specified conditions for a set period of time in order to complete diapause, just as a certain number of degree-days of temperature are needed to complete development. In the plant kingdom, the analogous process to diapause is referred to as 'vernalisation', and is needed to stimulate flowering.

If a species has diapause development in its lifecycle, its tolerance to the relevant adverse conditions will be taken into account automatically by the CLIMEX stress functions, as these are inferred from the geographical distribution. Hence, there is no need to consider diapause when estimating the four basic CLIMEX stress functions: hot, cold, dry or wet. However, diapause does need to be considered separately when a species' distribution is limited because climatic conditions do not allow diapause requirements to be met. For example, in temperate species, diapause cues involve exposure to certain low temperatures and short daylengths for a minimum period of time. Thus, if a temperate species with obligate diapause is placed in a sub-tropical or tropical environment, the required chilling conditions will not be experienced and this will prevent the species from occurring in these environments.

CLIMEX enables the user to model the limiting effects of obligate diapause requirements by establishing a required set of induction and termination conditions. Diapause can be induced by a specified combination of temperature and daylength conditions, and can be terminated by a specified temperature condition. There is also a parameter specifying the number of days below or above the diapause-inducing temperature necessary to complete diapause development. If the required number of days for diapause development is not achieved, the conditions for obligate diapause have not been met and the species will be excluded from the location. CLIMEX reduces the Weekly Growth Index (GI_w) to zero during the period of diapause development and it sets the annual GI_A to zero if the obligate diapause conditions are not satisfied. This has the effect of setting the Ecoclimatic Index to zero and excluding the species from that location. While a species is not experiencing diapause, the weekly Diapause Index (DI_w) = 1. When the species is in diapause, $DI_w = 0$.

5.7 Ecoclimatic Index (EI)

The Ecoclimatic Index (EI) integrates the Annual Growth Index (GI_A), which describes the potential for population growth, with the annual stresses that limit survival during the unfavourable season and with any limiting factors, ie PDD, diapause. As few climates are 100% suitable throughout the year, the value of GI_A rarely reaches its potential, thereby limiting the maximum value of EI. The EI thus gives an overall measure of the potential of a given location to support a permanent population Sutherst (Sutherst and Maywald, 1985). The index was designed to provide simple, concise information in a form that is readily interpretable by non-scientists, such as policy makers.

The EI is scaled between 0 and 100, with an EI close to 0 indicating that the location is not favourable for the long-term survival of the species. EI values of 100 are only achievable under constant and ideal conditions comparable in incubators. An EI of more than 30 represents a very favourable climate for a species, as it means that during the (say) six months suitable for growth with a maximum GI_A of 50, the species has achieved 60% of the potential population growth. The difference is due to the daily fluctuating temperatures as well as the likely absence of continuous optimal rainfall values.

$$\text{Ecoclimatic Index, EI} = TGI_A \times SI \times SX$$

Where

$$GI_A, \text{ the Annual Growth Index,} = 100 \sum_{i=1}^{52} TGI_{W_i} / 52$$

SI, the Annual Stress Index, = $(1-CS/100)(1-DS/100)(1-HS/100)(1-WS/100)$, and

SX, the Stress Interaction Index, = $(1-CDX/100)(1-CWX/100)(1-HDX/100)(1-HWX/100)$,

CS, DS, HS, WS are the annual cold, dry, heat and wet Stress indices respectively,

and CDX, CWX, HDX and HWX are the annual cold-dry, cold-wet, hot-dry and hot-wet Stress interaction indices.

The EI is also affected by the constraints of having sufficient degree-days to complete the lifecycle (PDD) and meeting the conditions required to complete diapause.

If CLIMEX is run with two species simultaneously, in addition to explicitly including competition via the moisture (Section 5.3.3), radiation (Section 5.3.4) and biotic (Section 0) indices (which will affect the weekly Growth Index, and hence the EI), competitive exclusion can be specified. When this is done, the Ecoclimatic Indices, EI1 and EI2, are compared, and the smaller of the two is set to 0, while the other is left unchanged. If they are equal, both are left unchanged. This method assumes that over a sufficient number of generations, a species that out competes another will displace the other species completely.

CLIMEX produces values for separate stages in evaluating the Ecoclimatic Index

- (1) The “raw” Ecoclimatic Index (EIA) that may include inter-species effects via the effects of moisture and radiation competition, but excludes the effects of the Biotic Index (BI) is calculated as follows:

$$\text{Ecoclimatic Index, EI}_A = \text{TGI}_A \times \text{SI} \times \text{SX}$$

- (2) An Ecoclimatic Index (EIB) that includes the effects of the Biotic Index (EIB) is calculated as follows:

$$\text{Ecoclimatic Index, EI}_B = \text{GI}_A \times \text{SI} \times \text{SX}$$

- (3) An Ecoclimatic Index (EIC) that includes the effects of the competitive exclusion:

$$\begin{aligned} \text{EI}_C(1) &= \text{EI}_B(1), & \text{EI}_B(1) >= \text{EI}_B(2) \\ &= 0, & \text{EI}_B(1) < \text{EI}_B(2) \end{aligned}$$

$$\begin{aligned} \text{EI}_C(2) &= \text{EI}_B(2), & \text{EI}_B(2) >= \text{EI}_B(1) \\ &= 0, & \text{EI}_B(2) < \text{EI}_B(1) \end{aligned}$$

5.8 Core Distribution

The Core Distribution is the complement of the Annual Stress Index, subject to the following two constraints:

1. if a positive PDD value is specified, and the number of generations is greater than or equal to 1, and
2. if an obligate diapause is set, the necessary conditions are met

If the constraints are not met, the Core Distribution value is set to 0, irrespective of the levels of stresses at the location.

In Version 3 of CLIMEX, the Core Distribution only calculated the absence of stresses – that is, the complement of the Annual Stress Index. In Version 4 it also includes consideration of the effect of the number of Generations and the need to satisfy the requirement for obligate diapause.

CLIMEX Model Theory

The Core Distribution variable can be quite useful during model-fitting, comparing a species known distribution with the Core Distribution variable focuses attention on the effect of stresses in limiting the range of a species.

6 Fitting CLIMEX Models

6.1 Parameter fitting via inference from distribution

Parameter fitting is the most skilful procedure that CLIMEX demands. An appreciation of the distinct and separate roles of the 'Stress Indices' (that limit geographical distributions) and the 'Growth Index' (that relates to seasonal population growth) is the first requirement for the successful estimation of parameter values. It is necessary to understand the nature of the stress functions, with their threshold and rate parameters and exponential accumulation rates over time, and the simple temperature and moisture functions in the Growth Index. The other requirement is to get a 'feel' for the iterative parameter fitting process by running and re-running a parameter file with different parameter values and noting the differences in indices as a result of the changes. By varying threshold values and stress accumulation rates, the user can quickly appreciate how changes in these values alter inferred species' distributions in different ways. The extent of the change in the distribution reflects the sensitivity of the species to both the particular climatic variable and the geography of the area concerned.

As with any model, the better the supporting data, the more reliable the results. Model projections will be more robust if they are based on accurate distribution maps, abundance and seasonal activity data, supporting experimental data and census data from specific locations. We can have more confidence in these results if they have been carefully validated using independent (geographically biased) data. CLIMEX is popular because useful projections (even if coarse on occasions) can be made with a minimum amount of information. However, there is a danger that a projections is credited with greater reliability than it deserves. It should always be remembered that CLIMEX is a tool used to help in decision making: its results are based solely on climatic data and consideration should also be given to other factors, such as competition and habitats, etc.

Nonetheless, CLIMEX can provide valuable insights into the biology and potential geographical distribution of a species from even the most basic distribution map. Manipulating stress parameters to fit a known distribution will often alert the user to aspects of the species' climatic preferences that were previously unknown. Fitting parameters and comparing climates can lead to changes in research direction by eliminating or supporting hypotheses regarding species' distributions. In broad terms the process follows the method of multiple working hypotheses, first expounded by T.C. Chamberlin in 1890 (Chamberlin, 1965).

Having collated all available information on the geographical distribution of a species, it is possible to develop a general picture of its climatic preferences. One obvious starting point is to classify the species as having a tropical (wet or dry), sub-tropical (wet or dry), temperate (wet or dry), alpine, arctic, arid or other distribution. A starting set of parameter values can then be imported using the templates provided or those from another species with a similar geographical distribution. Some templates are shown in Table 6-1 **Parameters templates, representative of different geographical distributions, that are provided with CLIMEX**. The templates are provided in the species list with CLIMEX to help the user begin fitting parameters.

Some general guidelines for estimating parameters are given below. Essentially, it is a process of examining what is known about the species in relation to the information required for a parameter description. It is usually more effective to estimate stress parameter values first, as these will give pointers to those conditions that are not stressful and thus suitable for population growth. N.B. The parameters in the templates were derived from a set of climate class maps, and do not have any meaning for any particular species being modelled. Hence they do not have any value in justifying the selection of a species parameter in any modelling exercise.

CLIMEX Model Theory

Table 6-1 Parameters templates, representative of different geographical distributions, that are provided with CLIMEX.

Parameter Group	Parameter	Wet Tropical Template	Mediterranean Template	Temperate Template	Semi-arid Template	Desert Template
Temperature Index	DV0	15	10	8	10	15
	DV1	28	16	18	20	25
	DV2	33	24	24	32	40
	DV3	36	28	28	38	44
	PDD	0	600	600	0	0
Moisture Index	SM0	0.35	0.1	0.25	0.1	0
	SM1	0.7	0.4	0.8	0.2	0.001
	SM2	1.5	0.7	1.5	0.25	0.2
	SM3	2.5	1.5	2.5	0.3	0.3
Cold Stress	TTCS	0	0	0	0	2
	THCS	0	-0.005	0	0	-0.001
	DTCS	25	15	15	15	15
	DHCS	-0.002	-0.001	-0.0001	-0.002	-0.0005
Heat Stress	TTHS	37	30	30	39	44
	THHS	0.0002	0.002	0.005	0.002	0.001
	DTHS	0	0	0	0	0
	DHHS	0	0	0	0	0
Dry Stress	SMDS	0.25	0.02	0.2	0.05	0
	HDS	-0.01	-0.05	-0.005	-0.005	0
Wet Stress	SMWS	2.5	1.6	2.5	0.4	0.3
	HWS	0.002	0.0015	0.002	0.01	0.1
Hot-Wet Stress	TTHW		23			
	MTHW		0.5			
	PHW		0.075			

After creating a new species file based upon an appropriate template, the parameter values are adjusted to better fit the species of interest. The program is run, the EI values are mapped, and the modelled distribution is compared with the information available regarding the species' known distribution. While the modelled distribution (EI) will initially vary from the known distribution, the pattern should be similar. The program is then run repeatedly to systematically estimate each different parameter value until a satisfactory fit is obtained in every direction away from the core of the known distribution. Depending on the detail of the data available for the species, EI values can be related to the field data.

To estimate the stress values it is helpful to map each stress index as well as the EI. This can be done using the Options button. The maps produced will show the locations where different stresses limit the species' distribution, and also the gradients from low to high stress values. The stress parameters can then be adjusted to give a gradient of sensitivity reaching 100% at the observed distributional limit. By examining the results for specific locations on the map (ie by moving the cursor over the cross or circle and clicking the left mouse button) or data in a table, the user can observe the value of each stress at each location.

The fitting of model parameters to match the known geographical distribution and seasonal population growth of a species is, in effect, generating a hypothesis on the factors that limit its geographical distribution or population. CLIMEX enables the user to test a number of different explanations for the same observations. Each of these hypotheses needs to be explored and, where possible, tested against independent data.

It takes practice to learn to fit parameters effectively. The procedure relies on the user becoming familiar with global geography and climatic patterns, the sensitivity of stress thresholds and rate value parameters, and the temperature and moisture parameters in the Growth Index. This and the following sections detail the procedures and techniques involved in deriving a set of parameters that best describe the climatic requirements for a species, given its geographical distribution. Instructions on the use of the command boxes are given in the CLIMEX Help file.

Step 1. Collate the known distribution data from the literature. This will usually be the observed geographical distribution data, seasonal population growth rate, or experimental observations on the developmental threshold temperature and moisture levels. If using the known distribution of the species, choose one continent, preferably its native distribution, to work from. Distributions in other countries or continents can be used to verify the parameters later. Check the distributional data for reliability and consistency.

Step 2. Create a new species file by copying the parameters from another species or template with a similar distribution. This is the quickest way to start building up a set of parameters and is done clicking on the right hand button of the Species panel and selecting **Add/Edit/Remove** from the menu and then choosing an appropriate species to copy by highlighting a name in the list. Enter a new name for the species and save the new file. The parameters can then be edited using the Parameter button (P).

CLIMEX provides several example species from published papers that cover a wide range of preferred climatic regions. A set of templates has also been provided for this purpose and can be used as a foundation for building a species file.

CLIMEX Model Theory

Step 3. Enter any comments regarding the data into the **Comments** field of the species parameter file by clicking on the **Model Parameters** in the **Initialisation** drop down menu, highlighting the new species and clicking the **Comments** button. Use this facility liberally so that your work will be of the maximum value to future users of your parameter file.

Step 4 Use the **Location** right hand button to choose a continent corresponding to the observed native distribution from the **Sequences**. Run CLIMEX for the new species using the **Compare Locations** option. Note the differences in the estimated distribution (EI), compared with the known distribution.

Step 5. Repeatedly adjust the parameters and re-run **Compare Locations** until the estimated potential range best matches the known distribution. Once you have a "feel" for how changing the parameters changes the modelled potential range, the process of iteration will become much faster. Any laboratory or field experimental data can be used to tune or confirm the parameter values at this stage. When the best visual fit has been achieved, carefully document the justification for fitting each value for future reference.

Step 6. Validate the parameter values against known distributions of the species in other continents. If the species is found in more than one country or continent, validate the parameter values by running **Compare Locations** using a new **Location Sequence** that covers an observed distribution not used to derive the original parameters. This will either confirm the parameter selection or require you to adjust some parameters until a good fit for all known locations is obtained. If further tuning is carried out, it needs to be carefully documented with appropriate justification. Such tuning may be necessary if the new **Location Sequence** includes a wider range of climate data than existed in the original set.

Step 7. Congratulate yourself for having developed an hypothesis explaining the species' distribution, while concurrently producing a tool of great practical value in further analyses of the risks posed by the species in new environments.

After selecting a starting parameter file (See **Step 2** above), make a **Sequence** from a **Location Selection** to work on or use a custom set (See 'Using MetManager' that covers defining Location Selections in Help). If using a known distribution to derive the parameters, the **Location Selection** must cover this area. If possible, choose a heterogeneous environment with the widest possible range of climates. Any continent will be suitable, but caution is needed in choosing tropical locations, as some countries have little heterogeneity. For example, Kenya in East Africa has a wide range of climatic zones, while Uganda has a limited climatic range.

Try to avoid using continents or countries with geographical boundaries that cause sudden climate changes that will make determining parameter values difficult. For example, the Himalayan Mountains to the north of India make it difficult to determine the Cold Stress boundaries accurately because the geographical changes are so dramatic that a wide range of parameter values will still give a modelled distribution limit at the base of the mountain range. Similarly, oceans often form geographical boundaries that make parameter estimation of limiting conditions difficult. If unsure of some parameters values because of geographic features, it is advisable to document the caveat and validate the parameters in an independent location whenever possible. In some cases, limits will have to remain undefined until further data is available. This is particularly so where the **Location Selection** is small, eg. Tasmania in Australia, where a species may extend beyond the island limits if there was suitable habitat instead of sea.

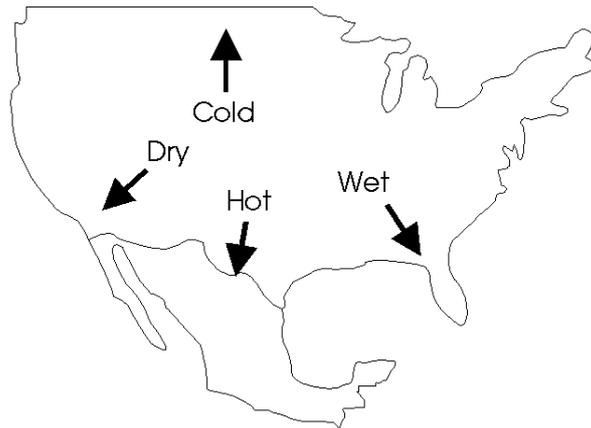


Fig. 6-1 A sketch map to reveal trends in climatic variables in North America as a guide to fitting parameter values in CLIMEX.

Using general knowledge of the climate of the continent or country, start by drawing a map of gradients of temperature and moisture using the 'Climatic-Gradients' utility. A simple example for the North America is given in Fig. 6-1. The process helps the user to visualise how gradients of stress accumulation will look.

Begin by adjusting the stress thresholds and stress accumulation rates. Using the example of the USA below (Fig. 6-2), there is a species with a distribution that is centred in central USA, spreading towards the national boundaries in each direction. The objective is to model a core distribution (locations with high EI values and little or no stress) that spreads outward to locations with progressively decreasing EI and increasing stress. At the limit of the distribution, one or more stress indices will be equal to or greater than 100, indicating that the location is not suitable for the long-term survival of the species in average years (as inferred from most distribution maps).

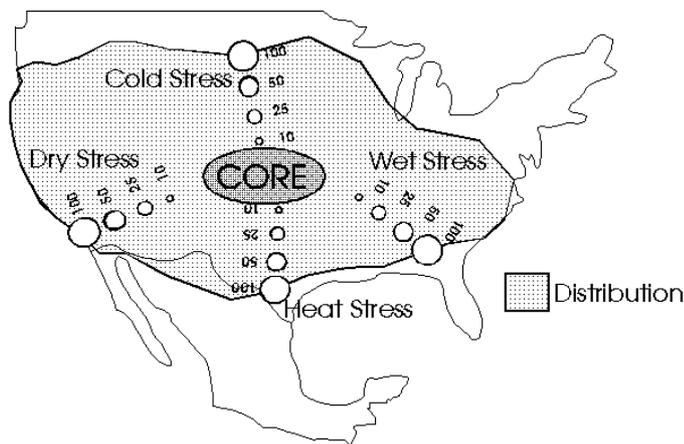


Fig. 6-2 Map of North America showing the hypothetical trends in each stress value in relation to distribution limits.

Pick one stress index (Cold, Hot, Wet or Dry) and identify the distribution limit defined by this stress index. In this example, Cold Stress is a likely variable that limits the northern distribution but the alternative hypothesis that the limit is determined by the availability of sufficient degree-days of development (PDD) also needs to be tested. Cold stress can be modelled either using a threshold

CLIMEX Model Theory

temperature model or a degree-day model (see Section 5.4.1). Begin by using the threshold temperature model, setting TTCS to 2 °C and the accumulation rate THCS to 0.001. This rate simulates a slow accumulation of stress during the winter and consequently indicates a low gradient in the total accumulated stress moving from the core to the northern distribution limit. Vary the stress threshold and/or the stress rate and re-run **Compare Locations** to try to match the northern distribution limit to that in Fig. 6-2. Then repeat the process using the degree-day based Cold Stress (DTCS) with an accumulation rate (DHCS) of 0.001 as above, with the THCS value temporarily set equal to zero to deactivate the threshold temperature model. In order to use the degree-day Cold Stress function, it is first necessary to estimate the value of DV0 (see below). Compare the results of the two models of Cold Stress, and choose the threshold and rate parameters that give the best description of the distribution in Fig. 6-2.

Repeat this process for each Stress Index. Once the stress levels are set, these values can be used to set the bounds for the Growth Index parameters for Temperature and Moisture.

6.1.1 Temperature parameters (DV0 - DV3)

The parameters DV0 - DV3 define the temperature range that is suitable for population growth and development. Temperature stress therefore cannot accumulate in this range. The Cold Stress Temperature Threshold (TTCS) must be \leq DV0 and the Heat Stress Temperature Threshold (TTHS) must be \geq DV3. Once the stress thresholds have been set, the temperature parameters for growth (DV0, DV1, DV2, and DV3) can be fitted. To fine-tune the values, include any experimental results or knowledge about how the species responds to temperature. The parameter values relate to long-term average temperatures and the Temperature Index is accumulated over a daily temperature cycle. Follow these steps:

1. Relate seasonal abundance/activity or population growth rate data to temperature at a number of locations. Low temperature requirements can be inferred from the timing of first activity in spring and high temperature tolerance by quiescence or high mortality rates in summer.
2. Check for day length effects operating to modify the GI values through the Light Index and diapause mechanisms. Diapause is taken into account by reducing the GI to zero in response to a given daylength (DLD0), or daylength in combination with temperature (DLT0). It is relaxed when the diapause termination temperature (DLT1) is reached later in the year, provided that the conditions for completion of diapause development have been satisfied.
3. Examine any experimental data relating to development rates at various temperature and moisture levels. If available, these can be extrapolated to give the four temperature indices in the absence of field data, or to confirm estimates derived from field data.

6.1.2 Moisture Parameters (SM0 - SM3)

The moisture range SM0 - SM3 defines soil moisture levels that are suitable for population growth and development. Moisture stress therefore cannot accumulate in this range. The Dry Stress Threshold (SMDS) must be \leq SM0 and the Wet Stress Threshold (SMWS) must be \geq SM3. Once the stress thresholds have been set, the

moisture development parameters (SM0, SM1, SM2, and SM3) can be fitted. To fine-tune the values, include any experimental results or knowledge about how the species responds to moisture. The parameter values relate to mean soil moisture in the soil layer within rooting depths of most crops and grasses. Moisture preferences must be estimated by relating them to population growth and abundance data in the field at different locations, with different amounts of rainfall. New CLIMEX users need to become familiar with the relationship between soil moisture on one hand and both rainfall and evaporation on the other, in order to be able to estimate these parameter values (Fig. 6-3). In order to help users visualise the regimen of soil moisture at any given location, as estimated using the hydrological model in CLIMEX, we have provided ‘_Soil Moisture Utility’ as a species file. The parameters have been set to ensure that the ‘moisture index’ appears as if it is directly proportional to the actual soil moisture content.

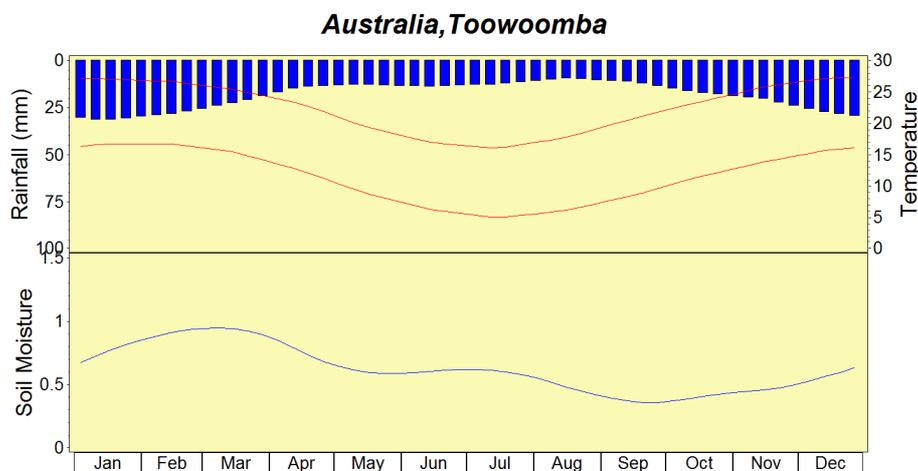


Fig. 6-3 Graph of the Moisture Index for the ‘_Soil Moisture Utility’ (SM0=0, SM1=1, SM2=4, SM3=10) at Toowoomba, Queensland, illustrating the relationship between soil moisture and rainfall.

The above figure shows that soil moisture levels can be compared with the rainfall to show the seasonal changes and also the values for reference when visualising the relationship between the species requirement and soil moisture rather than the traditional relative humidity measure.

6.1.3 General Parameter Fitting Example

The best way to understand the process of parameter estimation is to work through an example.

The codling moth, [*Cydia pomonella* (L.)], is a major pest of deciduous fruit around the world. This example shows the basic principles used to model the potential distribution of the moth from a simple distribution map, in this case its known range in Europe, North Africa and west Asia (Fig. 6-4).

CLIMEX Model Theory

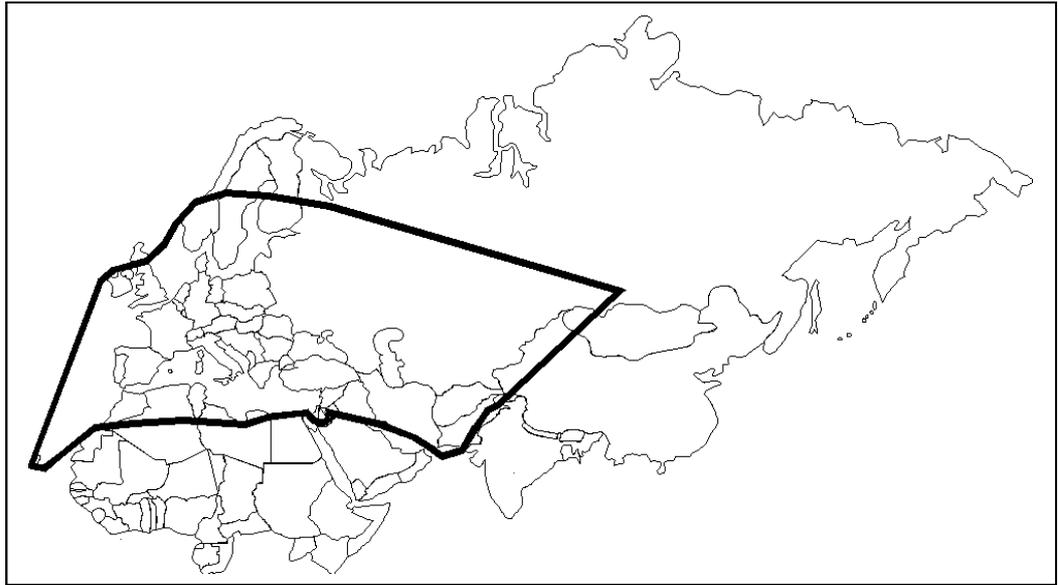


Fig. 6-4 The reported geographical distribution of the codling moth in Europe, northern Africa and western Asia (Commonwealth Institute of Entomology, 1976).

Several conclusions can immediately be drawn from the distribution map. Firstly, the moth survives in the cold weather of northern Europe and in the warmer dry Mediterranean region, but appears not to survive in the wet tropical regions of Africa. This indicates a low developmental threshold for both temperature and moisture, and that either high temperatures or high rainfall limit the favourableness of the tropical regions. Given these observations, an initial CLIMEX model could be based on parameter values of a temperate species (in this case from

on page 52), but with DTCS set to 25 to make the exercise more instructive. A map of a CLIMEX run with these parameter values is given in Fig. 6-5.

The map in Fig. 6-5 shows a much more restricted distribution than is observed in practice. Examination of individual locations on the map shows that Heat Stress, Dry Stress and Cold Stress figure prominently. Northern Africa and western Asia are all indicated as being too hot and dry for the moth. Northern Europe and eastern European countries are too cold. As the known range of the moth includes these areas, the parameters need to be adjusted.

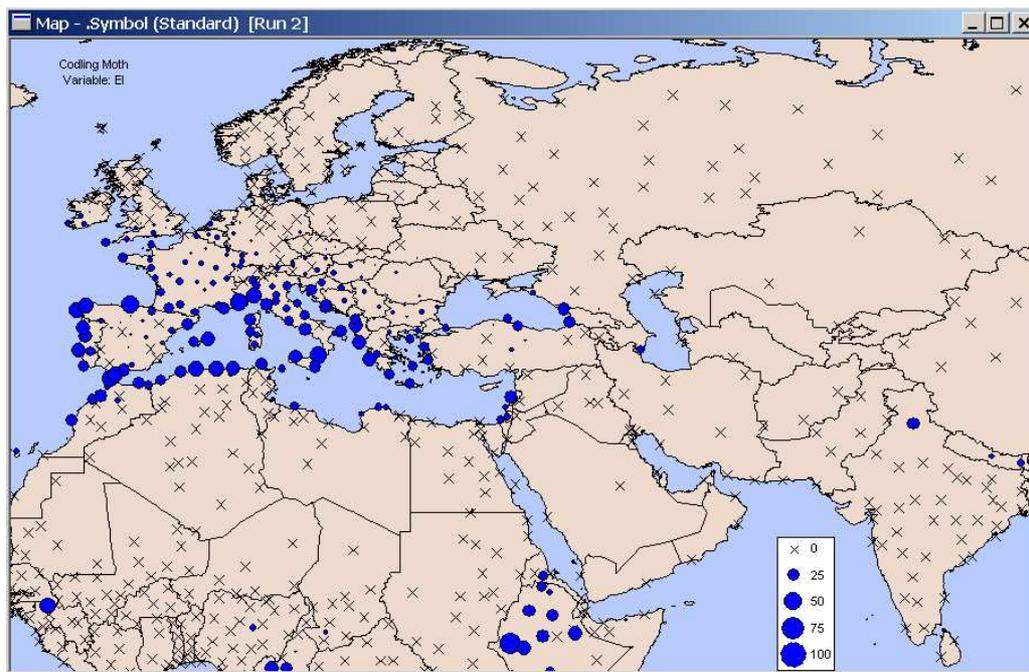


Fig. 6-5 The modelled distribution (EI) of the codling moth in Europe, northern Africa and Asia using the Temperate Template parameters in Table 6-1 Parameters templates, representative of different geographical distributions, that are provided with CLIMEX. with DTCS=25.

Cold stress makes much of Europe and Russia unsuitable. However, we know that the codling moth occurs in these areas, and that it also has an obligate winter diapause. . Thus, the first thing to do is to remove cold stress entirely, and to set the winter diapause functions. Setting diapause parameters must generally be done using data from the literature. Diapause is initiated with decreasing daylength (somewhere between 13 and 18 hours of daylight) and when summer temperatures drop below 15 °C (Shel'deshova, 1967; and other references cited within these; Bajwa and Kogan, 1998). A prolonged period of about 3 months below the developmental threshold is required for successful completion of diapause (Geier, 1963). Diapause is terminated when the necessary period of cooling has been completed, and temperatures begin to rise between 0 °C and 10 °C (Shel'deshova, 1967; Bajwa and Kogan, 1998). These findings indicate that the diapause parameters be set as follows:

These findings indicate that the diapause parameters be set as follows:

DPD0: set to 14 hours, so that as daylength decreases below this value, diapause will be initiated when the correct temperature (DPT0) is reached. Trial and error indicated this to be the best value to enable diapause to be correctly initiated in northern latitudes while also enabling diapause to be initiated in northern Africa (at

CLIMEX Model Theory

30°N, the maximum daylength is about 14.1 hours, so a higher value will preclude diapause from being initiated).

DPT0: set to 11 °C, as prolonged cooling below the developmental threshold (10-12 °C) is necessary (Shel'deshova, 1967). The value of 11 °C is also consistent with the indication that summer temperatures below 15 °C induce diapause (Bajwa and Kogan, 1998).

DPT1: set to 6 °C. This value is consistent with Shel'deshova (1967) and other references in Bajwa and Kogan (1998). This value of 6 °C is a compromise. A higher value precludes development in northern latitudes where minimum temperatures often do not exceed 6 °C whilst daylength is still increasing – ie prior to mid-June), and a lower value prevents the successful completion of diapause in northern Africa where minimum temperatures even in winter do not drop much below 6 °C.

DPD: set to 90 days, as a minimum of 3 months below the developmental threshold is necessary for successful completion of diapause (Geier, 1963). A lower value (eg, 30 days) extends the distribution range of the codling moth in southern Europe, South America, Australia and Africa, but more detailed information on the occurrence of the codling moth in these specific locations would be needed to justify lowering this value. Distribution maps only indicate the broad regions where a species occurs, and tend to include rather than exclude regions.

DPSW: set to 0 for winter diapause.

The literature search for information on factors inducing and terminating diapause also indicated that the temperature parameters of the Temperate Template need adjusting. In the Temperate Template, DV0 = 8 °C, DV1 = 18 °C, DV2 = 24 °C, DV3 = 28 °C and PDD = 600 degree-days (above 8 °C).

As we saw above, in fitting the diapause functions, the developmental threshold is between 10 °C and 12 °C, so DV0 must be adjusted. Several iterations, adjusting DV0 between each, indicated that 10 °C was the most appropriate value for this parameter. The optimal temperature for codling moth seems to be between 20 °C and 30 °C (Shel'deshova, 1967), so DV1 is set to 20 °C and DV2 is set to 30 °C. DV3 must also be adjusted, and 33 °C seems consistent with the literature (Shel'deshova, 1967; Pitcairn *et al.*, 1991; Bajwa and Kogan, 1998)

A value of 600 degree-days above 10 °C (ie PDD = 600) is the widely accepted minimum heat accumulation for the completion of a generation (e.g., Shel'deshova, 1967). So let's begin by seeing how these changes have influenced the potential distribution. The values we have altered in our new species file, copied from the Temperate Template, are as follows:

Temperature Parameters	Diapause Parameters	Cold Stress Parameters
DV0 = 10 °C	DPD0 = 14 (hrs of daylight)	TTCS = 0 °C
DV1 = 20 °C	DPT0 = 11 °C	THCS = 0 °C
DV2 = 30 °C	DPT1 = 6 °C	DTCS = 0 °C

DV3 = 33 °C	DPD = 90 days	DHCS = 0 °C
PDD = 600	DPSW = 0 (for winter diapause)	

With all other parameter values untouched, these adjusted parameter values provide the distribution shown in Fig. 6-6.

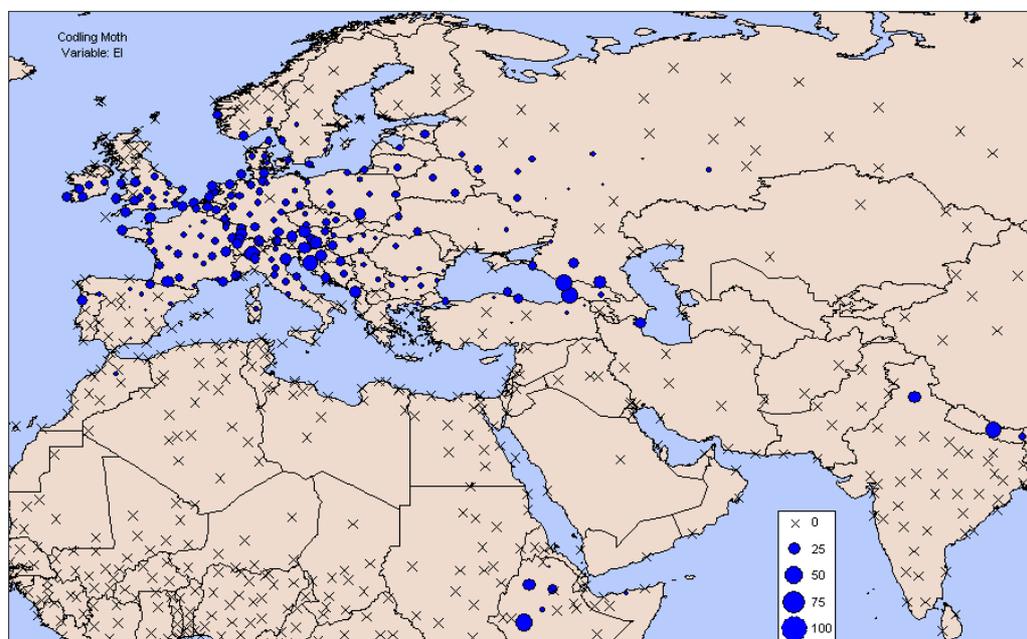


Fig. 6-6 Modelled potential distribution for the codling moth using the original Temperate Template parameters with the development temperatures and cold stress values adjusted, and the inclusion of winter diapause.

The distribution has extended further northwards, as we would have hoped by removing the cold stress parameters. Examination of locations in Scandinavia and Russia that are still unsuitable indicates that although the EI is zero, both the temperature and moisture parameters are suitable, and there is a positive GI value. Although temperature and moisture conditions are suitable for growth, and there is a positive GI value, these locations are not accumulating 600 degree-days above 10 °C. A generation cannot therefore be completed, and the species is excluded from the locations, with EI being set to zero. This throws into question the accuracy of the above estimate of PDD=600 degree-days > 10 °C.

If we reduce PDD to 450 days, we note that the potential distribution of the codling moth extends further north (Fig. 6-7), and now includes locations (eg Kirov, Perm and Tomsk in Russia) that are known to support permanent populations (Shel'deshova, 1967).

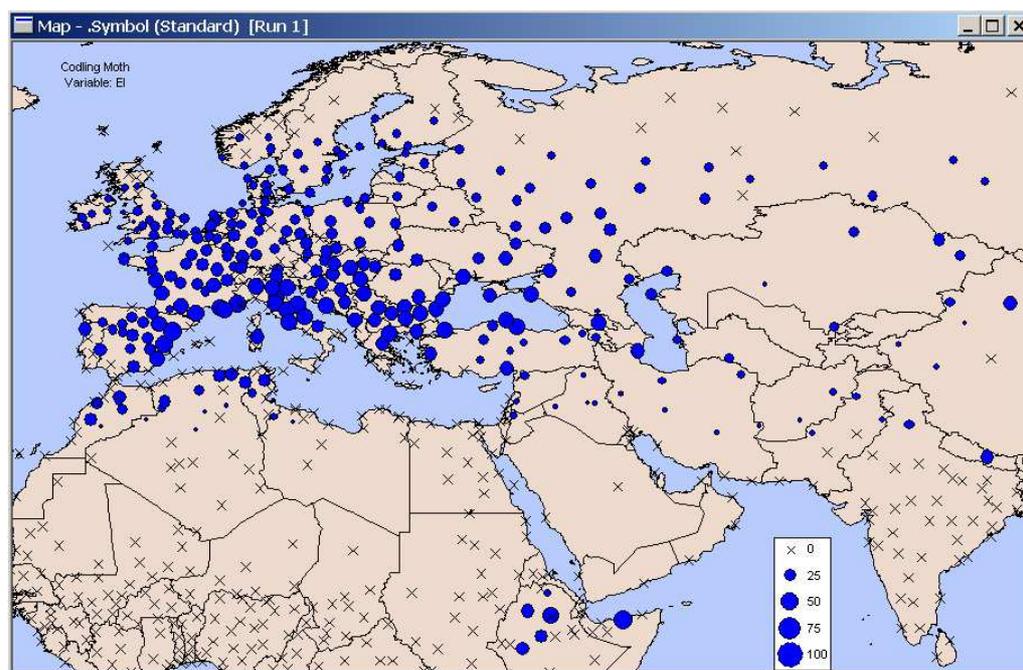


Fig. 6-7 Potential distribution of the codling moth with PDD reduced from 600 to 450 degree-days > 10 °C.

Many areas that should be within the distribution (Fig. 6-4) are still indicated as being unsuitable, either because of excessive heat or dry stress, or because the moisture index (MI) is limiting. Having adjusted all of the other temperature parameters, let's begin by adjusting the heat stress parameters. We now have an upper temperature threshold (DV3) of 33 °C, but the heat stress temperature threshold (TTHS) is still 30 °C, with an accumulation rate of 0.005. As TTHS cannot be below DV3, we must increase its value to at least 33 °C. Because we need to address the excessive heat stress, we also reduce the accumulation rate (TTHS) from 0.005 to 0.0003.

However, in order to see a major change in the modelled potential distribution of codling moth, we must also alter the moisture and dry stress parameters. Adjusting these parameters is somewhat complicated. Given that the codling moth is a pest of orchard fruit, the role of irrigation in maintaining a suitable microclimate in arid zones needs to be considered. Either we can assume that 1.5 mm day⁻¹ of irrigation are applied weekly in the summer months, and use this scenario to tune the CLIMEX moisture parameters. This could be done by selecting the **Irrigation** component module, and setting the appropriate value for irrigation (1.5 mm per day in summer). Alternatively, parameters can be fitted to the raw meteorological data on the assumption that the host fruit are grown in such climates. This carries the implicit assumption that irrigation must be occurring; otherwise the pome fruit could not be grown in these areas. We will try to fit the moisture and dry stress parameters to the raw meteorological data.

To reduce moisture stress, we have to alter both the moisture and the dry stress parameters. Let's begin by reducing the soil moisture parameters. The distribution map (Fig. 6-4) indicates that the codling moth occurs in dry areas, so let's reduce SM0 from 0.2 to 0.1, SM1 from 0.8 to 0.2, SM2 from 1.5 to 1 and SM3 from 2 to 1.2. Because we have reduced SM0, we must also reduce the dry stress threshold (SMDS) from 0.2 to at least 0.1 (Note that DYMEX forces consistency in the parameter values at all times, so the order in which you make these adjustments is important). This

in itself will reduce the dry stress accumulated, and reducing the accumulation rate (HDS) from 0.005 to 0.003 will further reduce the amount of dry stress accumulated. Finally, because we have reduced SM3 from 2.0 to 1.2, we should also alter the wet stress parameters. Let us reduce the wet stress threshold (SMWS) from 2.5 to 1.2. As this should increase the amount of wet stress accumulated, let's decrease the accumulation rate (HWS) from 0.002 to 0.0005.

The final result (Fig. 6-8) accords well with the overall distribution shown in Fig. 6-4. Some Mediterranean areas indicated as being within the distribution range in Fig. 6-4 are indicated as being unsuitable in Fig. 6-8 because obligate diapause conditions cannot be met (winters are not cold enough). Other areas in Europe and northern Russia are indicated as unsuitable in Fig. 6-8 because there are insufficient degree-days above 10 °C to enable a generation to be completed. To tune the model, it would be necessary to know whether or not codling moth occurs in these specific locations. It is likely that there are locations, not represented sufficiently well by the limited climate data set, that provide suitable conditions for survival.

The results of a world simulation (Fig. 6-9) can be compared to the known world distribution (Fig. 6-10). If suitability of northern Canada and Russia are thought to be too high, cold stress can be made to accumulate below temperatures of -20 °C at a rate of 0.001 week⁻¹. This would be consistent with results of various authors cited in Shel'deshova (1967).

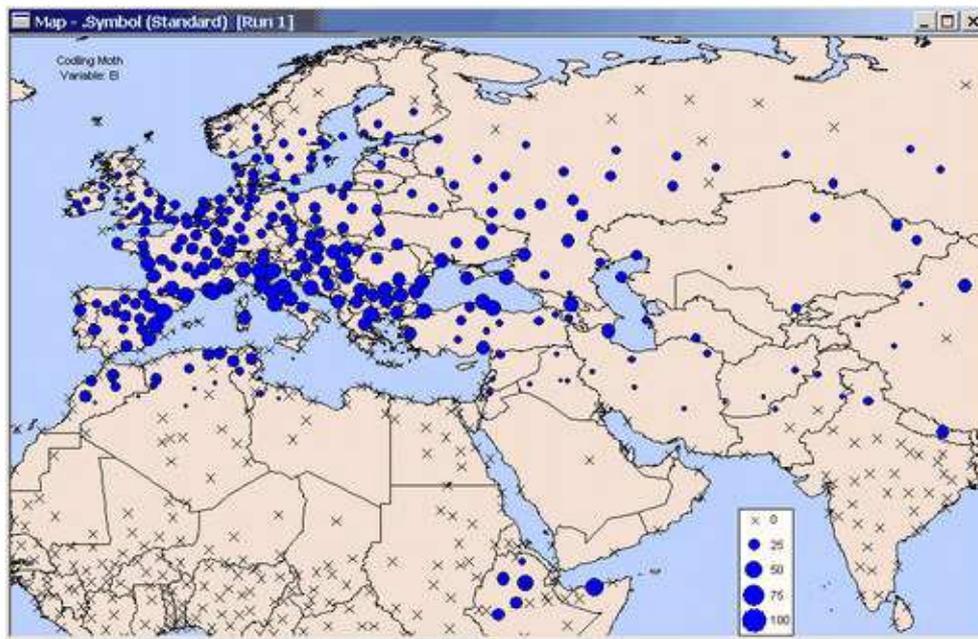


Fig. 6-8 Potential distribution of the codling moth adjusted for heat stress, moisture parameters, dry stress and wet stress.

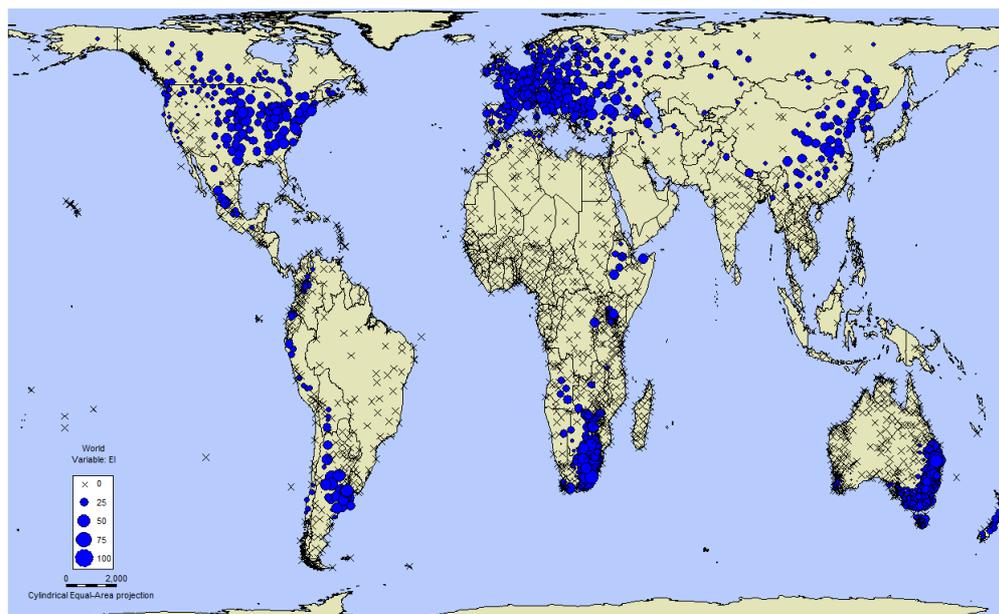


Fig. 6-9 Modelled potential world distribution of codling moth using final parameters.

There are some other discrepancies between the modelled and known distributions, but these may not be due to climate. For example, the CLIMEX model suggests that codling moth could occur in the south-western corner of Australia (Fig. 6-9), which has a Mediterranean climate that is clearly suitable for the moth (Fig. 6-4). Whilst the moth does not in fact currently occur in Western Australia (Fig. 6-10), it has occurred here (Wickens 1928 and Jenkins 1947, 1956, cited in Shel'deshova, 1967) and (Barnes 1991 cited in Bajwa and Kogan, 1998). Similarly, the more restricted distribution in Africa could be due to other factors that are quite separate to the climatic factors regulating the potential distribution. CLIMEX only considers the climatic suitability of locations for a species – not the presence of suitable hosts, or other factors which may be responsible for limiting actual distributions.

This example is obviously a condensed version of the parameter fitting process, as it would take too long to describe all the iterations. Until the user becomes familiar with the program, it is suggested that only one parameter be changed at a time to keep track of the changes in distribution. The value of each of the stresses affecting the species at each location can be examined by selecting the location on the map with the mouse, viewing the results in tabular form or by mapping any of the stress indices rather than the EI using the Options button. Mapping the stresses gives a much better indication of where alterations need to be made.

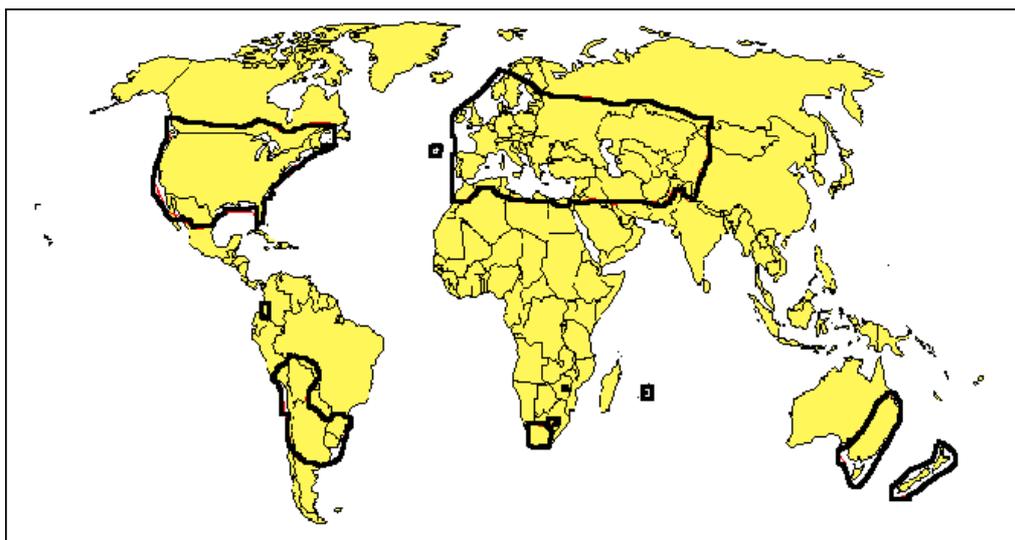


Fig. 6-10 The known world distribution of codling moth (Bajwa and Kogan, 1997).

Realistically, it will probably take a week to complete and document a CLIMEX model including literature search, parameter-fitting and examination and reporting of the behaviour of the model.

References

Ferro, D.N., Sluss, R.R. & Bogyo, T.P. (1975). Factors contributing to the biotic potential of the codling moth, *Laspeyresia pomonella* (L.), in Washington. *Environmental Entomology* 4(3): 385-391.

Hill, D.S. (1987). *Agricultural Insect Pests of Temperate Regions and their Control*. Cambridge University Press, Cambridge. p 402.

6.1.4 Diapause Parameter Fitting Examples

If a species has obligate diapause, the parameters should be set prior to estimating the stress factors. This applies especially to the “relevant” stress factors (eg Cold Stress if the species has winter diapause, and Heat Stress if the species has summer diapause). Heat and Cold Stress can still apply to species undergoing obligate diapause, but as diapause is a mechanism for avoiding unfavourable conditions, the conditions causing stress to occur during diapause are likely to be extreme. If the Heat and Cold Stress parameters are set too high, then excessive stress will accumulate during diapause development, thereby excluding the species from locations which are in fact favourable for its long term survival.

Users are referred to Section 5.3.6 and 7.1 of the User’s Guide for a description of the diapause functions and how they are implemented in CLIMEX. We start with the assumption that a requirement for diapause to complete a lifecycle is obligate in the species of interest. Thus the requirement will limit the range of habitats that the species can occupy to those that provide suitable climatic conditions for diapause development.

CLIMEX Model Theory

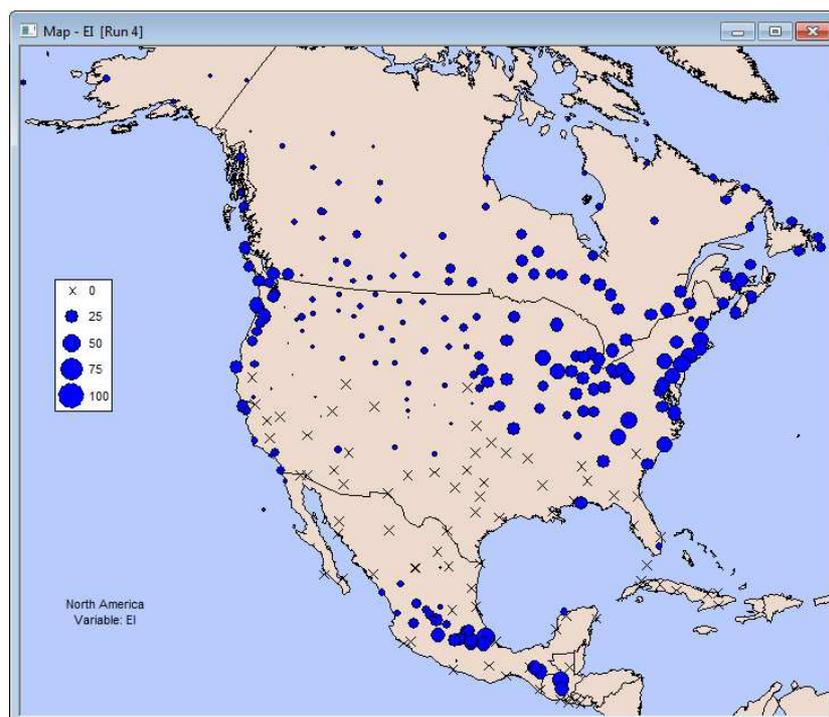


Fig. 6-11 Distribution of species with no diapause for North America, using parameter settings from Table 6-2.

Recall that in CLIMEX, short daylengths and low temperatures (winter diapause) or long daylengths and high temperatures (summer diapause) can trigger diapause. CLIMEX models the limiting effects of obligate diapause by establishing a set of induction and termination conditions and a number of days below/above a diapause inducing temperature to complete diapause development. CLIMEX reduces the Growth Index to zero during this period of diapause development. While a species is not in diapause the weekly Diapause Index (DI_w) = 1. $DI_w = 0$ for the time that the species is in diapause.

To assess the effect of obligate diapause on the geographical distribution of a species, let's consider a hypothetical species which has a broad distribution range and is only marginally affected by any stress factors, such that it can persist from northern Canada to Mexico (Fig. 6-11).

The parameter values used to obtain this distribution are given in Table 6-2, page 69. If we remove the marginal limiting effects of Cold Stress, and impose obligate winter diapause as a constraint (Table 6-2), we see that the species is now excluded from the warmer southern areas (Fig. 6-12).

Table 6-2 Values for parameter settings of species with and without obligate diapause.

Parameter Group	Parameter	Without Diapause	Winter Diapause	Summer Diapause
Temperature Index	DV0	2	2	2
	DV1	22	22	22
	DV2	26	26	26
	DV3	30	30	30
	PDD	0	0	0
Moisture Index	SM0	0.1	0.1	0.1
	SM1	0.7	0.7	0.7
	SM2	1.0	1.0	1.0
	SM3	2.0	2.0	2.0
Cold Stress	TTCS	0	0	0
	THCS	0	0	0
	DTCS	0	0	0
	DHCS	0	0	0
Heat Stress	TTHS	30	30	30
	THHS	0.01	0.01	0.01
	DTHS	0	0	0
	DHHS	0	0	0
Dry Stress	SMDS	0.1	0.1	0.1
	HDS	-0.01	-0.01	-0.01
Wet Stress	SMWS	2	2	2
	HWS	0.05	0.05	0.05
Diapause	DPDO	0	10	10
	DPTO	0	10	30
	DPTi	0	0	30
	DPD	0	90	30
	DPSW	0	0	1

CLIMEX Model Theory

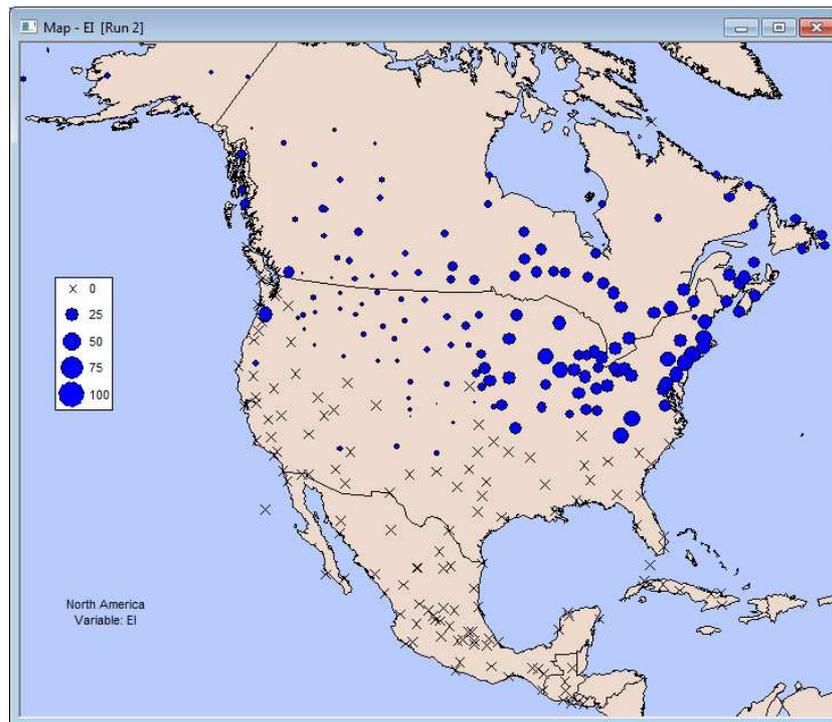


Fig. 6-12 Distribution of a species with obligate winter diapause in North America using parameter settings from Table 6-2.

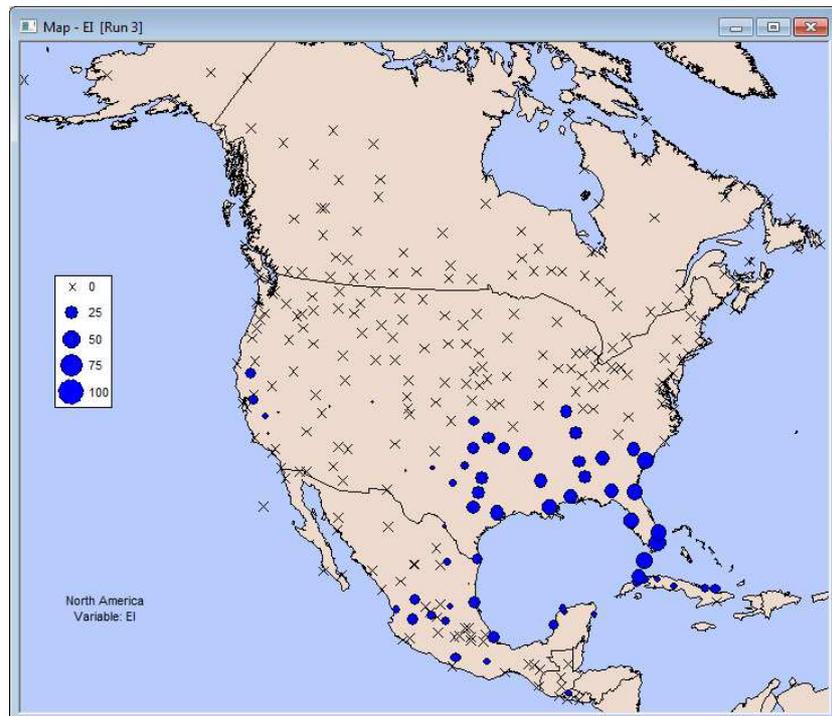


Fig. 6-13 Distribution of a species with obligate summer diapause in North America using parameter settings from Table 6-2.

If we impose obligate summer diapause (Table 6-2), the species is excluded from the colder northerly areas and the high altitude southern areas that do not experience the necessary heat conditions to initiate and maintain diapause (Fig. 6-12).

6.2 Compare Locations Model-fitting tips

When fitting models trust nothing, not even your instincts! As you fit models you will inevitably be faced with conflicts between your model results, published observations, theoretical expectations, and underlying climatic data. Your ability to resolve such conflicts is a mark of your abilities as an ecological modeller. With experience, your instincts will lead you more rapidly to the likely culprit, or at the least to an understanding of the conflict and a hypothetical framework for resolving it.

6.3 Distribution data

Distribution data is one of the most important sources of information we have to inform our understanding of the conditions a species can tolerate. Unfortunately it is fraught. Chapman (2005a; 2005b) identifies many of sources of error in species distribution data, and provides techniques for their identification and treatment. The most important ones to be aware of here concern location accuracy, location precision, taxonomy, and meaning.

6.3.1 Epistemic uncertainty

One very important source of data uncertainty that Chapman ignores is the epistemic error or meaning of the record. So, notwithstanding that the location record may be in the wrong location, or it may refer to an incorrectly classified taxon, but its observation at a location may represent an observation of an ephemeral population or one that is established. In either circumstance, it may be able to be present only by virtue of an anthropogenic or natural climate-modifying factor (eg irrigation, glasshouses, or natural microhabitat).

6.3.2 Confounded uncertainty

Our primary modelling covariates are climatic data (temperature, rainfall and humidity). Despite some tentative advances in the use of remotely sensed data, the primary source of our climate data still comes from meteorological stations with temperature and humidity sensors housed in shaded shelters (Stevenson screens) and tipping bucket rain gauges. These stations are an imperfect source of climate data: they are irregularly placed, suffer from sensor drift and inaccuracy, may stop working entirely from time to time, etc. These days it is unusual to see published niche maps that employ station data directly. It is more usual for researchers to use gridded climate datasets for fitting and projecting their models. When we are fitting our model using these gridded climate datasets to our distribution data we may find that we have troubles that are due to the distribution point not being located at the same position that the climate has been estimated. This is type of error is most frequently encountered in areas of high topographic relief (e.g., the Southern Alps of New Zealand; Kriticos *et al.*, 2004). In this case, the plants were observed in the valley floor while the climate stations were located at the top of mountains. It is easier to check for this type of error if the climate dataset is estimated for a specific point location on the lattice of points (e.g., CliMond), rather than being estimated for an average value within a defined grid cell (see Kriticos and Leriche, 2009 for a brief discussion). If the species distribution record includes an altitude value, this can be compared with the altitude value in the climate dataset. The adiabatic lapse rate can be used to estimate the temperature difference between the collection location and the climate data pseudo-station. A detailed knowledge of the topography and

CLIMEX Model Theory

consideration of the prevailing rainfall pattern may indicate whether the species is likely to be located at a site experiencing wetter or drier conditions than the climate data pseudo-station.

6.4 Growth Indices

Growth indices are best informed by experimental data or phenological observations, but for most applications the parameters can be set within the bounds defined by the fitted stress indices, combined with some additional considerations. For the Temperature Index parameters, DV0 corresponds to the lowest temperature that individual (and population) growth commences. This is sometimes called the base temperature or lower developmental threshold in the literature. In the literature separate developmental thresholds are often quoted for each lifestage. However, Jarošík *et al.* (2002) showed that the lower developmental threshold for a species does not vary with lifestage and that the proportion of time spent in a lifestage does not vary with temperature. This notion, termed rate isomorphy has been shown to occur in all ectothermal animals (Jarošík *et al.*, 2004). Rate isomorphy helps explain why a very simple model like CLIMEX, which does not distinguish between lifestages, can model poikilotherms successfully. The fact that selection on a particular developmental stage is related to the thermal adaptation of the other stages confers a degree of stability on selection processes. The remarkable similarities in some of the cardinal temperatures for related taxa indicate that in the absence of direct experimental observations, the lower developmental threshold (DV0) and the sum of effective temperatures (PDD) can be estimated from phylogenetically closely-related species (Jarosik *et al.*, 2011). As observed by Jarosik *et al.*, this is important for modelling the potential distribution of many pests of phytosanitary concern, for which there may be little scientific knowledge.

Jarosik *et al.* discovered that 90% of species they examined had values for DV0 between 3.6 and 16.5 °C (2011). Tropical species tend to have higher values for DV0 (Trudgill, 1995).

Typically, growth rates rise in an approximately linear fashion in relation to temperature, up to the optimal temperature for growth before dropping fairly sharply (Stinner *et al.*, 1974). DV1 and DV2 define the optimal range for population growth. These define the range of temperatures between which growth rates are 90-95%. Some authors mistakenly try to adjust the gap between DV1 and DV2 in order to obtain higher values of EI throughout most of the species range. This not only distorts the model parameters, but reduces the ability of the model to discriminate different levels of climatic suitability. In practice, the gap between DV1 and DV2 should normally be about 4 °C. An exception to this situation seems to occur in at least some plant pathogens, where the optimal temperature range appears far wider. There are remarkable similarities in some of the cardinal temperatures for related taxa (Jarosik *et al.*, 2011). Reanalyzing the dataset of Jarosik *et al.* (2011) revealed another useful pattern. The gap between DV0 and the optimum temperature for development was approximately 21 °C (average = 20.8 °C and median = 20.5 °C), with a standard deviation of 4.05 (n=232 species for which both the lower limit and optimum were available).

For 90% of species in the dataset of Jarosik *et al.* the heat sum required to complete a generation (PDD) fell in the range 201.1 and 472.7 °C Days (2011).

6.4.1 Soil Moisture

If you are modelling a pathogen or an insect, remember that they are likely to be dependent on plant hosts, and hence, while they may not be living directly in the soil environment, they may be strongly influenced by the soil moisture relations of their hosts. The permanent wilting point for plants is approximately 0.1. Below this point, plants wilt, and populations tend to become stressed. Hence, an appropriate starting point for modelling is an SMO and SMDS value of 0.1. Phanaerophytes (plants with very deep root systems adapted to arid environments) may be able to access water deeper than the bucket model in CLIMEX is simulating, and hence the appropriate Dry Stress threshold may be lower than 0.1.

Field capacity is 1.0. Many plants become stressed when subject to water-logging conditions. If so, an SM3 value near 1.0 may be appropriate.

6.5 Stresses

The stresses mostly define the species range boundary, and hence are typically best defined by fitting them to the species known geographical distribution. However, each mechanism has between two and three parameters. In many cases, at least one of these parameters can be informed by experimental observations (which may have been reported in the scientific literature). These observations may pertain to the species being modelled, or possibly an ecologically similar species. If you can anchor one of the parameters using such knowledge, it becomes easier to fit the remaining parameter(s) for that mechanism using the distribution data.

6.5.1 Cold Stress

There are three separate cold stress mechanisms (Threshold temperature, Degree-Day and Average temperature). Each mechanism has a characteristic fitting behaviour, geographical pattern and putative biological or ecological mechanism. The Threshold Temperature mechanism mimics the effects of direct tissue damage due to low temperatures. This is often associated with the phase change of aqueous solutions to a solid state (e.g. freezing). When this happens within cells, plants and animals tend to experience tissue loss or death. In insects, this phenomenon is known as the super-cooling point. Thus, when the cold stress threshold is exceeded, the stress accumulation rate tends to be quite high in magnitude.

The degree-day cold stress mechanism is associated with an energetic balance. If the conditions are such that the organism is unable to generate sufficient energy (e.g. through foraging or photosynthesis), it draws down on reserves until it dies or conditions improve. The threshold for calculating the degree-day heat sum for growth is usually the base temperature for development (DV0). The necessary minimum heat sum for offsetting respiration losses is typically in the order to 10-20 °C days week⁻¹. The accumulation rate is usually at least an order of magnitude lower than the Threshold temperature mechanism.

The third mechanism, Average Temperature, applies to organisms whose temperature experience is highly moderated. It was first developed in order to model the potential distribution of the Red Imported Fire Ant (Sutherst and Maywald, 2005). In that paper, Sutherst and Maywald were unable to get a satisfactory fit to the known distribution data using either of the other two mechanisms. When they considered the biology of the ants, they concluded that the

CLIMEX Model Theory

earthen nests allowed the ants to moderate the air temperatures, storing up heat accumulated during the day.

6.5.2 Heat Stress

When species overheat they generate heat shock proteins. Heat stress tends to be a fast-acting stress.

6.5.3 Wet Stress

Wet stress may be associated with waterlogging, or it may be associated with rainfall. It is unusual for the wet stress threshold to be lower than 1 (field capacity).

6.5.4 Dry Stress

Dry stress is usually thresholded at a soil moisture level of 0.1, which equates to permanent wilting point. Whilst many published models have a threshold different to this, in many cases a better fit can be obtained by changing the threshold. Some particularly hardy plants can clearly tolerate soil moisture levels lower than 0.1 before they start experiencing stress. These plants usually have specific adaptations such as drought deciduousness [e.g., *Vachellia* (syn. *Acacia*) *nilotica*], modified (and reduced) leaves, making use of photosynthetic stems (e.g., *Parkinsonia aculeata*), adopting a Crassulacean Acid Metabolic (CAM) photosynthesis mechanism (e.g., *Opuntia* spp.), having extremely deep root systems (e.g., *Prosopis* spp.), or tightly folded leaves (e.g., *Nassella trichotoma*). Animals require the presence of primary producers. Most animals that we will be interested to model will require the presence of growing plants in order to thrive. Their soil moisture threshold therefore follows a similar logic to that of their plant host(s).

6.5.5 Hot-Wet Stress

Hot-wet stress is likely associated with some form of biotic stress. This could come in the form of competition or predation. This follows the observation of Brown *et al.* (1996) that where (heat and moisture) resources are scarce, it is the ability to make use of these meagre resources that defines species range boundaries, whereas, where such resources are abundant, it is biotic factors that tend to limit species distributions. Hot-Wet stress therefore tends to be used to define the range boundary between warm Temperate and Sub-tropical regions.

Because this is a stress that is related to inter-specific factors, the thresholds are not always intuitively obvious. For example, what is the temperature and soil moisture levels where one species (or a suite of species) are better able to grow than the species being modelled. The telling factor here is that they are at lower levels than either the hot or wet stress factors.

6.5.6 Cold-Wet Stress

A cold-wet stress tends to be a biotic-mediated stress, discriminating between those species adapted to a Mediterranean climate from those adapted to a Temperate environment.

6.5.7 Hot-Dry Stress

This stress is seldom used, with heat stress and dry stress acting separately able to define a species range limits.

6.5.8 Cold-Dry Stress

This stress is seldom used, with cold and dry stress acting separately able to define a species range limits.

6.6 Model complexity and over-fitting

Deciding on an appropriate level of model complexity is often a difficult challenge, especially for inexperienced modellers. If you include too many range-defining factors in a model, the model is able to increase its specificity, you may find opportunities to “explain” additional facets of the species range. However, there is a risk that you are “explaining” sampling artefacts such as biases, biotic effects such as competition or predation, or teleological artefacts (e.g. an invasion lag). These could all lead to model projections that contain a conservative error. This phenomena is called model over-fitting. Fortunately, CLIMEX limits the number of range-defining factors that can be used, so there is less danger of over-fitting a model than when using a correlative species distribution model. Generally, in CLIMEX models interaction stresses are not required, and when present only one is likely to be needed.

6.7 Semi-automatic parameter-fitting using the genetic algorithm

In order to help CLIMEX users with the parameter fitting process, a semi-automatic parameter fitting procedure, based on a “genetic algorithm” (GA), has been included in CLIMEX. It is used to fit the Stress parameters only. This is done by defining that area of the species’ geographical distribution in which stresses are lowest (defined as the ‘Core Distribution’ (page 55)).



The GA procedure should be considered a supervised process. The resulting fitted species parameters should always be reviewed critically. Species distribution data are notoriously biased and incomplete, and hence the fitted resulting model will tend to reflect these biases. The model may have identified solutions that satisfy the matching function well, but use biologically implausible parameters. In such cases it pays to adjust the parameters to what you consider to be more realistic values and rerun the model, comparing the results with those from the automated fitting procedure. The differences in the output may suggest hypotheses that could challenge the veracity of the distribution data, the model formulation (which stress mechanisms are used), or your understanding of what is biologically reasonable. In the last case, the dominant ecological process limiting the species geographical range may not be the one that you imagined, or there may be some factor at play that alters the effect of climate (e.g. irrigation or the presence of glasshouses). Resources such as Google Earth can be very useful for exploring such issues.

In order to implement the GA procedure, you need to supply a set of localities where the species has been observed and some settings that determine how the GA operates. In the previous version of CLIMEX it was also necessary to specify a set of locations where the species was assumed to be unable to persist. The use of pseudo-

CLIMEX Model Theory

absence data is an inherently poor scientific practice. An initial set of species parameter values is also needed to initiate the fitting process. Such values can be derived from a species with a similar distribution or from one of the species templates supplied with CLIMEX. The genetic algorithm will then generate new “populations” of parameter sets and sift through them to find those that best describe the observed distribution.

General descriptions of the operation of a genetic algorithm can be found in Beasley *et al.* (1993); Renner and Ekart (2003); and Duboz *et al.* (2010). A GA is an optimization procedure that simulates natural evolutionary processes by creating ‘populations’ of parameter values (each corresponding to a hypothetical ‘species’ in nature) and then selects the best fitting values in each generation of the simulation. In each generation the GA introduces new variants into the populations by simulating genetic *cross-overs* and *mutations*. This allows new parameter spaces to be explored. After the selected number of generations, the GA provides the best fitting parameter set. If the simulation is interrupted the GA stores parameter values derived up to that time.

The fitness function is critical to the behaviour of the model. It defines the phenotype or model characteristics to which the solution will tend. Assessing the goodness of fit of species distribution models is a topic that has concerned many authors (e.g., Fielding and Bell, 1997; Anderson *et al.*, 2003; Hirzel *et al.*, 2006; Johnson *et al.*, 2006). Whilst the Area Under the Receiver Operating Characteristic Curve (AUC) is very popular amongst species distribution modellers, it has critical problems (Lobo *et al.*, 2008), especially for niche modelling purposes. Originally designed for characterising the nature of a radio antenna, the method found favour amongst statisticians for situations where models developed a probability surface for a phenomenon where there was a trade-off between model sensitivity and model specificity. In the niche modelling situation sensitivity relates to the proportion of qualified suitable locations that are modelled as suitable, and specificity is related to the prevalence, or relative area that is modelled as suitable. The calculation of the AUC involves mathematical integrating the model sensitivity and specificity for all levels of threshold for the output variable. Thus, almost all of the information in the AUC is gained from versions of the model that are irrelevant for most practical purposes. For niche models we have asymmetrical knowledge of the suitability of different locations for a species. We tend to know with high certainty that some locations are suitable, but we are far less certain of whether other locations are unsuitable, especially when we consider suitability outside of the species native range. Hence, high (or perfect) model sensitivity is a pre-requisite for a reliable model. That 99% of the information used to calculate the AUC comes from situations where the model sensitivity is less than 1 highlights how uninformative it is for judging the value of a model.

The characteristics of a good model of a species potential geographical distribution are as follows:

1. It should have high or perfect sensitivity, according with the known distribution of the species.
2. Given 1, it should have good specificity, with relatively low model prevalence.
3. It should be parsimonious (as few parameters as possible).
4. The structure should be plausible and the parameters agree with what is known about the species from other knowledge domains (e.g. experimental

observations and theoretical expectations based on physical laws and behaviours of similar species). That is, it should get the right answer for the right reasons.

In this version of CLIMEX we introduce a new information criterion, the CLIMEX Information Criterion (CXIC) that embraces the first three of the requirements above for a good model, but does not require the use of pseudo-absence data or a sampled background *sensu* MaxEnt. The CXIC shares a similarity of structure with the popular Akaike Information Criterion (AIC, Akaike, 1974), rewarding model accuracy, and penalising model complexity. In contrast to AIC, for a given set of weightings values, a higher CXIC score indicates a better quality model.

CLIMEX Model Theory

The CXIC is composed of three terms, *sensitivity*, *prevalence* (1-specificity) and *parsimony* (complement of complexity).

$$Sensitivity = \frac{N_{suitable}}{N_{input}}$$

Where $N_{suitable}$ is the number of climate stations where the Core Distribution values are > 0 , and N_{input} is the number of climate stations spatially associated with known suitable locations for the species being modelled.

$$Prevalence = 1 - \frac{N_{suitable}}{N_{selection}}$$

where $N_{selection}$ is the number of climate stations in the present run.

$$Parsimony = 1 - \frac{\max(P_{active} - 6, 0)}{P_{total}}$$

Where P_{active} is the number of parameters in the model that are activated. P_{total} is the total number of parameters activated in the model. Activated are those that are that they are turned on in the parameter grid, and have stress accumulation rates that are significantly different from zero. This allows for the GA to turn off parameters by setting rates that are effectively zero, contributing nothing to the sensitivity or specificity of the model. Counting the number of parameters used (Activated) in the model is a little tricky. Stress rate accumulation parameters largely determine the contribution of a stress function. Determining when a stress function is essentially turned off is somewhat arbitrary, and varies between stress functions depending on whether the variable is soil moisture or temperature. CLIMEX evaluates each stress function with the set of parameters used for each run and assesses whether the result varies appreciably when the input variable is adjusted. If the function is “flat” because the stress rate is near zero, it is considered to be turned off.

The terms in the CXIC are combined as follows, with each term accorded its own weighting parameter, which is set in the Genetic Algorithm Settings dialogue (Fig. 6-14).

$$CXIC = 100 \cdot \frac{W_{sens} \cdot Sensitivity + W_{prev} \cdot Prevalence + W_{pars} \cdot Parsimony}{W_{sens} + W_{prev} + W_{pars}}$$

Where W_{Sens} , W_{Prev} and W_{pars} are the weightings for the other indices, with defaults of 1, 1 and 0.05 respectively.

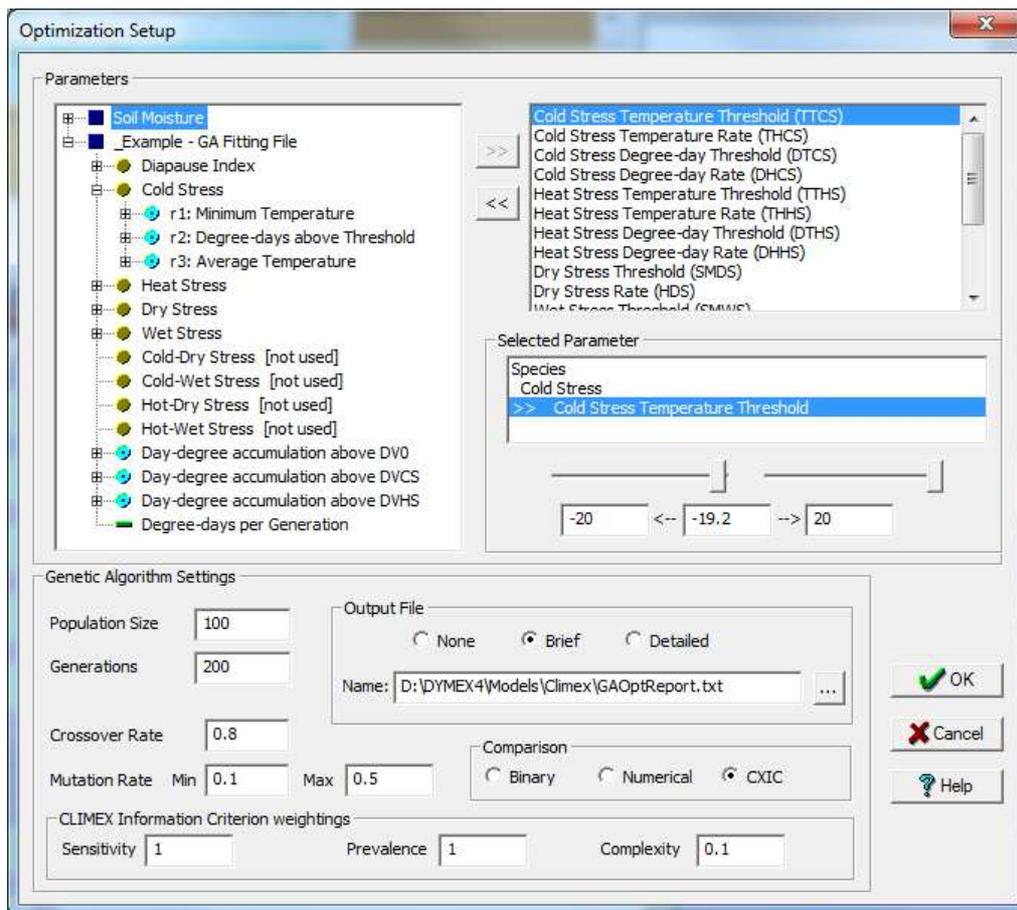


Fig. 6-14 The Optimization Setup dialog, used for configuring the settings for the automatic fitting of parameters using the Genetic Algorithm where the reference data file is a shapefile.

The steps involved in the fitting procedure are as follows:

- (1) Obtain a known distribution (the target distribution of the species) for a particular region (for example, a continent).
- (2) Create a Reference File that reflects the distribution of the species as found in step 1. The Reference File is used by the GA to judge how accurately a particular set of parameter values matches the known distribution. You can create a reference file in two ways:
 - a. Provide a point shapefile with the known suitable locations. This is the new means of specifying the species locations.
 - b. Create a file based on the Metadata selection that you want to use for your modelling. The shell for the file can be obtained by running the metadata selection with any species and saving out a csv file with only the Core Distribution variable. You then need to manually edit the file so that it contains the locations being used in the GA fitting and the status of the species (-1 = unknown suitability, 1 = present, 0 = unsuitable) for each of the locations. If relative abundances are known, these should be indicated with a range of numbers scaled between 0 and 100. The presence/absence data reflect the “Core

CLIMEX Model Theory

Distribution” of the species. This is the technique used to specify the species location information in Version 3 of CLIMEX.

- (3) Create a new species (the species to be fitted).
- (4) Decide which stress parameters are to be adjusted, and optionally set a range for each of the parameters to be included in the fitting. By default, the optimization routine will use the full allowable range of a parameter when searching for a best value. In order to save execution time (and to avoid “unrealistic” values of parameters being used by the fitting routine), it is strongly recommended that each parameter be carefully examined and an appropriate range set. Each parameter’s permitted range can be set separately.
- (5) Change the Genetic Algorithm operating characteristics, if required. By default, the routine operates with a parameter set population of 100, a “crossover rate” of 0.75, a “mutation rate” of 0.1, and will run for 250 generations before stopping.
- (6) Run the Genetic Algorithm. Note that the algorithm may take a long time to run (it requires many runs of CLIMEX), but may be interrupted at any time. If the parameters are then saved, the fitting routine may be restarted at any time with the most recently saved parameter values.

These steps are now described in more detail by an example of fitting a species (*Diuraphis noxia*, the Russian Wheat Aphid) for North America.

Step 1: Obtain a known geographical distribution for the target species.

A map of the aphid’s known distribution is shown in Fig. 6-15, which has been adapted from Figure 2 in Hughes and Maywald (1990).

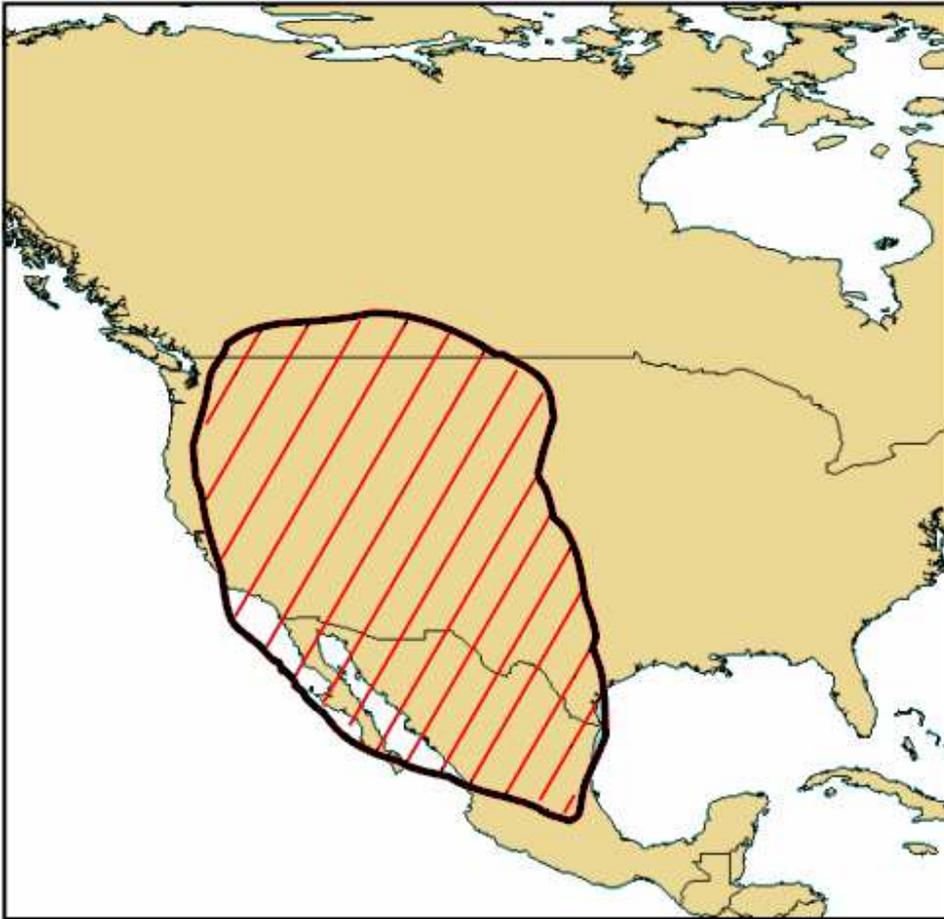


Fig. 6-15 Approximate known distribution of *Diuraphis noxia* in North America.

Step 2: Create a Reference File (Shapefile methods)

We need to present CLIMEX with a point shapefile. The suitable area for *D. noxia* in Fig. 6-15 is described by a polygon. If you can use a Geographical Information System (GIS) such as ArcGIS or QGIS (freeware) you can create a gridded point reference file by spatially intersecting a gridded dataset such as the CliMond CM10 points shapefile file that comes with CLIMEX (look in the Shapefiles directory), with a polygon shapefile.

If the data you are starting with is a set of point locations with latitude and longitude coordinates you can convert the popular csv file format to a set of shapefiles using a GIS or one of a number of converters such as the free one from http://www.mapwindow.org/apps/wiki/doku.php?id=csv_to_shapefile_converter.

Click on the **Initialization** menu and select the **Reference Data File** sub-menu. This opens the Reference Data File Setup dialog (Fig. 6-17) . Click on the small button at the top right of the dialog and in the resulting “File” navigate to the point shapefile that you have just created. Click on “Load” to load the file into CLIMEX. Depending on whether you intend to run the system using gridded climate data or from irregularly-spaced meteorological stations, you can specify how CLIMEX will search for nearby metdata locations to build the set of reference metdata. You can specify the search distance in terms of the spacing of the grid cells or a search radius. It is clearly more straightforward to make this decision in the case of gridded data. For station data you will have to decide how close is near enough. In areas of high

CLIMEX Model Theory

topographic relief and relatively dense station network, you can set the match radius fairly small (perhaps 0.2 degrees).

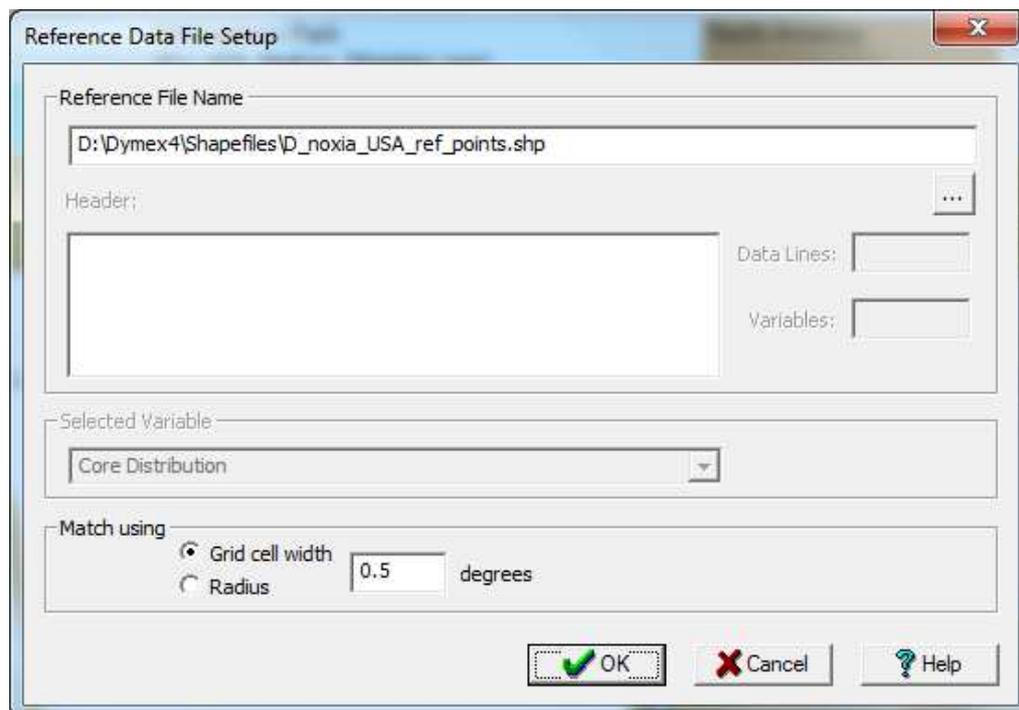


Fig 6-16 The Reference Data File Setup dialog with a shapefile of reference points for *Duraphis noxia*. Note that the Header and Selected Variable areas are greyed out because they are only relevant if you provide a Metadata based reference file (see below).

- ! You can check the stations that have been included in the reference dataset mapping the variable Reference Data. You can access this by running the Genetic Algorithm for at least one run and then stopping it manually (Fig. 6-18).

Step 2: Create a Reference File (Metadata method)

N.B. this is the most difficult method of generating a distribution file that CLIMEX can use. It should be avoided if at all possible.

We will use the CLIMEX station data (not the climate grid data) for fitting this species to expedite the GA calculations. We need to create a Reference File that contains the known occurrence of the species at each of the locations in the Location Selection that will be used in the fitting process. The most convenient way to create the reference file is to run CLIMEX with the appropriate Location Selection, and output a Results File which contains a column with the item we want, the “Core Distribution”.

For editing purposes it is easier to export a variable that has all zero values. To generate such a file, use the special “species template”, “**_Template – GA Reference File**”. Run the **Compare Locations (Stress)** application with the desired Location Set and this template, select **Results | File** from the menu and then select “**GA Reference File – Core Distribution**” from the popup menu. The resulting File Outputs dialog will be set up correctly for the required Reference File format. Click on the “Save” button, give the file a suitable name (“**Diuraphis noxia Reference**” for our example) and place it into the same directory as the Simulation (.dxs) file.

The first few lines of the resulting file should look as shown below (make sure that the header for the variable shows “Core Distribution”:

```
CLIMEX - Compare Locations (Stress Only)
_Template - GA Reference File
North America
<Climate Change scenario> / Irrigation: <Irrigation>
Run on Aug 25 2006 07:12

"Continent","Country","State","Location","Core Distribution"
North America,Barbados,,Bridgetown,0
North America,Canada,Alberta,Banff,0
North America,Canada,Alberta,Beaverlodge,0
North America,Canada,Alberta,Calgary,0
North America,Canada,Alberta,Edmonton,0
North America,Canada,Alberta,Embarras,0
North America,Canada,Alberta,Fort McMurray,0
North America,Canada,Alberta,Grande Prairie,0
North America,Canada,Alberta,Jasper,0
North America,Canada,Alberta,Keg River,0
```

The next part of the procedure is rather laborious. It involves replacing the default zero value of the “Core Distribution” for each location with one of the following values:

- 0 the species is known not to occur at that location
- 1 the species is known to occur at that location
- 1 the status of the species at that location is unclear

The changes can be made using programs such as Excel, Notepad etc. but will be much faster if there is a ‘global replace’ feature to substitute values into blocks of lines rather than editing each line separately which may be prohibitively time consuming.

Several things need to be borne in mind when coding this file. Distributions are generally poorly defined when using sources such as museum records. The fact that a species is not recorded at a particular location can mean either that it is not there or has not been found there yet. The danger is that if all such locations are coded as 0 (absent), the genetic algorithm will produce a modelled distribution that is too conservative (i.e., it will indicate some locations as unsuitable even if they are suitable for the species). This is a problem in all such automatic fitting procedures. Of course, such locations could be coded as -1 (species status unknown). This cannot be done for all such locations if you are going to use the Binary fitting method, however, as this fitting algorithm needs some negative records to be able to fit a boundary to the distribution. Careful consideration therefore needs to be given to how locations are coded, especially around the boundaries of the known distribution. It is recommended that the genetic algorithm be used to give a “ballpark” fit initially, with careful consideration of the data used in the fitting and manual fitting used to complete the process.

Once the file has been edited, it can be checked by loading the “**Compare Locations (Genetic Algorithm)**” application again and setting this file as the Reference Data File. Click on the **Initialization** menu and select the **Reference Data File** sub-menu. This opens the Reference Data File Setup dialog (Fig. 6-17) . Click on the small button at the top right of the dialog and in the resulting “File” navigate to the

CLIMEX Model Theory

reference file that you have just created. Click on “Load” to load the file into CLIMEX.

The file will be read during the loading operation, and if its format is correct, the file “Header” (i.e., the lines preceding the data) will be shown in the **Header** panel (Fig. 6-17). The two boxes to the right of the Header panel will be showing the total number of lines containing data in the file, and the number of variables that the file contains. If the data file does not have the correct format, DYMEX will display an error message indicating that a problem has occurred. Should all be well, select the required variable (“Core Distribution”) from the **Selected Variable** box by clicking on the small arrow to the right and locating it in the drop-down list.

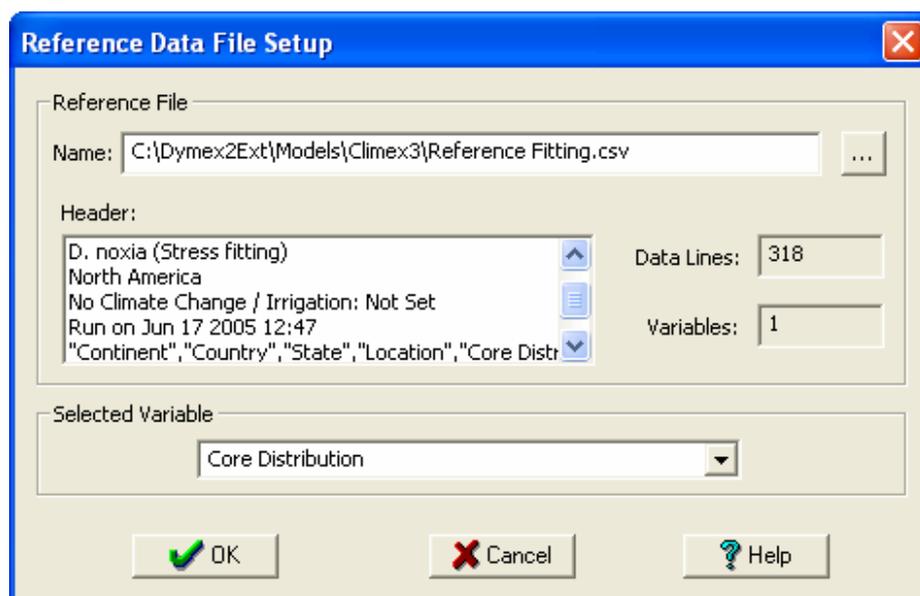


Fig. 6-17 The Reference Data File Setup dialog, with settings for the *Diuraphis noxia* fitting.

With the Reference File correctly loaded, CLIMEX can now be run (start the genetic algorithm for at least one run and stop it manually) to check the data in the file. With the run completed, map the data in the file by creating a map using the output variable **Reference Data** (Fig. 6-18).

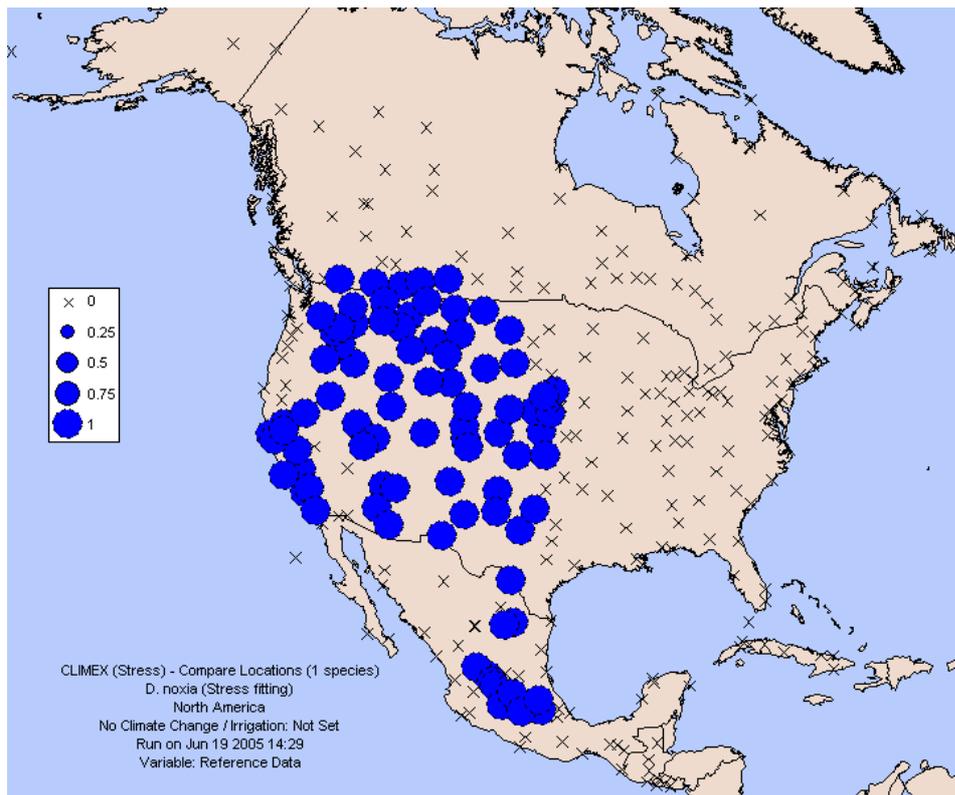


Fig. 6-18 A map of the data in the Reference File.

Step 3 (Create a new species):

Load CLIMEX (always use the **Genetic Algorithm** version of CLIMEX) and create a new species (see Section 8.2) that contains the initial ‘seed’ parameter values (these represent the first guess at a parameter set and will usually be one of the template “species” or a species with a known, similar distribution). Use that species for the fitting procedure.

Step 4 (choose GA operational characteristics) and 5 (select Stress parameters to be optimized):

We are now ready to set up the optimization run using the “genetic algorithm”. This must be done using the “Genetic Algorithm” application of CLIMEX, **CLIMEX – Compare Locations (Genetic Algorithm)**. This allows the core distribution to be fitted without the additional complications of the Growth Index, and speeds up running of the optimization. To set up the optimization run, click on the **Initialization** menu and select **Genetic Algorithm...** from the drop-down menu. This opens the **Optimization Setup** dialog (Fig. 6-19).

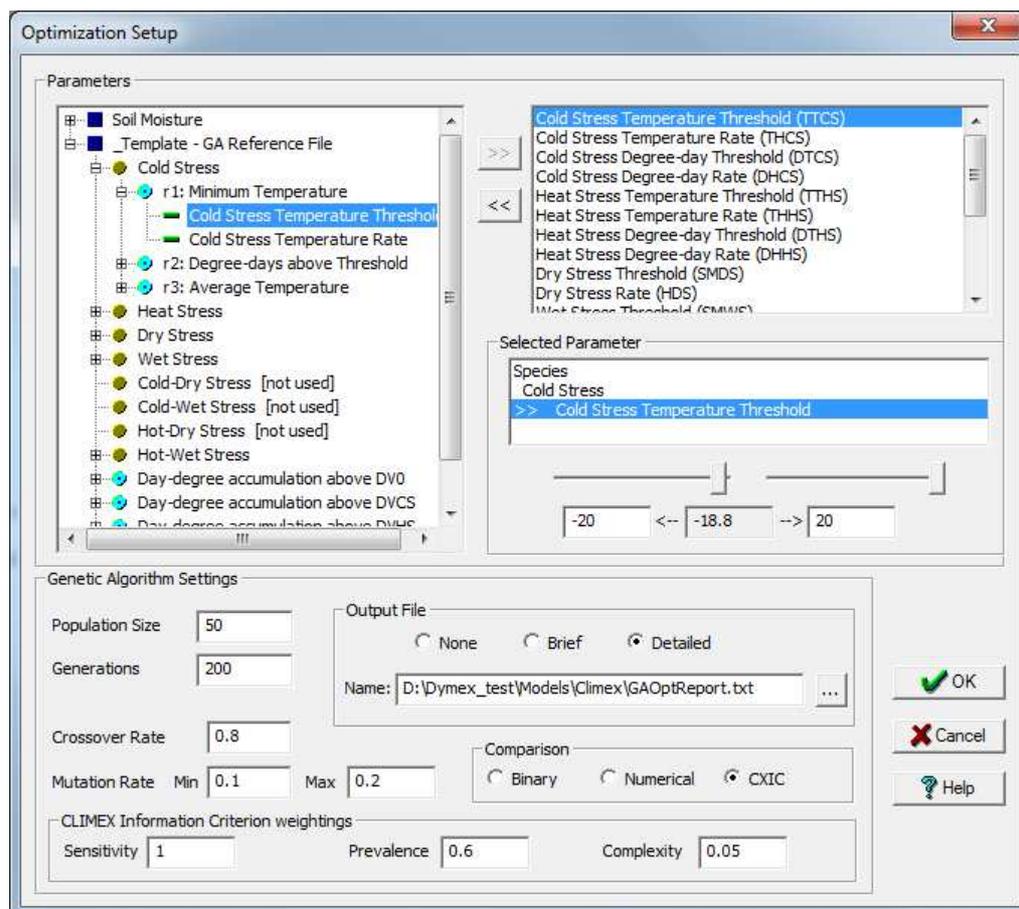


Fig. 6-19 The Optimization Setup dialog, used for configuring the settings for the automatic fitting of parameters using the Genetic Algorithm.

The first step is to select the parameters to be optimized with their upper and lower limits for the target species. The parameters are listed at the left, in a “tree” view, where various CLIMEX model components form the nodes of the tree, with the parameters forming the terminal nodes. This allows the “context” of a parameter to be determined. Nodes can be “opened” in this view by clicking on the small rectangle containing a + in front of the node symbol. The list box at the top right of the dialogs contains the parameters that are included in the fitting. To include a parameter, select it in the **Parameters** window and click on the **>>** button to move it into the list box. As many parameters as required can be fitted in the optimization run. The sliders and edit boxes in the **Selected Parameter** panel allow adjustments of the range over which the highlighted parameter can be adjusted by the GA. If a panel and slider are greyed out, the corresponding limit is not adjustable. When a parameter is first included in the fitting, the adjustment range is set to the full range allowed by the CLIMEX model for that parameter.



The diapause parameters do not include an accumulation rate as the stress parameters do. This means that the GA has no means of turning it on or off. The GA can simply adjust the thresholds for commencing and terminating diapause, or setting the minimum duration, which has the effect of making it obligate or facultative. If you are unsure of whether diapause is relevant for your species you may have to experiment, running the GA with and without the diapause parameters included in the run.

The operating parameters for the genetic algorithm are set from the Genetic Algorithm Setting panel at the bottom of the dialog. **Population Size** indicates the number of parameter sets being operated on at any one time, **Generations** specifies the number of generations for which to run the optimization, **Crossover Rate** is the crossover rate to be used, and **Mutation Rate** specifies the lowest and highest mutation rates. The mutation rate declines linearly from the specified maximum value when the ratio of the standard deviation of the population's fitness to the average is above 0.1, down the specified minimum value when the ratio is 0.

You can also select the format used to compare the data in the Reference file with the estimated values. It can be either **Binary**/categorical (representing presence or absence) or **Numerical**. In addition, there is a choice of format and destination for the **Output File** (which contains information about the fitting process). You are now ready to create your new species file and run the Genetic Algorithm.

Step 6:

The Genetic Algorithm can now be run by selecting **Run Genetic Algorithm** from the **Execution** menu. Such a run can take a long time (many hours), but can be stopped at any time by clicking on the **Stop** button. The species will then contain the best combination of parameters found so far, and these parameters can be saved in the usual way – overwriting the starting species parameter values - from either Parameter Window. The optimization run can then be restarted with these parameter values at a future time. After saving the parameter values, close the simulation, return to the simulation window and run the species as usual. You are then able to compare the Core Distribution with the Reference Data by mapping the two variables simultaneously.

CLIMEX Model Theory

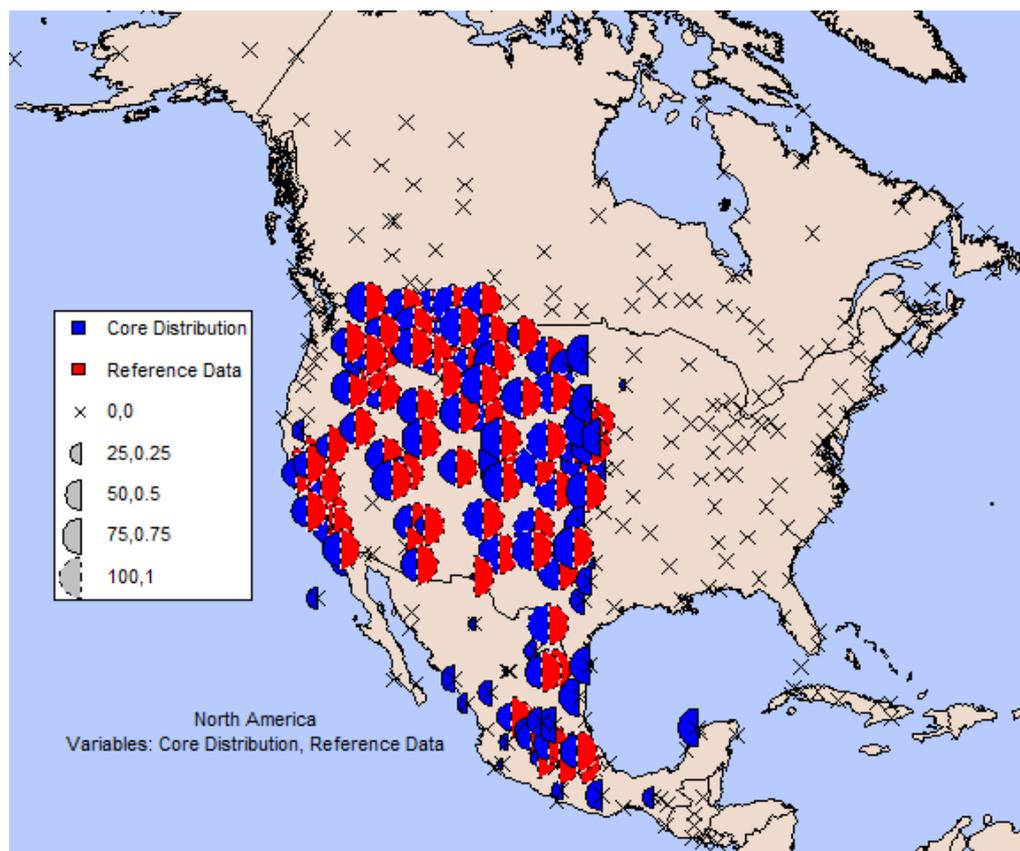


Fig. 6-20 Results of a 200 generation run, showing the modelled 'Core Distribution' of *Diuraphis noxia* in blue using the fitted parameter values and the Reference Data points for *D. noxia* (in red).

A run for 200 generations, with a population of 100 and all other settings as in Fig. 6-14 produced the results shown in Fig. 6-20. The run results window summarises the goodness of fit of the model (Fig. 6-21).

Run 28	
Model	CLIMEX - Compare Locations (Genetic Algorithm)
Run on	Feb 03 2015, 15:07
Parameter File	D:\.\Climex\Compare Locations (Genetic Algorithm).gmp
Run Type	Supervisor (Genetic Algorithm)
Description	Comparison method: CXIC Generations completed: 3 Best fitness: 78.87 Sensitivity: 0.988 Prevalence: 0.6183 Complexity: 0.5

Fig. 6-21 The run results window for a CLIMEX Genetic Algorithm run.

There are some important steps you should take before you insert the map into a report or journal paper. You should check the fitted parameters systematically to assess whether they are biologically reasonable. If some parameters look dubious, you should force them to take what you consider reasonable values, re-run the model and compare the results with those that CLIMEX fitted to the original data. The

discrepancies between these results reveal a conflict between the distribution data, the climate data, the stress mechanism(s) used in the model and your understanding of the species ecology. These are the usual suspects, and your task as a modeller is to investigate the conflict and decide how to resolve it. It may be helpful to think of this as applying the method of multiple working hypotheses (Chamberlin, 1965), whereby you have to consider the weight of evidence in favour of different sources of error or conflict.

CLIMEX has done some of the work in fitting the stress parameters to the geographical location data you gave it. However, location data often contains errors, omissions or otherwise misleading points. See section 6.3 for a discussion of some of these error types.

It is possible to overfit a model. In CLIMEX it is possible to identify many of these issues and deal with them appropriately. Because the stress mechanisms and parameters have ecological meaning, it is possible to check them in relation to your knowledge of the species ecology. This knowledge can come from scientific studies of the ecology or eco-physiology of the species itself or from similar species.

6.8 Parameter Sensitivity Analysis

There are many factors that affect model results (e.g., the structure of the model, the input data used to train it, and of course, the parameters). Each of these factors contributes to uncertainty in our model results. In recent literature there has been an increasing interest in trying to quantify and present uncertainty in model results, especially for potential distribution models presented in pest risk assessments (Venette *et al.*, 2010; Makowski, 2013; Yemshanov *et al.*, 2013). In order to assess which variables we should be most concerned about in terms of their accuracy, we can employ a parameter sensitivity analysis. In these analyses selected parameters for the fitted model are perturbed or adjusted by a defined amount either side of the fitted value, and the effect on the model results is gauged. Variables that are found to have a large effect on the model results when they are perturbed are said to be *sensitive*. These are the variables that we need to be most concerned about in terms of their accuracy. *Insensitive* variables are those that can be adjusted considerably before they have an important effect on the behaviour of the model. We may find that we have sensitive variables in the model that we are unable to specify with much certainty. In such cases, responsible modelling practice dictates that we should explore the effects of the uncertainty, and present the results. Knowledge of the uncertainty and its source can lead to better decision-making. If the stakes are sufficiently high, a decision can be made to invest in further research to increase our confidence in the uncertain and sensitive parameter.

It is generally inadvisable to choose a single state variable as a good indicator of model behaviour (Miller, 1974). In CLIMEX we present a table of state variables and their sensitivities in relation to the parameters that were included in the analysis.

CLIMEX Model Theory

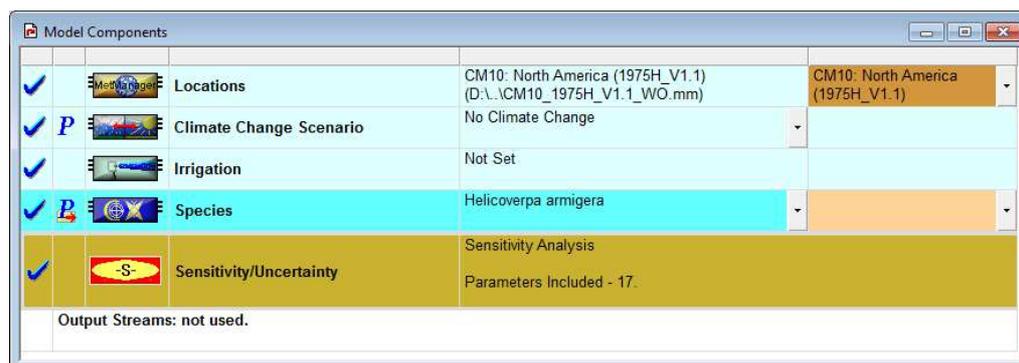


Fig. 6-22 Model Components window for sensitivity analysis

Because the variables in CLIMEX are anisotropic, it is unreasonable to adjust them by a fixed percentage: a ten percent adjustment of a soil moisture index parameter has a substantially different impact on a species modelled ability to grow or persist compared with a ten percent difference in temperature.

Following consultation of a range of experienced CLIMEX users, we settled on the following ranges for the default perturbations:

- Temperature +/- 1 degree
- Soil moisture +/- 10%
- Rate parameters +/- 10%
- Degree day temperature thresholds +/- 1 degree
- Degree Day Sum (PDD) +/- 20 degree days
- Diapause temporal thresholds (+/- 0.5 hour), and
- Diapause durations (+/- 7 days)

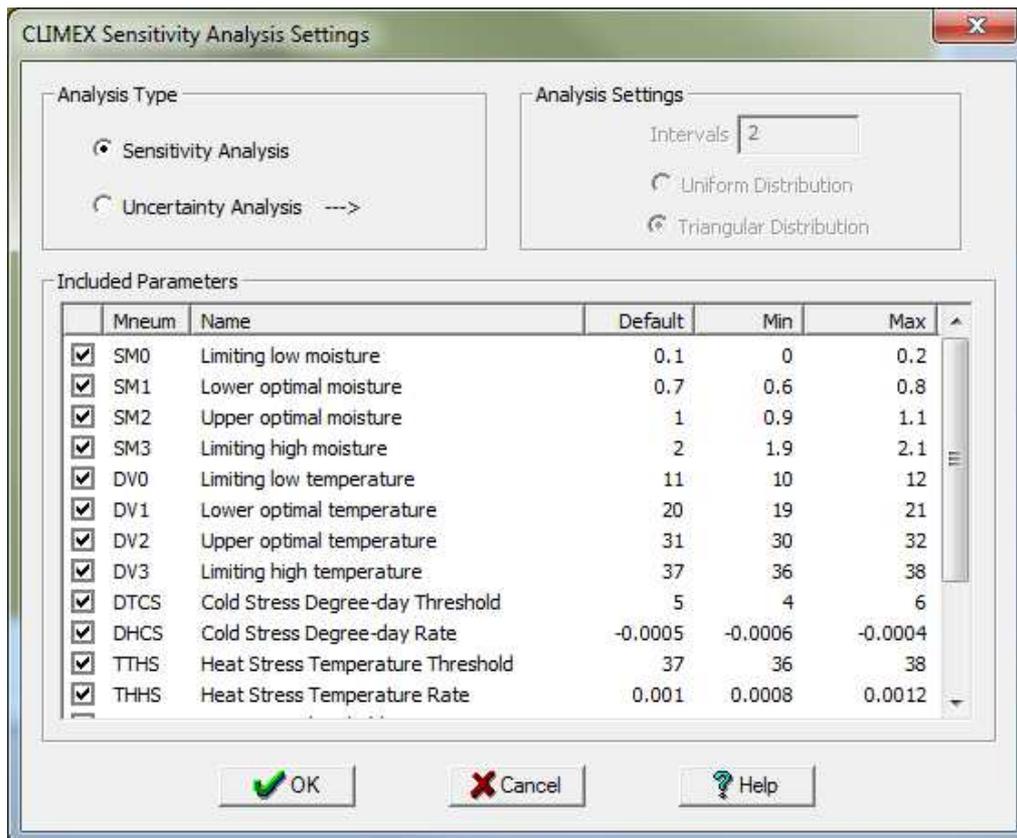


Fig. 6-23 Initialisation settings for Sensitivity Analysis

The selection of parameters to be included in the analysis is controlled using the Sensitivity Analysis Settings window (Fig. 6-23). The default perturbation values can presently be manually over-ridden by the modeller, but it requires manual editing of a CLIMEX include (*.gmi) file. The values were selected based on the *reasonable* uncertainty in model parameters, considering the interplay between geographical accuracy in species location records and splining errors in climate datasets. In order to be useful, these perturbation values need not be precise or *accurate*, but rather, they should be *meaningful* for modellers. In selecting these perturbation ranges we tried to identify the level of precision that would normally be attributed to each parameter type, and which would normally be explored manually once a model was nearly finished.

! In CLIMEX (and DYMEX), parameters for rate variables that are negative (e.g., Cold Stress or Dry Stress accumulation rates) are considered to be lower in value, the more negative they are.

This is illustrated in Fig. 6-23 where the minimum value for the Cold Stress Degree-day Rate is -0.0006 and the maximum value is -0.0004.

In the sensitivity analysis, the model is run for a scenario where each parameter is adjusted upward and then downward by the relevant perturbation value, and the effect on the species range and a set of selected state variables (EI, GI_A, etc.) is gauged against each variable by creating difference maps and integrating across them appropriately.

CLIMEX Model Theory

The sensitivity function is a simple average deviance function:

$$S = \frac{V_i^+ - V_i^-}{V_i^0}$$

Where the subscripts denote the increase, decrease or default value of the parameter i used to generate the variable V . The value of S increases as the model changes to a greater extent when the parameters are perturbed away from the fitted values V_i^0 .

The output also includes a table listing all of the parameters, the amount by which they have been adjusted upward or downward, and what effect this adjustment had on each of the selected state variables (Fig. 6-22).

The analysis generates a set of variables describing the change in relevant model state variables. Most of the variables (e.g. EI Change, CD Change, GI Change MI Change, TI Change CS Change, HS Change, WS Change, DS Change) describe the mean sum square change in the value of the Index (e.g. Ecoclimatic Index) for all of the locations included in the sensitivity analysis. The Range Change is unique in that it is the percentage of locations whose EI changes about the threshold value. Taylor and Kumar (2012) calculated the coefficient of determination R^2 to assess the level of agreement between model runs for the sensitivity analysis. We feel that it is more informative to present the mean sum of squared differences.

There are patterns in the sensitivity results that are expected, and probably unremarkable. The stress variables should each be relatively sensitive to the corresponding parameters that gave rise to them. The Growth Index (GIA) should be sensitive to the soil moisture and temperature growth parameter sets. For most applications the relative sensitivity of the Range Change, EI Change and CD Change variables to different parameters (in decreasing order to priority) will be of most interest.

Fig x (sensitivity analysis output)

6.9 Parametric Model Uncertainty

Where the sensitivity analysis is focused on the sensitivity of each parameter on a one-at-a-time basis, there could well be strong interactions between parameters that will be missed using this approach. In CLIMEX, uncertainty analysis involves simultaneously selecting random values from either a triangular distribution or a uniform distribution using a Latin hypercube sampling scheme. This technique is attractive for gauging overall levels of model uncertainty in a very efficient manner (Helton *et al.*, 2005).

As with sensitivity analysis, uncertainty analysis relies on the analyst specifying both perturbation parameters and a set of state variables against which variations are gauged. In CLIMEX, the uncertainty analysis generates a set of state variables:

- a) mean goodness of fit*

- b) SD goodness of fit*
- c) mean sensitivity*
- d) SD sensitivity*
- e) mean proportion of climate stations that are suitable (as an index of model specificity)
- f) SD proportion of climate stations that are suitable (as an index of model specificity)
- g) the number of locations in the climate selection

We will also generate maps for each state variable e.g., average values across 100 runs and a map of the proportion of runs where the models are above a threshold (e.g., $EI > 0$).

6.10 Assessing the goodness of fit of a model

Relying on a single statistic to assess the goodness of fit of an ecological model is like trying to judge someone's character based on their residential postcode. Despite quite a lot of attention to the matter, there is no universal goodness of fit metric that can be used to assess ecological models of species niches or their distributions. Models need to be evaluated in relation to their purpose and their meaning. If a model is built to address a specific question, the nature of that question and the consequences of different errors need to be borne in mind when evaluating the model. For example, in pest risk analyses, the consequences of underestimating the potential distribution of the pest in a new region could be that the organism is allowed to spread there in the absence of appropriate biosecurity efforts to prevent the spread. These consequences are probably larger than the costs of biosecurity efforts applied when they are unnecessary. In this case model sensitivity is to be preferred over model specificity.

In most biosecurity applications, calculating and interpreting model specificity is particularly problematical. The notion of model specificity has an assumption that we know with certainty a number of sample points where we are sure of habitat suitability and others where we are sure of habitat unsuitability. However, there is an additional implicit assumption that both of these samples are unbiased. If we consider a map of climate types across the globe (e.g., Kottek *et al.*, 2006) it immediately apparent that similar patterns of climate appear in widely dispersed locations. This patterning underpins the greatest part of both invasion biology and the international spread of agriculture. Because of natural dispersal limitations, species tend to evolve within one geographical region, spanning one or a few bioclimatic zones. When they are introduced into a similar climate in a different region, we find that many of them are able to become established and spread in the new region. If we were to build a bioclimatic niche model of such a species when it was restricted to its native range

The frequently cited AUC (area under the receiver operating characteristic curve) was one attempt to capture the overall accuracy of a model by simultaneously considering model *sensitivity* (the proportion of known positive records correctly modelled) and *specificity* (the proportion of known negative records correctly modelled). Unfortunately it suffers from a number of problems, including the

CLIMEX Model Theory

threshold problem and the need to invent species absence data (so-called pseudo-absence data). The use of pseudo-absence data is an act of statistical desperation. The pressure for it comes from the fact that with species distributions we can be much surer that we have observed a species to persist at a location than we are that a species cannot persist at a location. Using pseudo-absence data is bad science. It is in effect a set of hypotheses that are never tested, and worse, these assumptions get transformed into data that is used to both train the models and to test them. This is a logical circularity. The *presence-background* methods such as MaxEnt (frequently described misleadingly as presence-only models) typically embed the selection of the pseudo-absence points into an automated sampling procedure within a defined background. The definition of the pseudo-absence data in this case is through the definition of the region encompassing the background. The effect is nonetheless the same; the resulting models are highly sensitive to the definition of the background.

In the new version of CLIMEX we introduce the calculation of a set of statistics for the Compare Locations models. Considered together, these statistics help describe some of the model behaviour in relation to the known distribution.

- Sensitivity
- Prevalence
- Complexity
- CXIC
- Significance

However, as a modeller you should be mindful that the framing of a model can have far greater impact on a model than the statistics. The statistics describe how well the model matches the data that you have presented it. What would be the difference between a model fitted to data that includes points where a species is able to persist by virtue of irrigation compared to one where it is only presented with “natural rainfall” records? Similarly, what if the location records included locations where the species was observed during annual migrations compared to those locations where it was observed to be able to persist throughout the year. Clearly, assessing the performance of a model also demands an understanding of the framing and meaning of the model.

Invasive species pose particular problems for assessing the goodness of fit of niche models. Firstly, the pest risk assessment challenge requires us to project the potential distribution of species outside of their native range. The Enemy Release Hypothesis (Davis *et al.*, 1998; Keane and Crawley, 2002) alerts us to the fact that species may behave differently outside of their native range, typically expanding their range into warmer and wetter habitat than they occupy in their native range (Wharton and Kriticos, 2004). If a model is being fitted to a species native range and part of its invaded range, there may be significant invasion lags apparent in the invaded range. That is, there may be areas that are climatically suitable, but otherwise devoid of observation records simply because the invader has not yet spread throughout its suitable climatic range on that continent. Invasive species may also find limited hosts in an invaded range, depending on the distribution of, say, cultivated host crops. In these latter cases, the absence of location records does not have the same putative meaning as an absence of records in the native range, where we might expect the record to be more complete. For invasive species, the cost of

underestimating the potential distribution could lead to an unwanted incursion, and very expensive impacts on agriculture, biodiversity and agriculture (Pimentel, 2002). Hence, we should be wary about placing too much emphasis on model specificity for invasive species niche models, preferring a model that is sensitive and logically defensible.

7 Matching Climates

The 'Match Climates' applications compare the long-term meteorological data for each of the selected 'Away' locations with the climate of the 'Home' location. The level of similarity is given by the 'Composite Match Index, CMI', which is the product of six component indices, I_{tmax} , I_{tmin} , I_{rain} , I_{rpat} , I_{hum} and I_{sm} , indicating similarity in maximum and minimum temperature, total rainfall, rainfall pattern, relative humidity (RH) and Soil Moisture, respectively. Each of these component indices can range between 0 and 1, with a value of 1 indicating an exact match with the 'home' location. All indices can be weighted individually (using Composite Match Index Weights) to emphasise the more important variables. The default weight setting is 1.0 for each index, except the RH index that has a default weighting of zero.

Comparison of the climates of a home location can be made using a country, continent or the global database during a single run. When data from locations in different hemispheres are compared, CLIMEX automatically displaces the data from the Southern Hemisphere location by six months. The resulting Climate Match Indices can be displayed in maps, tables or as graphs. When maps are created, they show the degree of similarity as circles whose size is proportional to the difference between the value of the selected Match Index and a selected threshold (the Match Level). Only locations with an index that exceeds the Match Level are displayed in the resultant maps and tables. In the standard Simulation File, a selection of threshold Match Levels (0.6, 0.7, 0.8 and 0.9) is available for maps of the Composite match Index, while the Match Level for the **Standard** table is set at 0.5.

Parts of the year can also be masked out so that only certain weeks are compared. This way, comparison of the similarities of a season or month rather than the whole year can be made. If comparison of locations in relation to the distribution of a waterweed is to be made, rainfall and humidity will not directly influence survival of the weed. In this case temperature is the driving variable and locations with a similar temperature range and pattern will need to be identified. By masking out the rainfall and humidity indices, a match on temperature alone can be made.

Index	Weight	Default Value
I_{tmin}	W_1	1
I_{tmax}	W_2	1
I_{tav}	W_3	0
I_{rain}	W_4	1
I_{rpat}	W_5	1
I_{hum}	W_6	0
I_{smst}	W_7	0

The calculations used to derive the Match Indices are given below.

Maximum and Minimum Temperature Match Indices, I_{tmax} , I_{tmin} and I_{tav}

$$I_{tmax} = \exp(-k_T T_{dmax})$$

$$I_{tmin} = \exp(-k_T T_{dmin})$$

$$I_{tav} = \exp(-k_T T_{dav})$$

T_{dmax} , T_{dmin} and T_{dav} are the means of the weekly absolute differences in maximum, minimum and average temperatures, respectively, between the two locations. The constant, k_T , is set to 0.1, thus giving values for the index of 0.9 for a mean weekly difference of 1 °C, and 0.6 for a mean weekly difference of 5 °C.

Total Rainfall Match Index, $I_{rtot} = \exp(-k_R R_d)$

where, $R_d = \text{abs}(R_T - R_M) / (1+a (R_T+R_M))$,

R_T = annual rainfall at Home location and

R_M = annual rainfall at Matching location.

R_d is the difference in annual rainfall between the two locations, adjusted so that a (say) 100 mm difference in rainfall is more significant for locations with lower rainfall. CLIMEX uses values of 0.001 and 0.004 for the constants a and k_R , respectively. With these values, a difference in rainfall of 200 mm per year between two locations will result in an index of 0.64 if the average rainfall for the two locations is 400 mm, and a value of 0.85 if the average rainfall is 2000 mm.

Rainfall Pattern Index, $I_{rpat} = \exp(-k_P R_D)$

R_D is the mean of the absolute difference between the weekly rainfall of the 'Home' and 'Away' location, after the weekly rainfall of the 'home' location has been multiplied by R_z ($R_z = R_M/R_T$). R_M and R_T have the same meanings as in the previous section. The constant, k_P , is set to 0.005, thus giving values for the index of 0.9 for a mean weekly difference of 20 mm, and 0.6 for a mean weekly difference of 100 mm.

Humidity Index, $I_{hum} = \exp(-k_H H_d)$

H_d is the mean of the weekly absolute differences in relative humidity between the two locations. The constant, k_H , is set to 0.024 so that a difference of 10% gives an index of 0.8 and a difference of 20%, 0.6.

Soil Moisture Index, $I_{sm} = \exp(-k_{SM} S_d)$

S_d is the mean of the monthly absolute differences in estimated soil moisture between the two locations. The constant, k_{SM} , is set to 1.2 so that a mean weekly difference of 0.1 in soil moisture gives an index of about 0.9 and a mean weekly difference of 0.2 gives an index of about 0.8.

The combined **Temperature Index**, $I_t = \frac{I_{min} \times W_1 + I_{max} \times W_2 + I_{av} \times W_3}{W_1 + W_2 + W_3}$

The combined **Moisture Index**, $I_m = \frac{I_{rtot} \times W_4 + I_{hum} \times W_6}{W_4 + W_6}$

Rainfall Pattern Index, $I_{rp} = I_{rpat} \times W_5$

Soil Moisture Index, $I_{sm} = I_{smst} \times W_7$

The total '**Composite Match Index**' (CMI), is then calculated as

$$CMI = \left(I_t^{wt} \times I_m^{wm} \times I_{rp}^{W_5} \times I_{sm}^{W_7} \right)^{1/(wt + wm + W_5 + W_7)} \times 100$$

Match Climates

where **wt** is the maximum value of W_1 , W_2 and W_3 , and

wm is the maximum of W_4 and W_6 .

7.1 Matching climates using Region-fitting

The regional climate matching application (**Match Climates (Regional)**) is an extension of the **Match Climates** application. In effect, it takes the **Match Climates** application and runs it for many 'Home' locations. The results of all of these runs are combined by retaining the best result for each location from all of the runs.

If you have a region of interest, such as New Zealand, and you are responsible for assessing the risk of imported goods carrying live organisms that could pose a threat to New Zealand, you might like to know those regions of the world with a similar climate to anywhere in New Zealand. The region matching application allows you to select a set of locations, and to identify the best match between each of the locations in the 'Away' location selection, with all of the locations in the 'Home' region (e.g., Kriticos, 2012).

Another use of the region-fitting function is to refine the search area for biological control agents, to narrow the search area to zones that are climatically similar to an area into which you wish to import agents (e.g., Robertson *et al.*, 2008). By overlaying the native range of the target host with the regional climate match data, you can then concentrate your search efforts to the zone of overlap, confident that an agent found there is more likely to be able to persist in the introduced range should it be established there, compared with individuals or species collected from outside of that range.

Setting up a region-matching run involves several steps: preparing the location set, setting options in the **Match Climates (Regional)** application, select the weightings, and then running the analysis.

7.2 Preparing the Location Set

There are two means of specifying the location selection for the Home region. You can use the location selection procedure within the MetManager utility to develop a selection using a combination of filters, or you can create a comma-delimited text file. The latter option gives you the flexibility to select locations based on attributes other than those available in MetManager (continent, country, state, location name, bounding box, altitude or importance weighting). This latter option will probably be of critical importance for gridded climate datasets that do not have any of the geopolitical information associated with the location records.

To specify the location set you need to prepare a comma-separated input file containing locations from a MetManager file and a column containing a character string identifier that allows you to specify which locations to include in the run. The file must be in a specific format, with a line containing the strings "Continent", "Country", "State", and "Location" in the first four columns, indicating the start of the region specification section (line 8 in the example shown in Fig. 7-1). Any lines preceding this "label" line are ignored by CLIMEX but may be included to indicate the origin of the locations or any other relevant information. The locations which can be included in the run then follow the "label" line, one location per line. The first four columns in each of these lines contain the names of the continent, country, state and location that identify the specific location. The fifth column contains the

indicator that will allow you to select location subsets for inclusion in a model run. You may wish to use descriptions of the classes that are described by the classification character(s) in column five, as shown in Fig. 7-1. You may add additional information in the sixth and subsequent columns. Any such extra information will be ignored by CLIMEX for region fitting runs.

The file can be built manually using a text editor, or Excel (saving the output in comma-separated (csv) format. It can also be created by running a **Match Climates** or **Compare Locations** simulation and saving out the results in csv format with just the headers (“Continent”, “Country”, “State”, and “Location”) and a single output variable included (the file output format “EI” in Compare Locations will create such a file). You can then import this file into a text editor or Excel in order to replace the values in the 5th column with the identifying character or string.

Fig. 7-1 The first few lines of a file specifying sets of locations constituting “regions” for use in the Match Climates (Regional) application.

	A	B	C	D	E	F
1	Scientific name					
2	Common name					
3	Hosts					
4	Distribution					
5	EPP0 PRA					
6						
7						
8	Continent	Country	State	Location	LENZ	
9	CRU_World	No_countries	nostates	66679	R	
10	CRU_World	No_countries	nostates	66680	R	
11	CRU_World	No_countries	nostates	66681	R	
12	CRU_World	No_countries	nostates	66710	Q	
13	CRU_World	No_countries	nostates	66711	R	
14	CRU_World	No_countries	nostates	66712	R	
15	CRU_World	No_countries	nostates	66713	R	
16	CRU_World	No_countries	nostates	66742	O	
17	CRU_World	No_countries	nostates	66743	O	
18	CRU_World	No_countries	nostates	66744	Q	
19	CRU_World	No_countries	nostates	66745	Q	
20	CRU_World	No_countries	nostates	66747	R	
21	CRU_World	No_countries	nostates	66778	O	
22	CRU_World	No_countries	nostates	66780	L	
23	CRU_World	No_countries	nostates	66781	Q	
24	CRU_World	No_countries	nostates	66782	Q	
25	CRU_World	No_countries	nostates	66783	R	
26	CRU_World	No_countries	nostates	66784	R	
27	CRU_World	No_countries	nostates	66809	Q	
28	CRU_World	No_countries	nostates	66810	N	
29	CRU_World	No_countries	nostates	66811	Q	
30	CRU_World	No_countries	nostates	66812	Q	
31	CRU_World	No_countries	nostates	66813	R	
32	CRU_World	No_countries	nostates	66835	Q	
33	CRU_World	No_countries	nostates	66836	Q	
34	CRU_World	No_countries	nostates	66837	Q	
35	CRU_World	No_countries	nostates	66838	N	

Header

0.5 degree grid of points intersected with environmental classes

If you wish to select regions using geographic data, then you can use a Geographical Information System (GIS) to intersect the output file with a polygon or grid layer that contains geographic region information. To do this, first output a file from a **Match Climates** run or a **Compare Locations** simulation and include latitude and longitude in the output file. Save the file out as using csv format. To import the file into the GIS, you will normally need to delete the first seven lines in the location set header, leaving the column headings. Then import the file as a set of points, using the “longitude” and “latitude” columns as the X and Y coordinate source respectively. You can then intersect the resulting point file with your regional data source in order to identify the corresponding region for each metadata location. This file will then need to be output as a csv file in the above format, deleting columns as necessary.

Match Climates

7.3 Setting the Region Match Options

Start the DYMEX program and open the “CLIMEX – Match Climates (Regional)” model using the **MC (Regional).dxs** simulation file (see Section 8.1). The application’s modules (Fig. 4-1) will now be shown in the Model Components window.

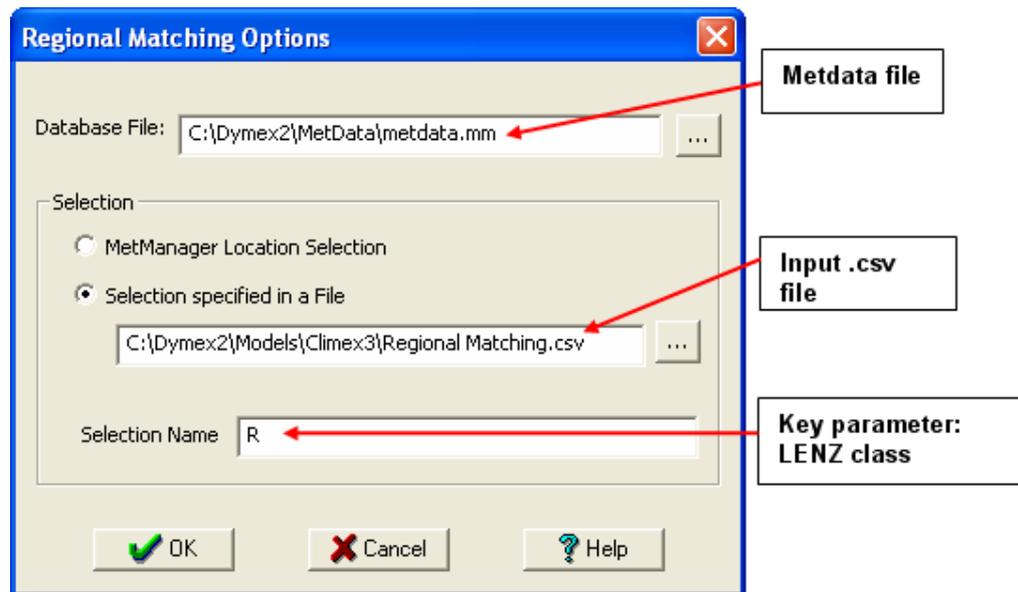
From the **Initialization** menu, select the **Regional Climate Matching** menu item. In the resulting **Region Matching Options** dialog, set the appropriate database file and Home Location Selection.

If you want to use the MetManager location selection method, click on the **MetManager Location Selection** option, and then choose the appropriate Selection Name from the dropdown list. If you have prepared a text file of the locations, click on the **Selection specified in a file** option and specify the name and location of the “csv” file. If necessary, use the **...** button and then navigate to the Location Selection file.

Fig. 7-2 The Regional Matching Options dialog, showing use of a Location Selection to specify the set of locations constituting the Home Region.



Fig. 7-3 The Regional Matching Options dialog, showing use of a Selection Name to specify the locations in a file constituting the Home Region.



When you are satisfied with the Regional Matching Options, initialize the remaining model components, if required. Each of the remaining model components can be set in exactly the same manner as for a single location **Match Climates** run. For example, the Home and Away locations can each be modified for climate change and irrigation scenarios, and the Away location dataset can be set separately from the Home Region selection.

When you have set all of the remaining options, run the climate match analysis by selecting the **Execution | Run Regional Climate Matching** menu item, or click on the double lightning bolt symbol  indicating that you want to complete a multiple location run.

7.4 Running and displaying the results

When the run is completed, a window giving a summary of the process parameters is displayed. Obtaining output is exactly the same as for match climates or Compare Locations runs.

Run 2	
Model	CLIMEX - Match Climates (Regional)
Run on	Jul 26 2006, 10:58
Parameter File	D:\Dymex2\Models\Climex3\Match Climates (Regional).gmp
Run Type	Single (104 weeks)
Description	Database: D:\Dymex2\MetData\CRU_HD\CRU_HD_61-90_World.mm Selection: D:\Dymex2\Models\Climex3\NZlocation.csv (A)

If desired, results can be exported from CLIMEX as a file that can be imported into a GIS for publication quality map production.

8 Tasks

8.1 Loading a CLIMEX application

Before a CLIMEX application can be run, the **DYMEX Simulator** must be started, the appropriate application (equivalent to a DYMEX Model) selected and a suitable Simulation File loaded.

1. Go to the **Start** menu and find the **DYMEX** program group within the **Programs** group.
2. Choose **Simulator** to start the DYMEX Simulator.
3. Either go to the **File** menu and select **Open**, or if the button bar is displayed, left click on the **Open** button .
4. Select the application to be opened from the list labelled “Models” in the left panel (Fig. 8-1). The **Model Information** panel shows more information about the selected application.
5. If more than one Simulation File is shown in the **Simulation Files** panel at the right, select the required file by clicking on it. If only a single Simulation File is listed for the model, it can also be selected by double-clicking on the model name.

Note: If the required Simulation File has been opened recently, it can be selected from the list at the bottom of the “File” menu. Only one Simulation File (and hence one application) can be loaded into the **Simulator** at any one time.

Fig. 8-1. Open Existing Simulation window, showing selection of the “CL – Station Data” simulation file of the ‘Compare Locations (1 species)’ application.



8.2 Operations on Species Files

All of the CLIMEX Model parameters for a species are stored in a CLIMEX Species Parameter File (with the extension .exp). This file also stores any comments provided for that species. To run the CLIMEX model, a particular species must first be selected. This is done from the Species module display (see below).



The name of the currently selected species is shown in the panel at the right. At the far right is a button, which, when clicked, shows a drop-down menu of available species. New species may be added or unwanted species deleted with the **Add/edit/remove Species Parameter File** menu option. The same menu option can be used to indicate the directory where the Species Parameter Files are stored.

Selecting a Species

To select a species, click on the grey selector button at the far right and click the required species name in the drop-down menu. If no species names are shown, it is possible that CLIMEX is not looking in the right directory. Click on the **Add/edit/remove Species Parameter File** menu option, and then navigate to the correct directory using the directory button at the top left of the **Parameter File Selection** dialog.

To create a new Species Parameter File, click on the **Add/edit/remove Species Parameter File** menu option. This opens the **Parameter File Selection** dialog. Select the template or species that you judge to be the most similar to the new species from those in the list at the left and select it by clicking on its name. Then click on the **Create/Copy** button. In the **Create CLIMEX Species dialog**, fill in the name of the new species in the corresponding edit box. Click the **Copy Comments** checkbox only if comments are to be copied from the source species. Click **OK** and the new species file will be created. To adjust the parameters of the new species, select the species as above and use the usual parameter changing procedure (see Section 8.3).

Deleting a Species

To delete a species file, click on the **Add/edit/remove Species Parameter File** menu option. This opens the **Parameter File Selection** dialog. Select the species to be deleted from those in the list at the left, and select it by clicking on its name and click the **Delete** button.

Creating a new Species

Changing the name of a Species

Sometimes it may be necessary to change the name of a species after it has been created (for example, it may have been misspelled). This is done from the **Model Parameters** window (not the **Parameter Grid** window), which is opened by selecting **Model Parameters** from the **Initialization** menu. Click on the line containing the species name in the tree diagram at the left and then edit the name in the **Description** panel at the top right of the window.

Printing the species list

Clicking on the **Print List** button within the **Parameter File Selection** dialog will print a list of all species and their corresponding Species Parameter File names.

8.3 Operations on Species Parameters

CLIMEX Species Parameter Files are text files with the extension **.csp**, stored in the species directory (generally `..\Model\Climex\Species`). Species parameters for the currently loaded species can be viewed and modified most conveniently from the **Parameter Grid window**. Clicking on the Species module's parameter icon or picking **Show Parameters** from the popup menu that results from by clicking on the module bitmap will lead to the Parameter "Grid" window (

). Note that the Parameter Grid View must be enabled by checking the **“Use Parameter Grid for non-Lifecycle modules”** option in the **Current Model** preferences (this is done with the included Simulation Files).

DISPLAYING OUTPUT

Parameters: Bufo marinus - Cane toad					
<input checked="" type="checkbox"/> Temperature Index					
DV0	DV1	DV2	DV3		
14	27	32	36		
<input checked="" type="checkbox"/> Moisture Index					
SM0	SM1	SM2	SM3		
0.3	0.7	2	4		
<input type="checkbox"/> Light Index					
<input type="checkbox"/> Diapause Index					
<input checked="" type="checkbox"/> Cold Stress					
TTCS	THCS	DTCS	DHCS	TTCSA	THCSA
0	-0.01	20	-0.0004	0	0
<input checked="" type="checkbox"/> Heat Stress					
TTHS	THHS	DTHS	DHHS		
37	0.0015	0	0		
<input checked="" type="checkbox"/> Dry Stress					
SMDS	HDS				
0.25	-0.002				
<input type="checkbox"/> Wet Stress					
<input type="checkbox"/> Cold-Dry Stress					
<input type="checkbox"/> Cold-Wet Stress					
<input type="checkbox"/> Hot-Dry Stress					
<input type="checkbox"/> Hot-Wet Stress					
Day-degree accumulation above DV0					
DV0	DV3				
14	36				
Day-degree accumulation above DV3					
DV3	*DV4				
36	100				
Day-degree accumulation above DVCS					
DVCS	*DV4				
14	100				
Degree-days per Generation					
PDD					
0					

Fig. 8-2. Parameter Grid Window showing *Bufo marinus* parameters.

The window shows all parameters, grouped by component indices (the blue cells, for example “Moisture Index”). The parameter names are listed in their abbreviated form. The full name of any parameter can be obtained by moving the mouse cursor over the name field and leaving it there for about a second. This will cause a small window with the full name to appear next to the cursor.

Parameter values are modified by double-clicking on the relevant parameter value field and typing the new value. Typing within a field should be terminated with either the “Enter” or “Tab” key (or clicking on another grid cell). The window is closed using the “Close” button at the top right.

To avoid clutter by “unused” parameters, groups of parameters can be specified as “unused” by un-checking the small box in front of the component index name. This will close then hide the associated parameters, and set the value of that index to its “unused” value (which will be 1 for the Growth Index components and 0 for the Stress Index components). For example, if the Wet Stress box is un-checked, the Wet Stress will always be 0, no matter what the current setting of its parameters.

If a “protected” parameter File is selected, its protected status is indicated by the word “[Protected]” after the species name in the window title, and by the fact that

the component index fields are coloured grey rather than blue. These files cannot be modified using the Parameter Grid Window.

When the Parameter Grid Window is active, several File menu options become available to perform functions on the Species Parameter File. These are:

Save	Saves the modified species parameters
Revert to Saved	Retrieves the most recently saved values from the file
Print Setup	Allows changes to print settings to be made
Print	Prints the species parameters



Comments can be added to the Parameter File by clicking on the **Edit Comments** button in the upper right of the Parameter Grid Window. Such comments should always be added when a species file is created, and any time modifications are made.

CLIMEX provides an alternative way of accessing and changing the species parameter values, using the **Initialization | Model Parameters** menu item from the main window. This opens the **Model Parameters** window. The Model Parameters window provides more information about each parameter than the Parameter Grid Window, while also giving more information about the context in which the parameter is used. If you made changes to parameter values using the Parameter Grid Window can save them by using the **File | Save** menu item or by saving them when closing CLIMEX. Note that saving the Simulation File does not automatically save the changed parameter settings.

The large panel at the top left shows the modules that have parameters. In the “Compare Locations” model, it will show three modules, (1) The Climate Change Scenario (labelled with the selected scenario), (2) the Soil Moisture model and (3) the Species (labelled with its name). Each module’s parameters are arranged within a “tree” of module components. To modify (say) the “Dry Stress Threshold” parameter, click on the small rectangle (#) in front of the species name to show all of the component indices of the module. Then do the same with the rectangle in front of the “Dry Stress” component. This will show the component function (labelled with its driving variable, which is the Soil Moisture in the case of Dry Stress). Repeat with the function branch, and the two Dry Stress parameters will now be visible (Fig. 8-3). Now click on the Dry Stress Threshold node. The panel below the selection window will now show the value of the parameter, where the value can be changed either by using the slider bar or by typing a new value into the edit box. Note that the parameters description and its “context” (i.e., the function within which it is located) are shown at the right of the dialog.

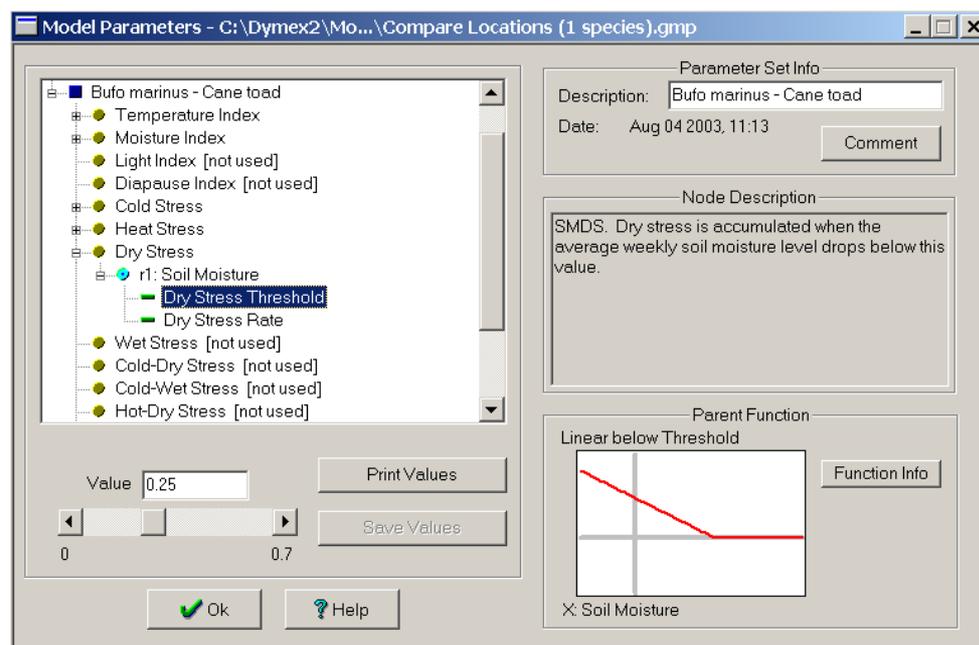


Fig. 8-3 Parameter Window showing *Bufo marinus* Dry Stress Threshold.

The **Print Values** button will print the parameters, while the **Save Values** buttons will become available for saving the file if either a parameter or parameter file comment is changed (either from this window or from the Parameter Grid Window). Component indices can be inactivated (i.e., set to “unused”) or activated by clicking on the component in the selection panel and either checking or un-checking the box marked “**Do not use this Process**” that will appear at the lower left.

Changing the species name

Note that the name of the species can only be adjusted from within this window. Select the Species module (for example, “*Bufo marinus* – Cane toad” in the example in Fig. 8-3) and then type the adjusted name into the **Description** box at the top right.

8.4 Using Climate Change Scenarios

The CLIMEX software enables users to consider the potential implications of either climate change or irrigation practices on both the similarity of meteorological data of different places (Match Climates) or the potential geographical distribution and seasonal phenology and survival of species (CLIMEX model). These changes may have implications in a number of areas. As species’ distributions change, some areas currently unsuitable for the species may become suitable. Alternatively, a change in the climate may force a species’ range to contract or move – this has implications



not only for pests but also rare and threatened species of conservation value. Changes in the climate may affect important agricultural industries if areas become unsuitable for the crops currently being grown there. By allowing the user to assess likely situations in the future, this option provides the rationale to initiate management or adaptation strategies to cope with the projected changes.

The **Climate Change** scenario module enables users to examine the impact of different temperature and rainfall conditions. The module display (shown above) shows the name of the currently selected scenario in the right-most panel. At the

far right is a button, which, when clicked, shows a drop-down menu of available scenarios. New scenarios may be added or unwanted scenarios deleted with the **Add/edit/remove** menu option. The same menu option can be used to indicate the directory where the Scenario Files (.ccs) are stored. Scenarios can apply changes to summer and winter months separately in temperate zones and a different change can be made to a user-defined equatorial zone.

Once a scenario has been selected, its parameters can be changed using the usual procedure (see Section 8.3). These parameters are listed in Table 8-1.

Table 8-1 Climate Change Scenario parameters.

Min. Temperature Change (Winter)	W_{tmin}	The change in the minimum temperature that occurs during the <u>'Winter'</u> months outside the Equatorial zone.
Min. Temperature Latitudinal Change (Winter)	WL_{tmin}	The change in the minimum temperature for each degree of latitude away from the equator that occurs during the <u>'Winter'</u> months outside the Equatorial zone.
Min. Temperature Change (Summer)	S_{tmin}	The change in the minimum temperature that occurs during the <u>'Summer'</u> months outside the Equatorial zone.
Min. Temperature Latitudinal Change (Summer)	SL_{tmin}	The change in the minimum temperature for each degree of latitude away from the equator that occurs during the <u>'Summer'</u> months outside the Equatorial zone.
Min. Temperature Change (Equatorial Zone)	E_{tmin}	The change in the minimum temperature that occurs inside the Equatorial zone.
Max. Temperature Change (Winter)	W_{tmax}	The change in the maximum temperature that occurs during the <u>'Winter'</u> months outside the Equatorial zone.
Max. Temperature Latitudinal Change (Winter)	WL_{tmax}	The change in the maximum temperature for each degree of latitude away from the equator that occurs during the <u>'Winter'</u> months outside the Equatorial zone.
Max. Temperature Change (Summer)	S_{tmax}	The change in the maximum temperature that occurs during the <u>'Summer'</u> months outside the Equatorial zone.
Max. Temperature Latitudinal Change (Summer)	SL_{tmax}	The change in the maximum temperature for each degree of latitude away from the equator that occurs during the <u>'Summer'</u> months outside the Equatorial zone.
Max. Temperature Change (Equatorial)	E_{tmax}	The change in the maximum temperature that occurs inside the Equatorial zone.

DISPLAYING OUTPUT

Zone)		
Rainfall Change (Winter)	W_{rain}	The change in the rainfall (in %) that occurs during the 'Winter' months outside the Equatorial zone.
Rainfall Latitudinal Change (Winter)	WL_{rain}	The change in the rainfall (in % per degree latitude) for each degree of latitude away from the equator that occurs during the 'Winter' months outside the Equatorial zone.
Rainfall Change (Summer)	S_{rain}	The change in the rainfall (in %) that occurs during the 'Summer' months outside the Equatorial zone.
Rainfall Latitudinal Change (Summer)	SL_{rain}	The change in the rainfall (in % per degree latitude) for each degree of latitude away from the equator that occurs during the 'Summer' months outside the Equatorial zone.
Rainfall Change (Equatorial Zone)	E_{rain}	The change in the rainfall (in %) that occurs inside the Equatorial zone.

The **Equatorial Zone** (z) is a band of latitude centred on the equator and its exact extent is defined in the **Model Preferences** dialog. "Winter" and "Summer" seasons are defined with reference to the Equatorial Zone and the day of year. North of the Equatorial Zone, "Summer" is defined as days 92 to 274 (i.e., March 2 – September 30, inclusive), with "Winter" being the rest of the year. South of the Equatorial Zone, these are reversed.

In the following descriptions, I represents the current climate variable and O the variables after the climate change scenario is applied.

The temperature output variables are derived from the input variables and factors as follows (the "t" subscript can be either "tmin" or "tmax", while l is the latitude, in degrees):

in the **Equatorial Zone**, $O_t = I_t + E_t$

elsewhere, in "Winter", $O_t = I_t + W_t + WL_t \times (abs(l) - z/2)$

elsewhere, in "Summer", $O_t = I_t + S_t + SL_t \times (abs(l) - z/2)$

The rainfall output variables are derived from the input variables as follows:

in the **Equatorial Zone**, $O_{rain} = I_{rain} \times \left(\frac{E_{rain}}{100} \right)$

elsewhere, in "Winter",

$$O_{rain} = I_{rain} \times \left(1 + \frac{W_{rain} + WL_{rain} \times (abs(l) - z/2)}{100} \right)$$

elsewhere, in "Summer",

$$O_{rain} = I_{rain} \times \left(1 + \frac{S_{rain} + SL_{rain} \times (abs(l) - z / 2)}{100} \right)$$

8.5 Using Irrigation Scenarios

The **Irrigation** scenario builder acts in a similar way to the **Climate Change** scenario builder with two exceptions. Firstly, it treats all regions equally and it only affects the amount of additional, simulated rainfall. There are options to apply irrigation on a monthly rather than seasonal basis and to apply ‘top-up’ amounts of irrigation to increase the effective rainfall to a set amount, such as 3 mm per day of (rainfall + irrigation), for example. It is particularly useful when considering the suitability of climates for a species that may only be present as a result of irrigation in agricultural or urban environments. Irrigation can affect both the growth and survival of species by enabling populations to persist and grow in what would otherwise be an unfavourable dry season.

8.6 Specifying Location Selections

In normal operation, the Compare Locations and Match Climates functions are run on a set of locations (often referred to as a “Location Sequence”). For example, in the Compare Locations illustration below, a Location Sequence with the name of **Australia** is being used [The equivalent module in Match Climates is named **Locations (Away)**].



The module display (shown above) shows the name of the currently chosen Location Sequence in the right-most panel. At the far right is a button, which, when clicked, shows a drop-down menu of available sequences. New sequences can be added or unwanted ones deleted with the **Add/edit/remove** menu option.

Creating a Location Sequence

A **Location Sequence** in CLIMEX provides a link to a MetManager **Location Selection** (see Section 28 of the DYMEX Simulator User’s Guide) that has been prepared for use by DYMEX. Hence, to create a new Location Sequence, the corresponding Location Selection must first be created in the MetManager. Section 10.2 of this Guide gives a detailed description of how to do the latter. Then, to create the new **Location Sequence**, select the **Add/edit/remove** menu option from the drop-down menu obtained by clicking on the button at the far right of the Locations module display. This opens the **Sequences (Locations)** dialog.

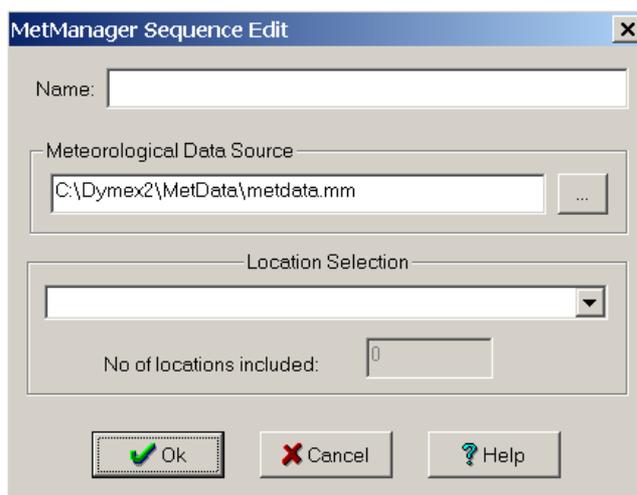


Fig. 8-4 The MetManager Sequence Edit dialog.

The Sequences (Locations) dialog lists the existing Sequences. To add a new one, click on the New button, which opens the MetManager Sequence Edit dialog

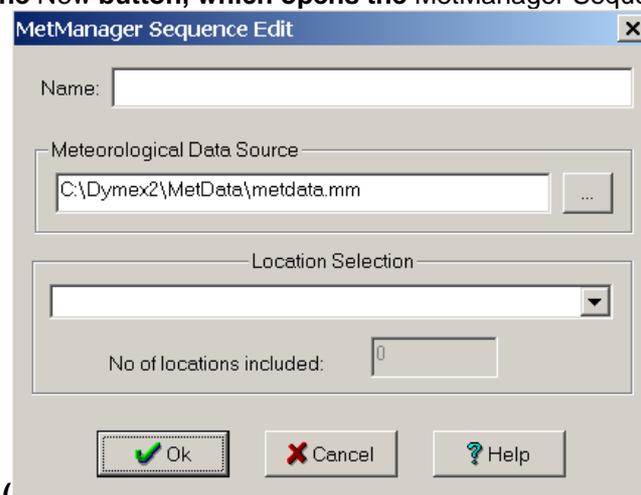


Fig. 8-4). Here, the MetManager database (.mm file) containing the data that the Sequence will use must be specified. Click on the Browse button () in the **Meteorological Data Source** panel and navigate to the required file. Then click on the button () in the Location **Selection** panel. A list of Location Selections defined for the chosen database will appear, from which the required one can be selected. The Name field at the top of the dialog is automatically set to the name of the Location Selection – however, it can be changed to any other name if desired. Then click **OK** to return to the Model Window, and the new Location Sequence is ready for use.

8.7 Adding new Locations to the Database

Climate data can be imported into the MetManager for use by CLIMEX from several sources.

1. Text files in .loc/.met format (the same as used by previous versions of CLIMEX), whose format is detailed in Appendix 3 of the DYMEX Simulator User's Guide.
2. CLIMEX Version 1 MetManager database (.bmt) files
3. CLIMEX Version 2 MetManager database (.mm) files

The procedure for importing from any of these sources is described in Section 28.4 of the DYMEX Simulator User's Guide.

8.8 Using the grid-based Climate data

A set of climate data derived from a global splined grid of data with a resolution of 0.5° from the Climate Research Unit (CRU) at Norwich UK (<http://ipcc-ddc.cru.uea.ac.uk>, New et al. 1999) has been included with the CLIMEX program. This data set can be conveniently used with the Compare Locations function by loading the model with the "CL – Grid Data" Simulation File. The "CL – Grid Data" Simulation contains Location Sequences on the grid data for every continent as well as the whole world. In addition, it defines map formats (using colour gradations to indicate index values) that are most suitable for mapping the grid data.

8.9 Creating and modifying map backgrounds

The DYMEX **Map Manager** is used to create Region Definitions that are then used as the base onto which the data layers of the map are drawn. The Map Manager is accessed from the File menu. Map features specified by the Map Region definition include the boundaries of the region, the included shape files that determine the geographical features visible on the map and the map projection. Before any DYMEX output variable can be mapped, the appropriate Region Definition must have been created. Region Definitions are stored in Region Definition Files (.mdf) in the Map Directory, as specified in the **Operating Preferences**.

To understand how a display map in DYMEX is related to the Region Definition, it is necessary to understand how such a map is structured. A map consists of a number of layers, as follows:

Data Layer

Shape File n

...

Shape File 1

Background

This is illustrated in Fig. 8-5. The final map produced when selecting the "**Results/Map**" menu item after a run of DYMEX is the result of drawing all these

layers (in order, from bottom to top), within the coordinate boundaries specified and with the specified map projection. Every layer except the Data layer is defined in the Region Definition. The Data Layer is added from the specifications provided in the Map Format (set in the **Map Specification** dialogs).

The Background layer is used to provide a background colour for the map. Each Shape File layer is associated with a single file containing spatial data in the ESRI® Shapefile Format (see the ESRI Shapefile Technical Description at www.esri.com/library/whitepapers/pdfs/shapefile.pdf). Another layer, not shown in Fig. 8-5, contains the positions of the title, legend and scale marker. This information is also stored in the Region Definition, though the actual contents of the title and legend are added when the map is actually drawn, again from information provided in the Map Format.

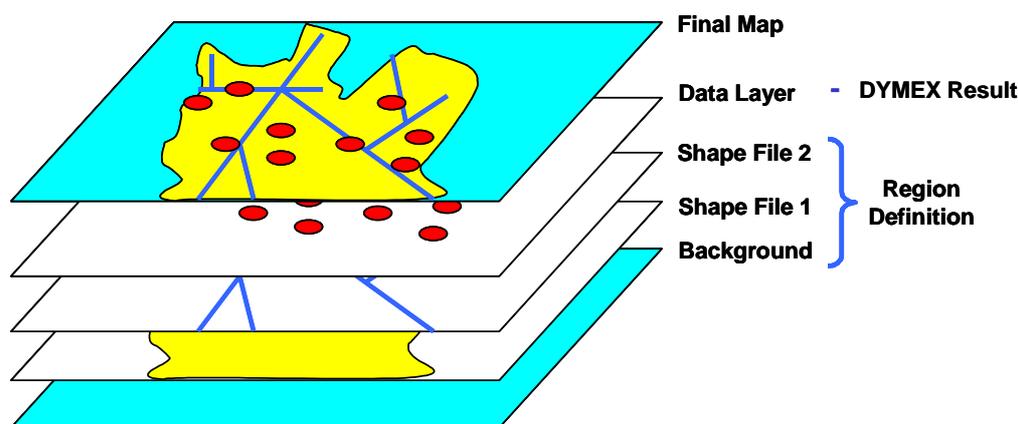


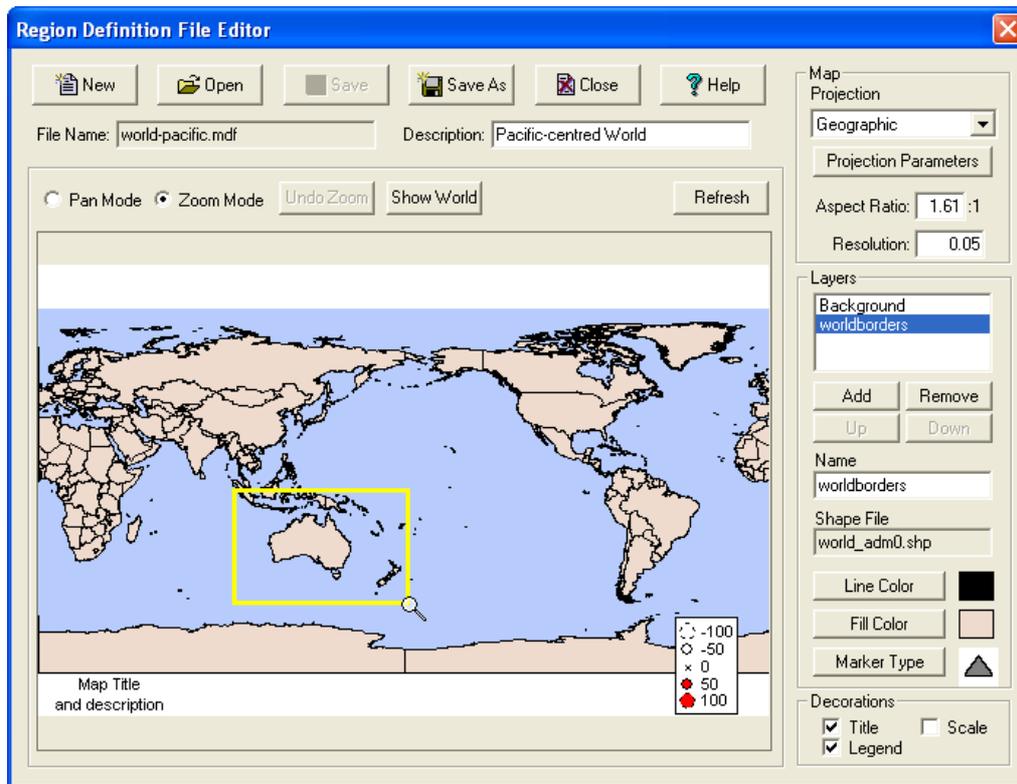
Fig. 8-5 The layers that constitute a map displaying a variable in DYMEC, and the map that is produced from these layers.

When the MapManager menu item is selected from the File menu, the **Region Definition File Editor** dialog is opened (Fig. 8-6). This is a rather complex dialog, which allows the user to create a Region Definition File by adding the required shapefiles layers to a background, specifying colours for each layer, and selecting a projection for display. The large display panel at the lower left shows the map as it will be shown (but without the data layer). The row of controls above this panel is used to pan and/or zoom the map until it is cropped to the required coordinates and to select label, title and scale marker decorations.

➤ To create a new Map Region File

1. Select the **MapManager** item from the **File** menu.
2. This will open the **Map Definition File Editor** dialog (Fig. 8-6), with only a background layer present in the **Layers** list at the right. Change the background colours, if required, by clicking on the **Line Colour** or **Fill Colour** buttons, and then choosing the desired colours.

Fig. 8-6 The Region Definition File Editor dialog, showing a map of the world, using the Geographic projection with Central Meridian set to 180°.



3. Add a Shapefile layer by clicking on the **Add** button, and navigating to the required shape file. Note that shape files generally have the extension **.shp**.
- ! Note that shapefiles should be stored in the **Map** directory, which is specified in the **Operating Preferences** dialog.
4. To see the effect that the addition of the layer has on the appearance of the map, click on the **Refresh** button, which redraws the map with current settings applied. You can change the colours of the currently selected layer using the **Line Colour** and **Fill Colour** buttons.
 5. Add more layers as required by repeating steps 3 and 4. If a layer has been added by mistake, it can be removed by selecting it and clicking on the **Remove** button. Similarly, the ordering of layers can be changed with the **Up** or **Down** buttons (however, the background layer will always be the first).
- ! A wide variety of spatial data resources are available, including many that are free. Administrative region data can be found here (<http://www.gadm.org/>).
6. If a projection other than the default **Geographic** projection is required, click on the small arrow button to the right of the projection name and select the required projection from the drop-down list. The set the projection parameters as required using the **Set Parameters** button.
- ! The choice of a projection and different projection parameters can make a big difference to the appearance of a map. If you are unfamiliar with projections, you may wish to consult a reference such as Maling, D.H., *Coordinate Systems and Map Projections*, 2nd ed, rev., Woburn, MA: Butterworth-Heinemann, 1993 or Mulcahy, K., "The Map Projection Home Page." at <http://everest.hunter.cuny.edu/mp/>.

7. Click on the Refresh button to update the map display, then use the **Pan** and **Zoom** options to size and position the map to your exact requirements.
8. By default, the map displays a label, legend and scale marker. If any of these are not required, uncheck the corresponding box in the **Decorations** panel at the lower right of the dialog. These ‘decorations’ can be moved to any desired spot on the map by clicking on them and ‘dragging’ with the mouse button down.
9. When the map region display shows all features exactly as they are required in the final map, the map region can be given a description by typing the appropriate text into the **Description** box.
10. Click on Save to save the Region Definition as a file with the **mdf** extension.
11. Close the dialog using the **Close** button.

The map produced by the Region Definition is shown in the large panel that takes up most of the dialog. Most of the controls in the dialog are used to alter the Region Definition. To avoid lengthy delays after settings are changed on slower computers, the display is not updated every time a change is made to the Region Definition. The **Refresh** button is provided for this purpose.

The shape of the cursor indicates the currently set operation, either “pan” () or “zoom” (). When in pan mode, clicking on the map and moving the mouse with button down will move the map into the same direction as the mouse. In zoom mode, two methods can be used to zoom in onto a point or region. Left-clicking on a map point causes the map to be redrawn at an about 50% larger scale, thus showing a smaller area, positioned so that the point that was clicked stays in the same absolute position on the screen. Similarly, right-clicking on a point draws a 50% smaller version of the map, again positioned so that the clicked point does not move. Note that there are limits to the enlargement and reduction allowed for a map. An alternative way to zoom into an area is to left-click on the point that is required as the top-left corner of the zoomed area and, leaving the mouse button down, move the mouse to the desired lower-right corner of the area. A yellow rectangle outlining the zoomed area will appear as the mouse is moved on the screen (Fig. 8-6). When the mouse button is released, the area outlined by the zoom-rectangle will be drawn to occupy the full size of the map display area. Zoom operations can be reversed by clicking the **Undo Zoom** button, which takes the map back to the state it was in before the previous zoom operation. **Undo Zoom** can be invoked as often as required to reverse multiple zoom operations.

The **Show World** button changes the boundaries of the region to show the whole world (i.e., -90° to $+90^{\circ}$ latitude, -180° to $+180^{\circ}$ longitude).

Map Region Definitions can contain up to three extra features (decorations). A legend, a title and a scale marker can be specified by checking the appropriate box in the **Decorations** panel. The actual contents of both the legend and title are determined by the data plotted on the map (a “dummy” content is shown in the map panel). Each of these decorations can be “dragged” to the required position using the mouse.

The buttons along the top of the dialog allow a new Map Region File to be created, an existing one to be opened, or the one currently being worked on to be saved (**Save**) or saved with a different name (**Save As**). The Close button closes the Region Definition File Editor, prompting the user to save the file if it has been altered.

8.10 Extracting run information for use in a spreadsheet or GIS

The mapping facilities within CLIMEX are intended for convenient and rapid presentation of results and may not be suitable for developing publication quality maps. In addition, the user may wish to present CLIMEX results in the same map as say a species distribution and other information that provides spatial context (e.g., roads, rivers and place names).

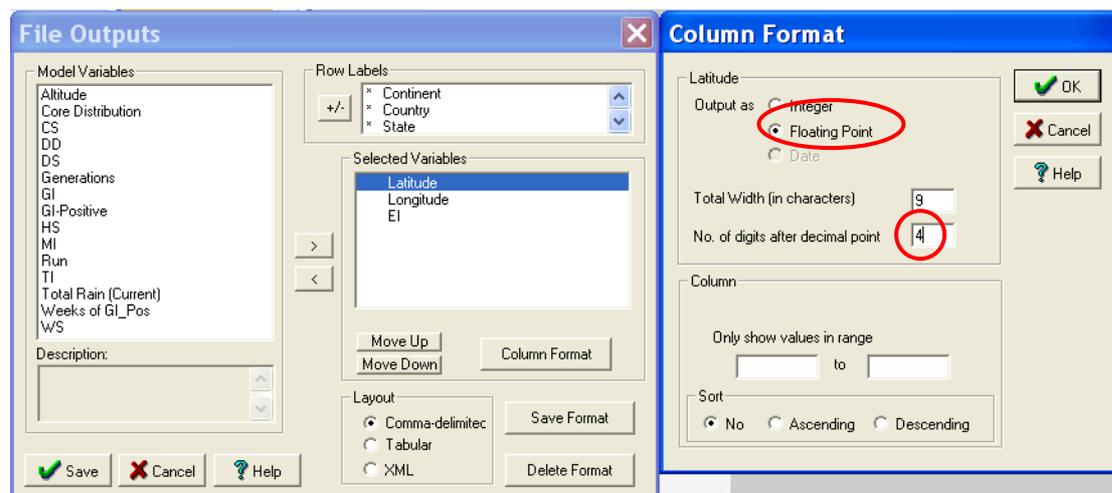
To extract information from a CLIMEX run, you first need to select or define a file format. With the run results window active, select **Results|File|New** to open the File Outputs dialog (Fig. 8-7). Define a new file format by selecting variables in the **Model Variables** box and clicking on the  button to transfer them into the **Selected Variables** box. Include both the Latitude and Longitude and any variables that you wish to map (e.g., EI, GI and CS). Select the Latitude variable and then click on the column format button. Set the output as floating, and the number of digits after decimal point between 2 and 4, depending upon the precision of your Meteorological locations (4 decimal points should be adequate for most applications, giving a precision of about 10 m). Click **OK**, select the Longitude variable and repeat setting the column format.

For most Geographical Information System (GIS) packages, the best layout type will be comma-delimited. When you are satisfied with the format settings, save the format using the "**Save Format**" button, navigating to a location where you will do your analysis or mapping, and give the file a new name.

The first step in using the file in a mapping situation will normally involve opening your file within a spreadsheet or text editor, and deleting the first seven lines of the file header, leaving only the column label line as the file header. After saving your file, remember to close it, otherwise some GIS packages will fail to import it. Open your GIS package and import your file as a set of points. To import the file into ArcMap 9.x, open the ArcMap application with a new or existing map, and then select **Tools|Add XY Data**. Click on the  button to navigate to the .csv file that contains your CLIMEX point location information. ArcMAP will automatically recognise the Longitude and Latitude column labels as the X and Y coordinate columns.

DISPLAYING OUTPUT

Fig. 8-7 The File Outputs dialog, showing latitude, longitude and EI selected for writing to an output file.



Once imported into the GIS, there are multiple approaches that can be used to store and represent the data. As with CLIMEX, point locations can be thematically mapped using a graduated symbol size, with increasing dot radius representing increasing values of the mapped variable. Alternatively, gridded variables can be thematically mapped using either a raster or a polygon fishnet. If you are only interested in mapping one or a few variables, and you are unlikely to need to use data from the same meteorological database again, then it may be simpler to convert the points into a raster for the variable(s) of interest. If, on the other hand you want to map several variables or you could use the same meteorological database again, then it may be worth converting the point data to a raster using the location name as the raster variable. The location field can then be used to do a database join with any other files extracted from CLIMEX for the same selection. Following the database join, any of the CLIMEX variables can be mapped by selecting it as the variable of interest, classifying it and defining its symbology. Using this approach, you need only import the file once, saving considerable time in preparing maps for the second and subsequent variables. By representing the location data as polygons, the polygon shapes can be smoothly projected or clipped to geographic boundaries such as country outlines, providing an aesthetically pleasing product.

8.11 Converting species files from Version 1 to Version 2

In Version 1 of CLIMEX, parameter values for all species were stored in a single file (species.par) while species comments were stored in separate files (one per species) with the extension '.cmt'.

In Version 1 of CLIMEX, parameter values for all species were stored in a single file (species.par) while species comments were stored in separate files (one per species) with the extension '.cmt'.

Version 2 uses a different scheme, with each species having its own parameter file (with the extension .exp) that stores both the parameter values and associated comments. These files are located in the 'species directory', generally '\Dymex2\Models\Climex\Species'.

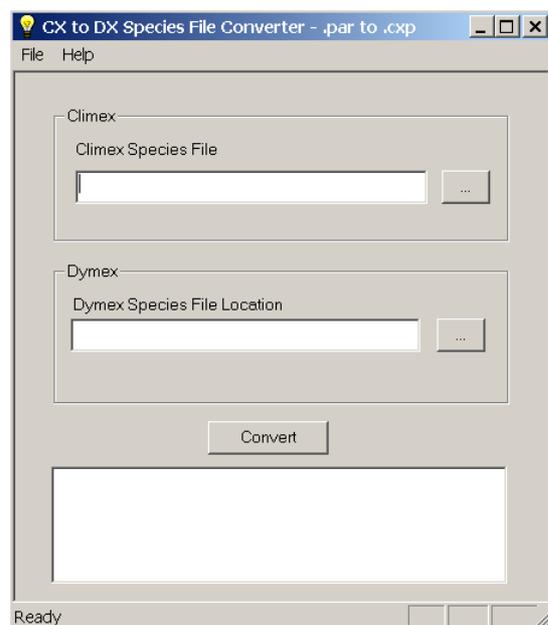
To facilitate the conversion from Version 1 to Version 2, the CLIMEX release includes a program that converts the species files from Version 1 format to Version 2. The program (**cxdxconvert.exe**) is situated in the '\Program Files\Dymex2' folder, and can be started by double-clicking on it in Windows Explorer (Fig. 8-8).

In Version 1 of CLIMEX, parameter values for all species were stored in a single file (species.par) while species comments were stored in separate files (one per species) with the extension '.cmt'.

Version 2 uses a different scheme, with each species having its own parameter file (with the extension .exp) that stores both the parameter values and associated comments. These files are located in the 'species directory', generally '\Dymex2\Models\Climex\Species'.

To facilitate the conversion from Version 1 to Version 2, the CLIMEX release includes a program that converts the species files from Version 1 format to Version 2. The program (**cxdxconvert.exe**) is situated in the '\Program Files\Dymex2' folder, and can be started by double-clicking on it in Windows Explorer (Fig. 8-8).

Fig. 8-8 The program that is used to convert CLIMEX species files from Version 1 to Version 2 format.



Use the button at the right in the **Climex** panel to browse to the location of the Version 1 species file. Then use the equivalent button in the **Dymex** panel to locate the directory into which the converted files are to be stored. This should generally be a directory named 'Species' within the CLIMEX model directory. Then click the **Convert** button to start the conversion process.

Parameter values are kept exactly as they are in the original species file. Since Version 2 of CLIMEX is stricter about the relative values of parameters, this may cause some warnings or errors when converted files are opened in CLIMEX. For example, Version 1 of CLIMEX allowed a value of SMDS to be greater than SMO, which is not permitted in Version 2. Any such files will generate an error and will not be accepted by Version 2 of CLIMEX. In such cases, the species file will need to be opened in an editor such as Notepad and the invalid parameter changed manually (Once changed to comply, they will need to be checked to re-fit the geographical distribution). See Appendix 3 for a description of the species file format.

9 Advanced Tasks

9.1 Saving out detailed results (Output Streaming)

In Version 4 of CLIMEX it is possible to save out detailed data into text files for subsequent analyses. The process is called *Output Streaming*.

The output stream properties dialog allows setting of the variables to be saved and the format for output files (Fig 9-1).

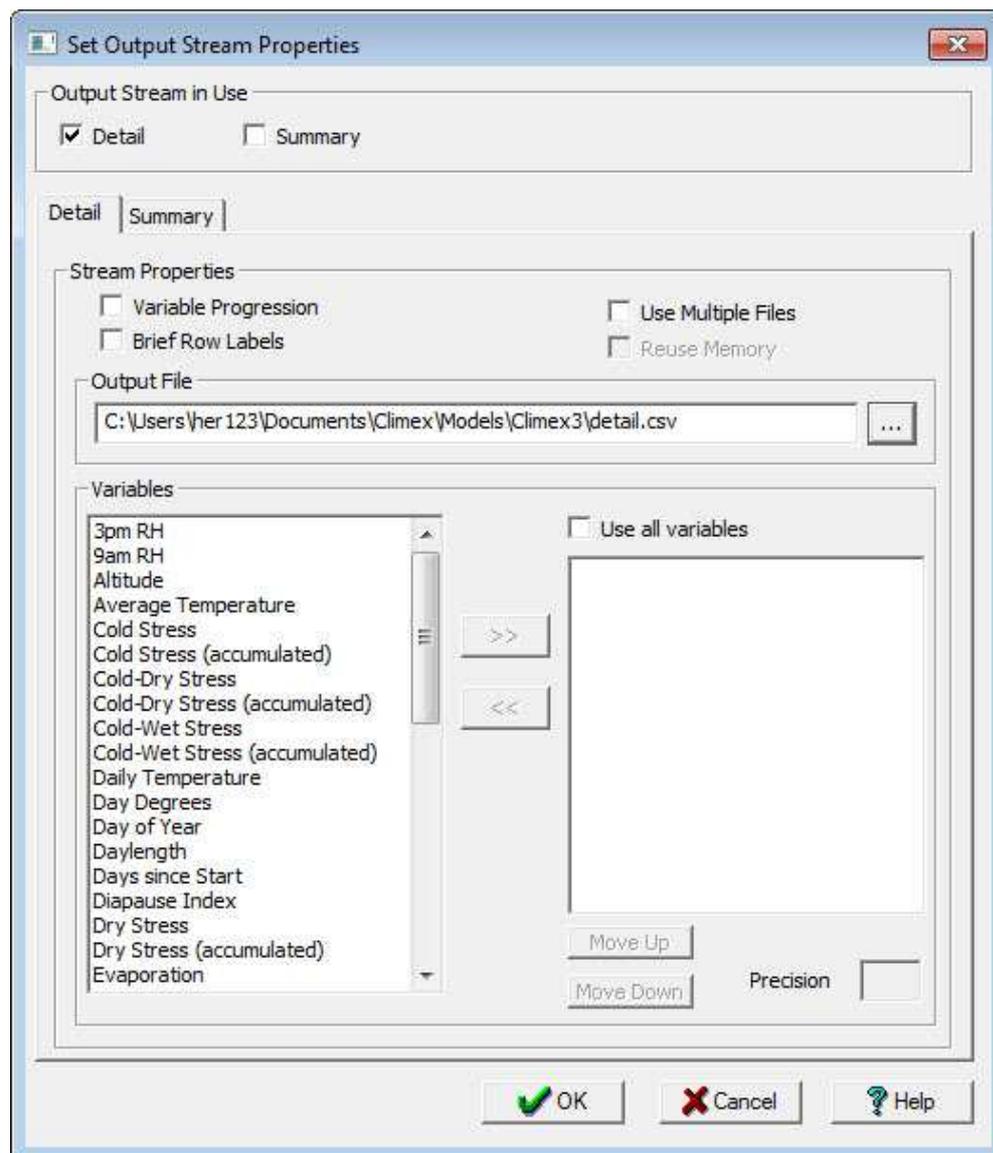


Fig 9-1 Output Stream Properties Dialog

9.1.1 Output Stream in Use

Output can be **Detail** (including variables for the model timestep) or **Summary** (annual variables). You can elect to save out *none*, *Detail*, *Summary* or both

Detail and Summary. N.B. saving out this data can consume very large amounts of computer disc space and slows the model run significantly. You should take the time to understand the implications of each of these choices before implementing them.

9.1.2 Detail Stream Properties

Selecting **Variable Progression** will partition (group) the output by the output variable. If you specify the use of multiple files, each file will contain results for a single variable. If you don't specify multiple files, Variable Progression will partition the variable outputs into the sequence specified by the order of the variables listed in the right-hand column.

Selecting **Brief Row Labels** will strip out the Continent, Country and State information in the output file, leaving just the Location Name and the selected output data. For gridded datasets, the Continent, Country and State information is extraneous. The location number is a unique code for each location in a met data file.

Selecting **Use Multiple Files** will separate out the results for each variable into a separate file.

<i>Variable Progression</i>	<i>Use Multiple Files</i>	<i>Description of output</i>
Yes	Yes	One file per variable, one run per line, one data element per column. In CLIMEX, 52 weekly values in columns
Yes	No	One file per variable, one run per line, one data element per column. In CLIMEX, 52 weekly values in columns
No	Yes	One file per location, one variable per column. This can generate a very large number of files, which can be quite unwieldy to manage.
No	No	One file, one variable per column, blocked by run/location. That is, following the header with the variable names, the first location is described, followed by the values for the variable for each timestep (in the case of CLIMEX 52 weekly values). Even moderately large files can generate extremely long csv files.

9.1.3 Output File

The location and name for the output file. Presently, this is limited to comma-separated variable (CSV) files.

 DYMEX will try to overwrite any existing CSV output stream files specified in the Output File box. If there are any that are open in another application, the write will fail (silently).

Variables

This contains the list of state variables that are maintained by the program. If a desired variable is not listed you may need to generate it. Be aware that the desired variable may be maintained only as a summary variable (e.g., the Ecoclimatic Index). Variables can be selected individually by clicking on the relevant variable in the left hand column, and clicking the  button. Variables can be deselected by clicking on the variable to be removed from the right-hand column and clicking the  button.

Selecting **Use all variables** will transfer all of the variables into the selected set on the right-hand side of the window.

The **Move Up** and **Move Down** buttons adjust the order in which the selected variables is output in relation to the other variables.

Precision

The number of decimal digits required in the formatted datum. For example, a precision of 2 would result in the value 12.456 being output as 12.46. A Precision of 0 produces integer output.

9.1.4 Summary Stream Properties

In the Summary Output Stream there is the option to *Reuse Memory*. This has the effect of discarding results for all but the most recent run. This may be useful for extremely large runs.

Use Multiple Files

When applicable, if this option is selected, CLIMEX will use multiple files to partition the data. For example, if the model is running a compound sequence of Species by Location and Use Multiple Files is selected, CLIMEX will generate one file per species. This is probably a useful outcome, allowing you to build maps for each species.

The current state of the Output Streams settings are presented at the bottom of the Model Components window (Fig 9-2).

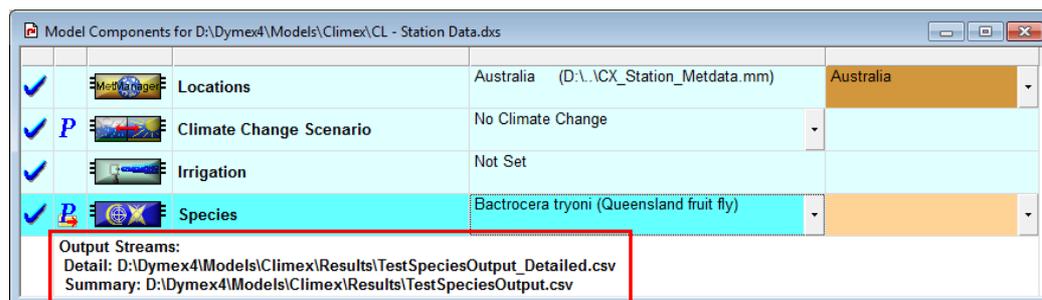


Fig 9-2 Model components window highlighting the selected summary stream properties

9.2 Creating automated run sequences

In Version 4 of CLIMEX we introduce a powerful modification to the underlying DYMEX simulation engine. This new feature allows you to create complex sets of model simulations that can be specified to be run in an automated sequence, removing the need for manual intervention. These sequences can include meteorological data selections, species, sets of parameters, or combinations of these variables.

There are four types of sequences: *Locations*, *Parameters*, *Species*, and *Compound*. The *locations* sequence is the most popular and is used by default in every Compare Locations model run. A *parameter* sequence can be used to explore the effects of adjusting one or more parameters using either a sensitivity analysis method based on a Latin hypercube, or an exhaustive exploration of the defined parameter space. These adjustments are in relation to the selected species. A *species* parameter sequence can be defined so that multiple species models can be run in sequence. Finally, a *compound* sequence can be defined by combining two or more of the other sequence types in a nested fashion.

Running batches of models relies on specifying an Output Stream so that the results can be saved. This is specified in a comma separated variable file. To specify the output stream file, select Output Streams... from the Execution menu. The Output Stream Properties dialogue can then be used to specify the data to be captured in the output stream files.

-  Output streams can capture very large amounts of data, and writing the information to disc can slow down model processing runs considerably. Not only could the runs take a long time to run, but it is possible that computer disk storage could be filled. Select a location for the output file(s) where there is a large capacity to expand disc storage, and consider whether you really need to capture the data. The output can be either detailed (weekly), summary (annual) data or both. Depending on this selection, the variables available for capture change. The Detail generates far more data, and can be useful for exploring seasonal patterns, either within a single year, or within several years.

The dialogue allows you to specify using brief row labels, which minimises the capture of extraneous location information. You can also specify whether you want the data from each weekly or yearly timestep to be saved into an individual file (Use Multiple Files option checked), or for all data to be saved into a single (potentially very large) file.

The Detail output stream option provides you with the option of specifying *Variable Progression*. This has the effect of partitioning output by the output variable. If you specify the use of multiple files, each file will contain results for a single variable. If you don't specify multiple files, Variable Progression will block the variable outputs.

In the Summary Output Stream there is the option to *Reuse Memory*. This has the effect of discarding results for all but the most recent run. The Output File specifies the name and location of the output file.

! If the output file name and location is specified in the species parameter sequence file, it overrides whatever is specified in the user interface.

9.2.1 Parameter sequence

A new parameter sequence can be defined by selecting **Define Run Sequences...** from the **Execution** menu, clicking the **New** button, and then selecting **Parameter** from the **Sequence Types** window (Fig. 9-3).

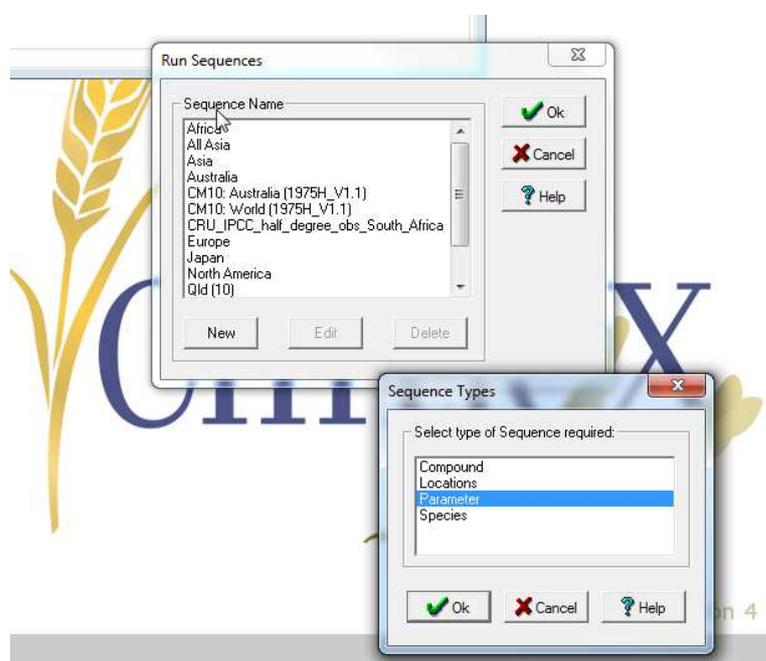


Fig. 9-3 Selecting sequence type to create

Use the Parameter Sequence Edit window to specify the parameters to be edited and the manner in which they are to be adjusted between runs (Fig. 9-4). The type of sequence can have an important influence on the number of runs to be executed. Exhaustive is a full-factorial combination of parameters and the number of steps specified for each parameter. The Sensitivity Analysis method on the other hand results in the sum of the steps specified for all of the parameters. The **Update characteristics for selected parameters** panel allows the modeller to select whether to modify the parameter in direct units or as a percentage change. For some parameters it makes more sense to use one method or the other. For example, soil moisture and rate variables will usually be adjusted proportionally.

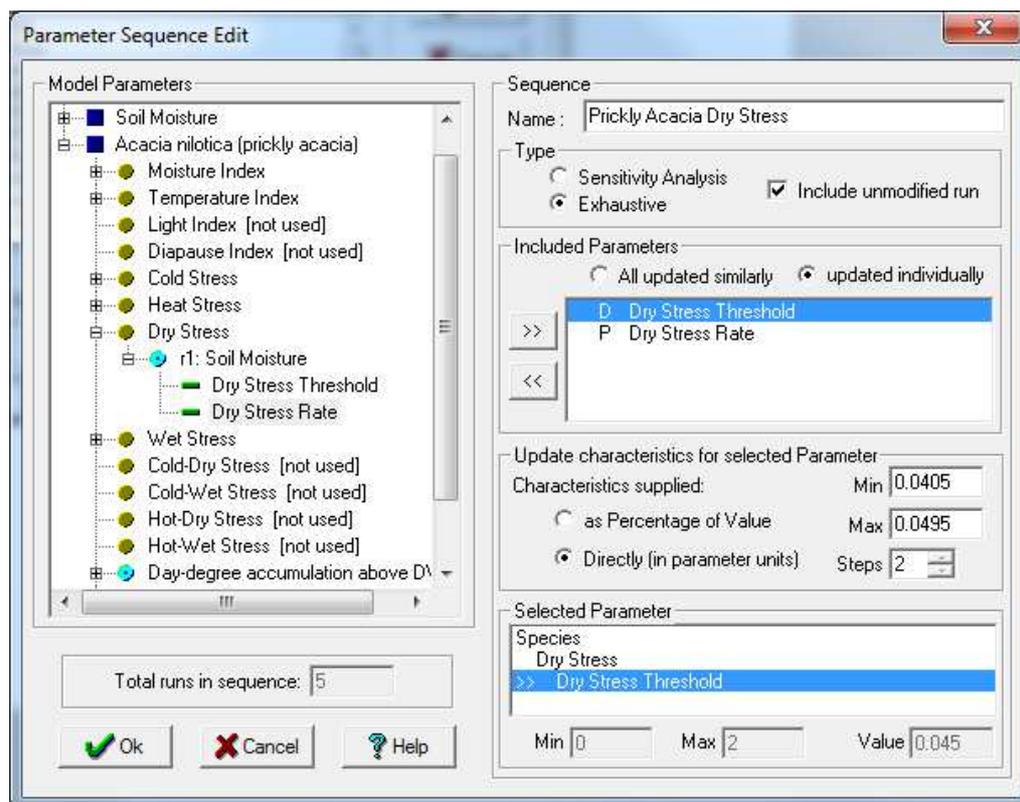


Fig. 9-4 Specifying the parameter sequence

9.2.2 Species sequence

The *Species Sequencer* can run simple (different parameters at a single location) or Compound (Species x Location) sequences. You may want to do this to explore the effects of a range of parameter values for a single species in order to explore the effects of a range of parameters on the fit of a model to explore a hypothesis about a range-limiting factor. This could be very useful for deciding between different types of cold stress mechanisms, or exploring parameter options for interaction stresses, which can be quite difficult to fit. You might also want to run a large list of species automatically across a climate dataset. You may want to do this to understand the risks to a certain jurisdiction from a set of species, or to understand the effects of climate change.

Invalid parameter combinations can be optionally skipped over, with a log file reporting all the runs and whether they worked or were skipped (with the reason for the failure). An example species parameter file is provided to test the system (Fig. 9-5). The fifth entry for species “EE” should generate an error.

Output	"Summary"	"Temperature Index"	"Moisture Index"	"Wet Stress"
Ident	"Limiting low temperature"	"Limiting high temperature"	"Limiting low moisture"	""
AA	14	35	0.32	0
BB	14	35	0.32	0
CC	14	35	0.34	0
DD	14	36	0.36	0
EE	14	35	0.32	1

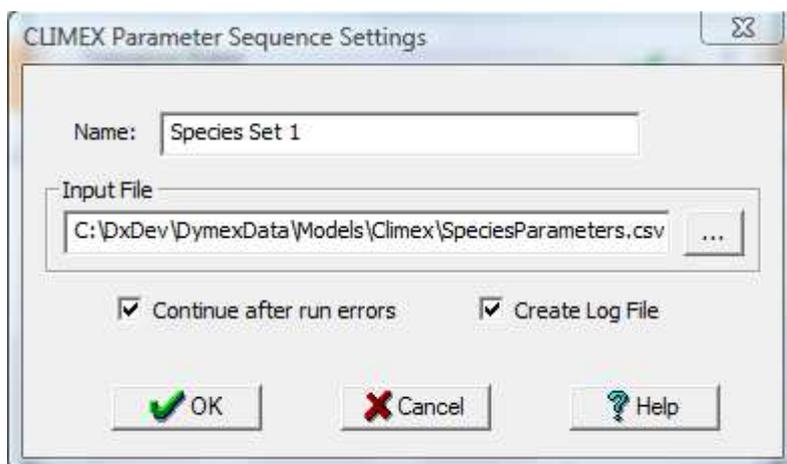
Fig. 9-5 An example species parameter file sequence for testing the system.

Species Sequences are associated with the “Species” module within CLIMEX. They are used to vary the values of one or more parameters from the values of the currently loaded species. A “Sequence Specification File” is used to specify the changed values of parameters and whether particular processes are to be included or excluded.

! The Species sequences overwrite the selected species parameter file. However, it only does so for the parameters that are specified in the Sequence Specification File. If you want to run a sequence of separate species, you have to provide the full set of parameters for each species, including those that are not used. To assist you, an example file is provided in the CLIMEX distribution. It is called SpeciesParameters.csv.

➤ **To specify a CLIMEX Species sequence**

1. In the CLIMEX Sequence Settings dialog, enter a descriptive **Name** for the Sequence – this name will appear in the list of Run Sequences for the model.
2. Enter the name of the *Sequence Specification File* (see below) into the **Input File** list box. The file can also be entered by searching for it using the **Browse** () button.
3. If the **Continue after run errors** button is checked, the sequence will continue on to the next run in the sequence even after errors (such as inconsistencies in parameter values) are found. The run that contains the error will be skipped.
4. The **Create Log File** check box, if checked, will create a log file that reports the status of each run in the sequence. This is particularly useful if the previous option has been specified, as it will report any skipped runs with the reason why the run was not included in the sequence.



Sequence Specification File

A typical Sequence Specification File takes the following form:

MetManager

```
1. "Output", "Summary", ".\Results\SomeParameters.csv"
2. "Output", "Log", ".\Results\SomeParameters.log"
3. "Ident", "Temperature Index", "Temperature Index", "Temperature
   Index", "Temperature Index", +
   "Moisture Index", "Moisture Index", "Moisture Index", "Moisture Index", +
   "Dry Stress", "Dry Stress", "Dry Stress", +
   ""
4. "", "Limiting low temperature", "Lower optimal temperature", +
   "Upper optimal temperature", "Limiting high temperature", +
   "Limiting low moisture", "Lower optimal moisture", +
   "Upper optimal moisture", "Limiting high moisture", +
   "", "Dry Stress Threshold", "Dry Stress Rate", +
   "Degree-days per Generation"
5. "Parameter Set 1", 18, 25, 37, 45, 0.045, 0.2, 1.2, 1.5, 1, 0.045, -0.025, 1200
6. "Parameter Set 2", 12, 25, 30, 33, 0, 0.01, 5, 10, 0, 0, 0, 0
7. "Parameter Set 3", 12, 23, 31, 36, 0.25, 0.4, 0.9, 1.3, 1, 0.25, -0.005, 0
8. "Parameter Set 4", 14, 26, 28, 30, 0.2, 0.6, 1, 2, 1, 1, 0.024, -0.041, 2, 0
9. "Parameter Set 5", 12, 25, 33, 36, 0.1, 0.5, 1.75, 2, 1, 0.1, -0.005, 380
10. "Parameter Set 6", 8, 20, 32, 39, 0.1, 0.4, 1.5, 4, 1, 0.1, -0.02, 0
11. "Parameter Set 7", 14, 26, 31, 36, 0.2, 0.6, 1, 2, 1, 0.2, -0.01, 0
12. "Parameter Set 8", 12, 22, 30, 35, 0.1, 0.3, 1, 1.5, 1, 0.02, -0.05, 622
13. "Parameter Set 9", 10, 20, 30, 33, 0.02, 0.08, 1, 1.2, 1, 0.02, -0.003, 450
```

The file is in comma-separated variable (.csv) format. **N.B. any line is considered to be a continuation of the previous line if the last character in the line is a “+”.** Character strings should be delimited by quote (") characters. Blank lines are permissible except between continued lines. Items may be spaced within lines (as shown in the example) to help in making the file more readable.

Line 1 specifies that the name of the output file that will contain the “Summary” information, i.e., the values of the indices selected for output in the **Summary Output Stream** specification (see Section 9.1).

If Multiple Files have been selected for this output stream, the specified name will form the root of the file name, with the identifier of each run appended to that root (for the example, the first output file would be `..\SomeParameters_Parameter Set 1.csv`).

Line 2 specifies the name of the log file. Information about each run in the sequence will be written to this file.

Lines 3 and 4 in conjunction specify the parameters and processes to be varied from the loaded species file. If column 1 of line 3 contains the phrase “Ident” (and column 1 of line 4 contains an empty string, as in the example above), then column 1 of each run specifier is taken to be the identifier label for the run. Otherwise, the runs will be labelled “Run 1”, etc. Each column in line 3 and 4 contains the name of a process/parameter pair. In the example above, column 2 specifies the “Temperature Index”/“Limiting low temperature” pair. If only a process is given, with the corresponding column in line 4 blank, then this indicates that this column specifies whether that process will be used or not used (see for example, column 10 in the example above). A “1” indicates that it is used, a “0” indicates that it is turned off. Some parameters do not have an associated process (for example, “Degree-days per Generation”) – in that case, the process should be a blank string.

The next lines each specify a parameter/process combination to be used in a run. If a run identifier is used, this must be the first string on each line. Each other column specifies the value for the corresponding parameter. If a column specifies whether or not a process is to be used, the 0 indicates the process is not used, 1 that it is used.

Generally, a CLIMEX Species Sequence will be run in conjunction with a MetManager (Location) sequence. To do this, a Compound Sequence needs to be created that contains both the CLIMEX Species Sequence and the appropriate Locations sequence. The CLIMEX Species Sequence will generally be the first sequence in order, so that the location set is run for each combination of parameters.

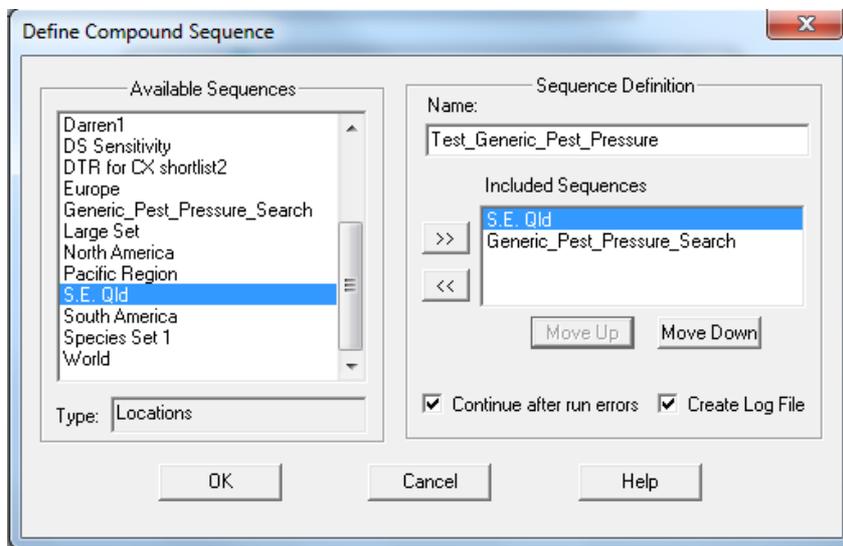
Note that each run in the sequence will use the parameters of the loaded species as a base set, varying those parameters specified for each run in the sequence. Care must be taken

that the specified parameter variations do not cause any parameter inconsistencies. If a large number of parameters need to be varied over a wide range, it may be useful to create a species parameter file that specifies a wide tolerance for parameter values. Even then, care is needed to ensure that relative values (for example, $DV1 > DV0$) are not violated.

The Species Sequence is an implicit compound sequence. If you select the Species sequence and run it, you will find that the results are ordered by location in the CLIMEX Metman Location selection. This is inconvenient for operations such as mapping the results, where we would be most interested in looking at the results one species parameter set at a time. To change this behaviour, we need to take a further step, and explicitly tell the program to run a compound sequence, species by location. This process is simple, and is explained in the next section.

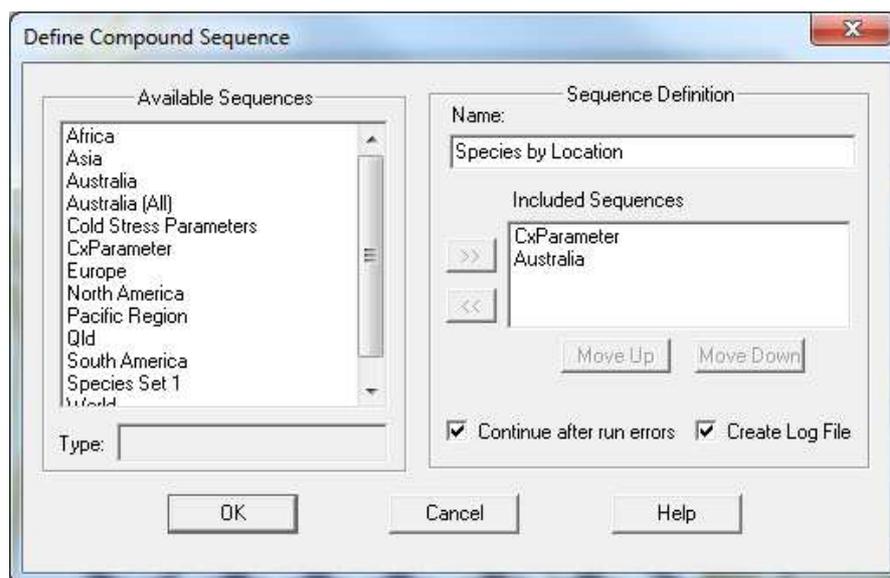
9.2.3 Compound sequence

The order of the included sequences is important. Generally, you will want to run species first, then locations.



If you specify a location selection and then specify a species sequence, the programme will create an automatic sequence on the fly. The order of this sequence will be in the order of appearance of the modules; locations, then species. In this case it will run each species sequence for the first location in the location set before moving to the next location. If you have specified an output stream the output will be partitioned (grouped) by individual location. This is probably not a very useful output for most situations.

To create a compound sequence you first need to define the individual sequence elements. The most common will probably be the Species Sequence by Location, or the Parameter Sequence by Location.



9.3 Creating composite risk models

It is becoming increasingly apparent that non-climatic anthropogenic factors can have a strong influence on species distributions and this is being reflected in collections of species location records, although the fact that the existence of a species in an area is due to the anthropogenic factor is usually not recorded, and has to be inferred. Irrigation is one such factor. The soil moisture regime experienced by crop plants in an irrigated situation will usually not reflect the pattern of soil moisture availability that would be modelled under a natural rainfall scenario. This has the effect of extending the potential distribution of the crop into otherwise inhospitable climatic habitat. The presence of the irrigation also has the effect of extending the potential distribution of pests, weeds, and diseases that are typically associated with the cropping system.

When confronted with distribution data that includes location records in such inhospitable habitat it is possible to either distort the species parameters to fit the data. In the case of the irrigation scenario, the parameters for the Dry Stress threshold and stress accumulation rate would have to be relaxed, and their meaning distorted. Another option is to try to model the processes explicitly (e.g., Beddow *et al.*, 2013). Fortunately, we now have access to valuable spatial datasets that describe where irrigation is practiced globally.

Here we describe an example of how to build a composite risk map. In addition to a working copy of CLIMEX, you will also need access to a Geographical Information System (GIS) or the R statistical package, and a spatial dataset describing irrigation practices.

1. Firstly, fit a Compare Locations model to the records you believe to be growing under natural rainfall conditions. You might use values of 0.1 for SMDS and SM0, to approximate permanent wilting point and a stress accumulation rate that seems reasonable for your species. Distribution points that appear outside of the suitable climate envelope may be growing under irrigated conditions.

2. Run an irrigation scenario using the top-up irrigation option with a reasonable rate, say 2.5 mm day⁻¹.
3. Save the results of the natural and irrigation scenarios to external files. You may have to ensure that you export the Latitude and Longitude values for each site, depending on how you decide to undertake the spatial analyses.
4. Find a spatial dataset of irrigation practices and spatially intersect the fishnet file that comes with CLIMEX, or use the pre-prepared fishnet file supplied with CLIMEX. It indicates the area of irrigation within each 10' grid cell based on the Global Map of irrigated Areas.
5. In the GIS or a spreadsheet, combine the results of the two scenarios using the irrigated area information. For each cell, if the area is irrigated, take the value from the irrigation scenario, else take the natural rainfall scenario result.

In the future, this procedure will likely be automated within CLIMEX.

10 The Meteorological Data Manager facility (MetManager)

The **MetManager** program is an application that can be used to store and manipulate long-term average meteorological data that is then accessible to CLIMEX through the **MetManager** modules. For the rest of this section, any reference to **MetManager** will be to the **MetManager** program.

The MetManager stores its data in Access databases with the file extension “.mm”. Each location’s data is identified by its name, longitude and latitude. The following fields are stored for each location:

Continent	The name of the continent that the location is in.
Country	The name of the country that the location is in.
State	The name of the state that the location is in.
Location Name	The name of the location
Longitude	Longitude of the location (in degrees) (-90 to +90)
Latitude	Latitude of the location (in degrees) (-180 to +180)
Elevation	Elevation (altitude) of the location (in metres)
Importance Level	An integer (0-5) that is used to select subsets of locations
Minimum Temperature	The monthly Average Daily Minimum Temperature (12 monthly values or 52 weekly values)
Maximum Temperature	The monthly Average Daily Maximum Temperature (12 monthly values or 52 weekly values)
Rainfall	The monthly Total Rainfall (12 monthly values or 52 weekly values)
9am RH	The monthly Average Daily 9am Relative Humidity (12 monthly values or 52 weekly values)
3pm RH	The monthly Average Daily 3pm Relative Humidity (12 monthly values or 52 weekly values)

Several user-defined fields are also available for storing such variables as, for example, Evaporation.

The database can contain data for a large number of locations, and a mechanism for retrieving or accessing a subset of these locations has been provided. A **Location Selection** is the set of locations each of which satisfies a specified number of conditions (the **Selection Criteria**). For example, if the Selection Criteria are that (1) the Country is Australia and (2) the Latitude is less than -30°, then the Location

Selection will include all the Australian locations in the database whose latitude is less than 30°. Many Selection Criteria can be defined, named and stored within the database. CLIMEX uses the resulting Location Selections as sequences. Note that the term Location Selection (or Selection) is commonly used for both a set of locations as well as the Selection Criteria that define that set.

The MetManager is started by clicking on the MetManager2 icon in the DYMEX Program Group. This displays the MetManager main menu, as shown (Fig. 10-1).

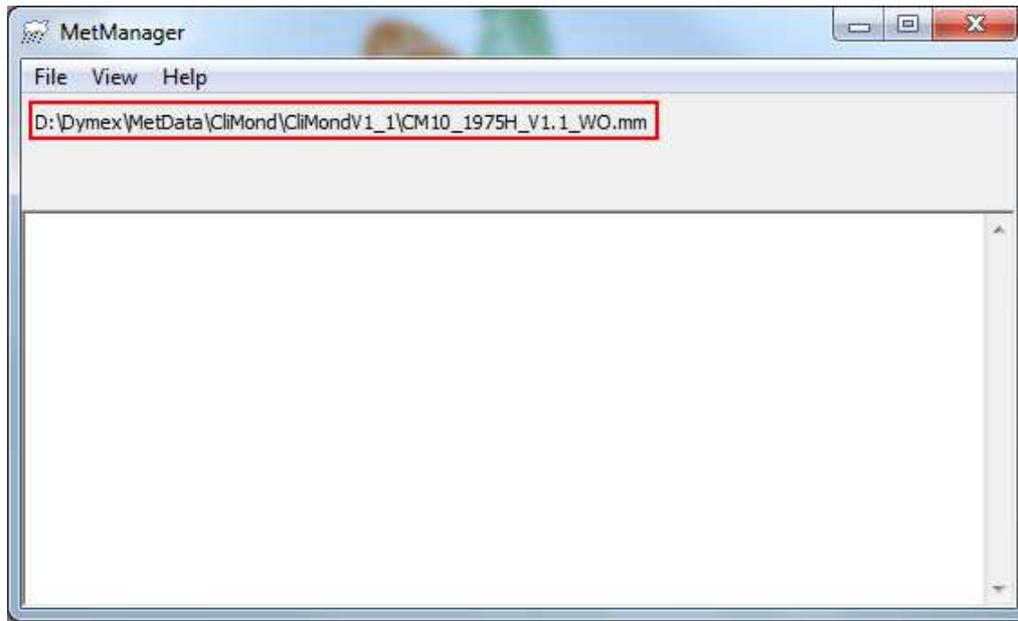


Fig. 10-1 The MetManager main menu, highlighting the currently open file.

The name of the current data-file is shown just below the menu list. A data file must be specified before any other actions are possible (either by creating a new data-file, or by opening an existing one).

10.1 Creating a new database

To create a new (empty) **MetManager** database file, select File | Create database... It leads to a file dialog, from whence the user can navigate to the appropriate directory and enter the name of the file. The file should be created as a MetManager file (with extension “.mm”). When the new file has been created, it is automatically opened as the current database file within **MetManager**.

Having created a new database, you will need to create at least one Metdata Location Selection. Firstly select the Edit Selections.... Option from the File Menu. This will open up the blank Edit Selections dialogue (Fig. 10-2 **Blank Edit Selections Dialogue**). Click on the Add... button to open the Add New Selection dialogue. Give the new selection a name that is meaningful to you, select the Metdata Entities for inclusion and any other constraints.

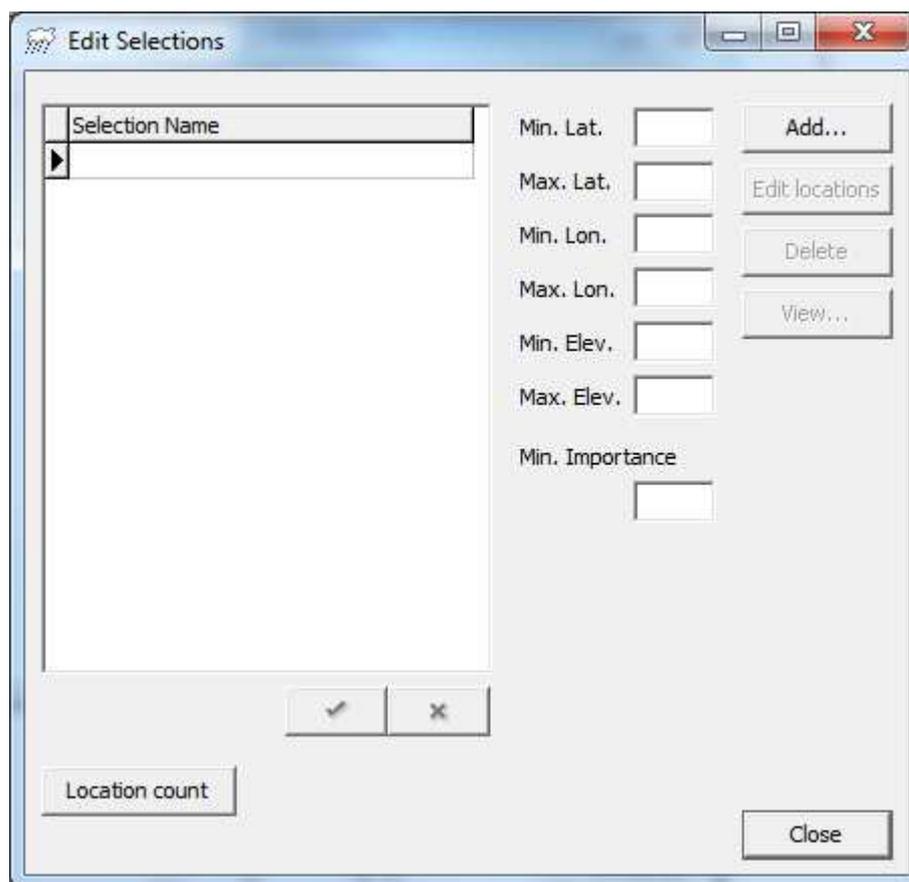


Fig. 10-2 Blank Edit Selections Dialogue

The **Add New Selection** dialog (Fig. 10-3) contains a cascading tree of Continents, Countries, States and Locations alongside the geographical and quality selection parameters. The selection indicator boxes have three states (empty=unselected, x=selected, and grey square, indicating that that element is partially selected). Clicking on a box selects the data element, and any sub-elements. In Fig. 10-3 Australia has been selected, which automatically selects all of the states within Australia. Because Australia is the only country selected above, Oceania is only partially selected.

Metadata **Entities** (Continents, Countries, States or Locations) are added to the Selection Criteria by clicking on the checkbox so that it shows a tick. Entities are deselected by clicking until the checkbox is empty. An entity that is only partially selected can only be unselected by removing the included, lower-level entities from the Selection Criteria.

The **Add New Selection** dialog also specifies other criteria that a location must conform to in order to be included in the set of locations making up the **Location Selection** (Fig. 10-3). The latitude and longitude limits can be used to set a bounding rectangle within which locations must lie, while the elevation limits set an elevation band. The **Min Importance Level** can be used to limit locations to those with a specified level of “importance”. This attribute is usually excluded from gridded datasets. When satisfied, click the OK button to return to the **Edit Selections** dialogue.

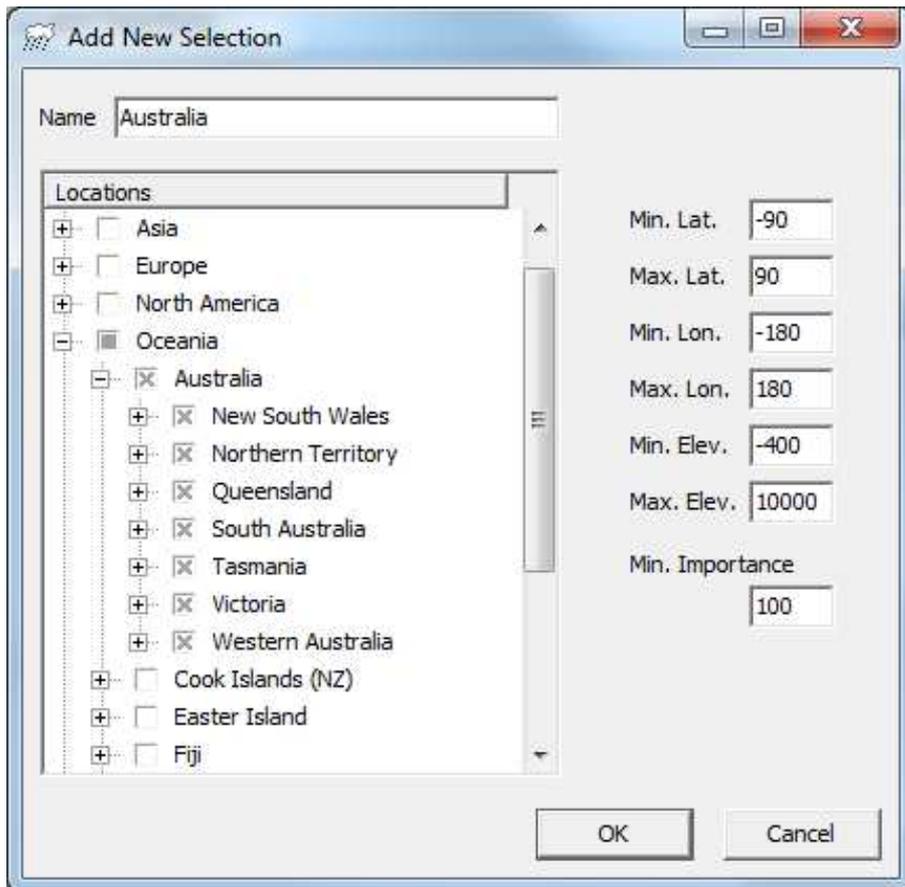


Fig. 10-3 The Selections dialog, used for specifying the Selection Criteria.

10.2 Edit Location Selections

When a database has been selected, the **Edit Locations** menu option on the File menu opens the **Edit Locations** dialogue. If no selections have been created, you first need to create them using the Add... button, which will open up the Add New Selection dialogue. This allows you to provide the Name for the Selection and to identify which Locations you want included in the Selection.

To create a new selection, click on the Add... button. The user is first prompted for the name of the new selection. Use a unique name that clearly describes the set. The Selections dialog is then presented (the **View** button goes directly there), where the Selection Criteria are specified (Fig. 10-3).

The **Edit Locations** option in the **Edit Selections** dialogue (Fig. 10-4) is used to open the Edit Selection Details dialogue (Fig. 10-5), which allows you to edit stored Location Selections (Selection Criteria).

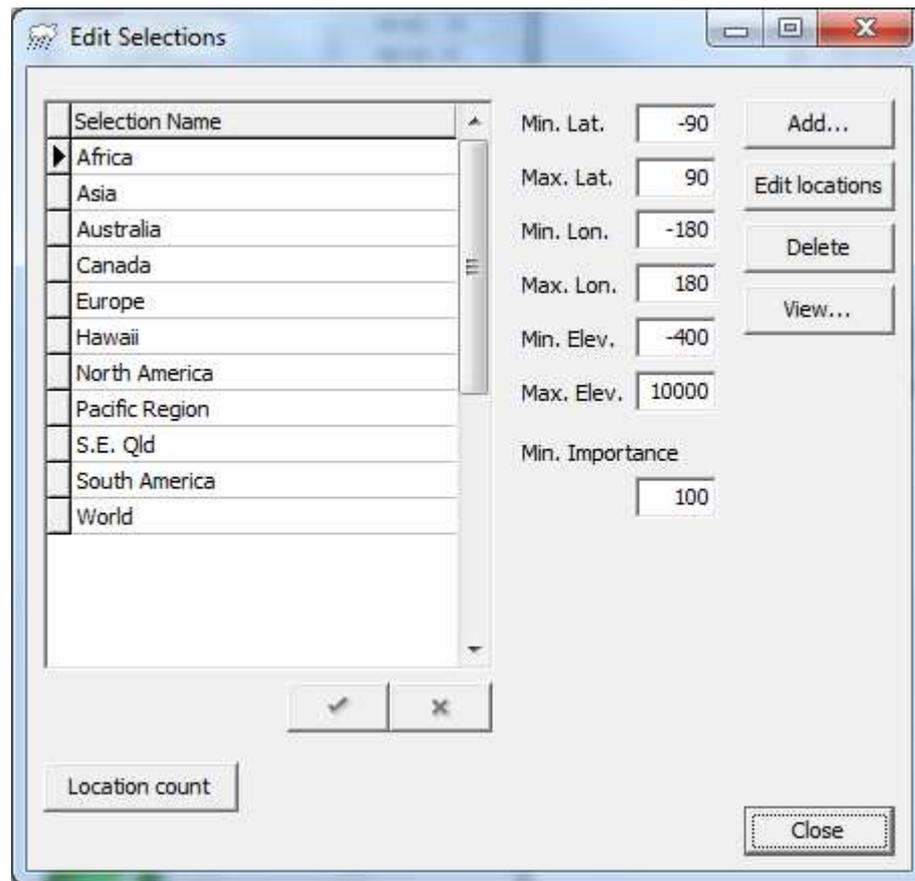


Fig. 10-4 The Edit Selections dialog.

To edit or delete a Selection, click on the appropriate Selection name in the list to highlight it and then click the **Edit Locations** or **Delete** button. Note that deleting a selection only deletes the Selection Criteria (it does not delete any data from the database). To see the number of locations in the database that match the selection criteria, click on the **Location count** button. For large datasets it may take a short while for a pop up window to appear with the count.

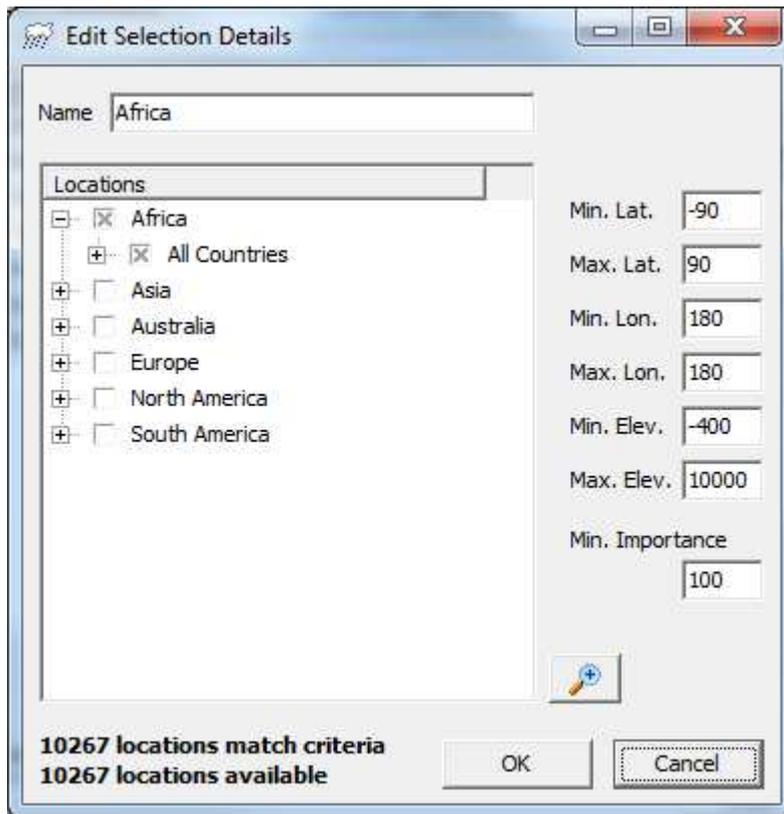


Fig. 10-5 Edit Selection details dialogue

10.3 Import meteorological data

In Metman2, the only met file formats that can be imported into CLIMEX are LOC/MET files and a MetManager Version 2 database file (.mm files). BMT files are no longer supported. The LOC/MET file format is a plain text format, detailed in Appendix 3.

To import data from a file, simply select **File | Import | From file...** menu item and select the appropriate *.loc file. The data will be imported and parsed according to the options you have set in the **File | Import | Configure Warnings...** dialogue. You will then be asked to select the location file (.LOC). Choose the directory where the file is located and select the file by clicking on its name and then the 'OK' button. There must also be a file containing the corresponding meteorological data (.MET file) in the same directory. The MetManager will now import the new data into the database.

The structure of your data files is very important (See Appendix 3). Example files of each (bolivia.loc and bolivia.met) are included in the ".\Dymex2\MetManager" directory. Use these files as a template for creating your files. The column setup will need to be EXACTLY the same as these example files otherwise MetManager will not import the data. Generally, an error message during the importation will identify any problems in the files. Note that if problems are found at any point during the import operation, no changes will be made to the database.

Review any import warnings after data has been imported into the database, this option allows review of the warnings that were generated during the import process.

MetManager

Several user-defined data fields are available for provision of additional data that describes a location (such as, for example, soil characteristics). These fields could be used to supply the data that drives the Physical and Biotic Substrate indices.

11 Useful modelling resources

Here is a set of resources for modelling species potential distributions, and an indication of where you can obtain them.

11.1 Species distribution data

11.1.1 Global Biodiversity Information Facility (GBIF)

GBIF is one of the most popular sources of species distribution data. You should be aware however that GBIF is a data harvesting portal that relies on the data integrity of the underlying databases (usually museums and the like). The data suffer from a wide variety of errors, including:

- miscoding of geographical coordinates for collection records (a typographical error).
- misrepresentation of geographical precision (e.g., attributing records where the collection locality was the country name to the geographical centroid of the country).
- using a coordinate of 0,0 where no location information has been specified.
- miscoding the sign on the latitude or longitude value. This places the record in the wrong hemisphere when it is mapped.
- attributing the collection location of the record to the laboratory where it is housed.
- misclassification of the taxonomic identity of the record

11.1.2 CABI

The CABI Invasive Species Compendium provides country-level mapping of species distributions for crop pests, as well as a useful guide to the apex literature about the major pests. It recently became freely available at <http://www.cabi.org/isc>.

Climate data

11.2 CliMond

Free climatological data in a range of formats including CLIMEX. The data are available at 10' and 30' spatial resolution for historical data (centred on 1975) and a set of future climate scenarios up to 2100. Available for free at www.climond.org.

Non-climatic habitat factors

Parameter Reference

11.3 Irrigation

The Global Map of Irrigated Areas (GMIA) was first produced by Siebert *et al.* (2005) and has been updated by Portmann *et al.* (2010). The database contains information on the area equipped for irrigation, the area actually irrigated, area equipped for irrigation with groundwater, with surface water and with non-conventional sources of water. The dataset is available from <http://www.fao.org/nr/water/aquastat/irrigationmap/index.stm>. This dataset is useful for defining specific areas in which an irrigation scenario should be applied, supporting the construction of CLIMEX models where the potential distribution is a composite mosaic of results derived from natural rainfall and irrigated scenarios (e.g., Guichard *et al.*, 2014).

11.4 Land Use

Finding the FAO land use data through the “front door” can be a little difficult. The data was compiled through the Land Degradation Assessment in Drylands project, which has since closed. The LADA Land Use system maps can be found here: http://www.fao.org/nr/lada/index.php?option=com_content&view=article&id=154&Itemid=184&lang=en.

11.5 Land Cover

The Global Land Cover (Globcover) project describes land cover. Users should be aware that landcover represents a snapshot of the vegetative state of the environment, which ignores the dynamic nature of landcover at a range of temporal scales. Consider that a broadscale production forestry landscape may be a mosaic of landcover types depending on where in the rotation cycle each forestry block falls. Thus, the landscape may consist of wasteland, ruderal grassland and shrubland or forest. If you are modelling a species that inhabits such a mosaic you may find problems with using this type of dataset as it may give unacceptable levels of model sensitivity or specificity as these land cover classes extend outside the forest landscape (e.g., Pitt *et al.*, 2011). The GLOBCOVER dataset may be found here: <http://due.esrin.esa.int/globcover/>.

11.6 Cropping areas

In cases where a species distribution is tightly constrained by the presence of a crop species it may be useful to frame the pest risk model so that it includes consideration of those locations where the crops are grown. The HarvestChoice project developed a series of maps portraying the distribution of the major crops of the world. They are available here (www.mapspam.org). Despite the comical name, the data are available at 5' spatial grid resolution as well as interactive maps. They are derived by allocating area wide production values into the habitat types where cropping is undertaken. This data was applied to a CLIMEX model of Wheat Stem Rust in order to estimate the likely global cost patterns if wheat resistance were low (Beddow *et al.*, 2013; Pardey *et al.*, 2013)

11.7 Crop Calendar

When fitting models it can be useful to understand the phenology of a crop-pest system. The Center for Sustainability and the Global Environment (SAGE) at the

University of Wisconsin-Madison has produced a series of useful datasets relating to cropping phenology. These are available as both digital maps and gridded spatial data of planting dates, harvesting dates, days between planting and harvest, and a few climatic metrics for each crop, such as the average temperature at which the crop is planted in each region. The data are available from this website http://www.sage.wisc.edu/download/sacks/crop_calendar.html , and the methods are described in Sacks *et al.* (2010).

11.8 Potential Vegetation

Occasionally it is useful to know the potential vegetation in an area so that you can explain why a species may or may not be found in an area. Ramankutty and Foley (1999) describes the modelling of this dataset available from http://www.sage.wisc.edu/download/potveg/global_potveg.html.

Mapping data

11.9 Global Administrative Areas (GADM)

GADM is a spatial database of the location of the world's administrative areas (or administrative boundaries) for use in GIS and similar software www.gadm.org/.

11.10 Placenames (Gazeteer)

The GeoNames geographical database (<http://www.geonames.org/>) is available for download free of charge under a creative commons attribution license. It contains over 10 million geographical names including 2.8 million populated places and 5.5 million alternate names. It is useful for geocoding locality names mentioned in the scientific literature and for checking the veracity of locations included in databases such as GBIF.

11.11 Digital elevation data

Digital elevation data allows you to check the lie of the land

<http://vterrain.org/Elevation/SRTM/>.

12 Frequently Asked Questions

What do different EI values mean, and how should I classify them?

Despite the fact that the Ecoclimatic Index is presented as a percentage value, it should be thought of as a quasi-ordinal variable. Technically, an EI value of (say) 10 is supposed to represent a climate whose suitability is 10% of the value of one in which the climate was optimal for the modelled organism throughout the entire year. However, the maximum value of EI in any given model is usually well short of this because climates tend to be seasonal, varying in their suitability for any given organism. Usually, the only situations receiving EI values in excess of 50 tend to be in the wet tropics.

Parameter Reference

The CLIMEX team noted that often CLIMEX users were misunderstanding the nature of the EI variable and were employing a false sense of precision in the percentile values, concerning themselves with very small differences in EI values. In an attempt to influence this, we started applying coarse somewhat arbitrary classification schemes (e.g., marginal, suitable, highly suitable and optimal, Kriticos *et al.*, 2003b; Kriticos *et al.*, 2003a). In an unforeseen quirk, the conversations shifted to what the class boundaries should be for the classes. Whilst Bob Sutherst attempted to provide some guidance (Sutherst, 2003; Sutherst and Maywald, 2005), the indicative ranges are frequently interpreted literally.

Ideally, the EI values should be interpreted in relation to meaningful data on the species abundance or impact. There have been few attempts to relate CLIMEX EI values to the abundance of organisms or pest impacts (Pinkard *et al.*, 2010; Kriticos *et al.*, 2013). In practice, there are many ways to gauge the abundance of an organism. The example with *Thaumetopoea pityocampa* highlighted the fact that impacts can be measured in a multiplicity of ways, and considerable effort was required to create a conformal dataset of impacts that could be related to the modelled EI values (Kriticos *et al.*, 2013). Clearly, the area of science concerned with measuring pest impacts is still in its infancy.

So, in practice, how should you classify your EI values? Firstly, consider how you fitted your stresses. Did you fit EI values of 1 or 2 to the most marginal sites? If so, you can say that values of 0-5 or 0-10 are marginal. Remember, the classification is arbitrary, and designed to convey meaning to the reader accurately, if not precisely. Next, working back from the highest values in your model, split them into 2 or 3 approximately equal interval, equal quantile or equal area classes. This decision can and should be guided by any information you may have on the relative climate suitability of different locations. The qualitative method of Pinkard *et al.* (2010) deserves further application and exploration. It offers a rapid means of calibrating the model of a pest species in terms of how difficult it is to manage.

How do I change the precision of the values presented in the output file?

With the Run Results window selected, click on the Save to File button and select an output format. In the Selected Variables box, click on the variable whose output precision you want to adjust and then click on the Column Format button. Set the No. of digits after the decimal point value to the desired level of precision.

My models seem to be running quite slowly, what could be wrong?

The latest version of CLIMEX has been sped up remarkably compared with the previous versions. There are four things known to slow the model performance:

1. A slow computer, with little memory.
2. Running very large climate datasets
3. Saving **Output Streams** to disk, and
4. Saving data to create **Map Sequences** (time series maps).

There is little you can do about the first issue. Regarding the second issue, very large datasets take longer to run than smaller ones. Generally, it is advisable to do the majority of your modelling using relatively coarse datasets, reserving the finer scale datasets for fine-tuning your results after you have your model looking approximately right. You may be surprised at how little gain in overall model accuracy there is with using finer resolution datasets (e.g., Kriticos and Leriche, 2009). In this study we found that significant gains were only apparent in areas of

high topographical relief at the range margins, where isolated pockets of suitable habitat were sometimes identified.

Unless you intend to save detailed output to disk for archival purposes or subsequent analyses outside of CLIMEX, you should uncheck all of the **Output Streams** options (Figure 12-1). This dialogue can be accessed from the **Execution** menu.

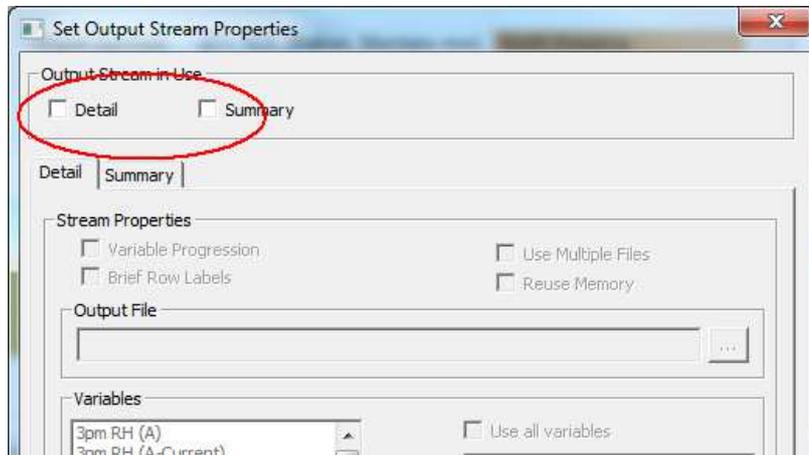


Figure 12-1 Output streams dialogue

In order to create Map Sequences, it is necessary to save intermediate results during a model run, slowing down the simulations. To speed up the simulations, ensure that the Retain results for map sequences option is turned off (Fig. 12-2). This dialogue is accessed from the **Preferences | Current Model...** menu option.

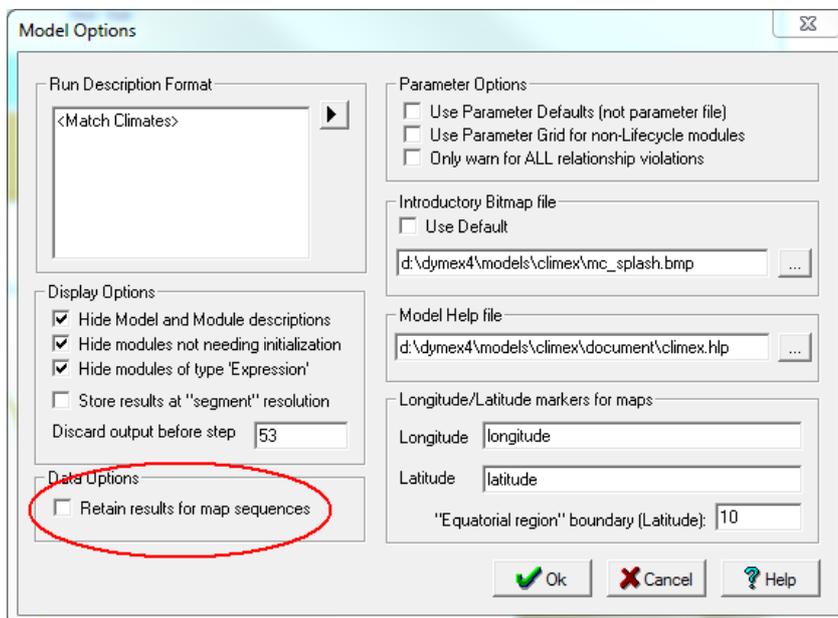


Fig. 12-2 Model Options dialogue, indicating where to select the “Retain results for map sequences” option.

13 References

- Akaike, H. (1974) A new look at statistical-model identification. *IEEE Transactions on Automatic Control*, **AC-19**, 716-723. doi:10.1109/tac.1974.1100705.
- Anderson, R. P., Lew, D. & Peterson, A. T. (2003) Evaluating predictive models of species' distributions: Criteria for selecting optimal models. *Ecological Modelling*, **162**, 211-232.
- Andrewartha, H. G. & Birch, L. C. (1954) *The distribution and abundance of animals*. University of Chicago Press, Chicago.
- Bajwa, W. I. & Kogan, M. (1997) Geographical distribution of the Codling Moth (*Cydia pomonella* L.). *Codling Moth Information Support System (CMISS)* (eds. Integr. Plant Protect. Center, Corvallis, OR, USA.
- Bajwa, W. I. & Kogan, M. (1998) Codling Moth Information Support System (CMISS): Bionomics of Codling Moth. (eds.
- Baskerville, G. L. & Emin, P. (1969) Rapid estimation of heat accumulation from maximum and minimum temperatures. *Ecology*, **50**, 514-517.
- Beasley, D., Bull, D. R. & Martin, R. R. (1993) An Overview of Genetic Algorithms: Part 1. Fundamentals. *University Computing*, **15**, 58-69.
- Beddow, J. M., Hurley, T. M., Kriticos, D. J. & Pardey, P. G. (2013) Measuring the global occurrence and probabilistic consequences of wheat stem rust. pp. 24. HarvestChoice, St. Paul.
- Brown, J. H., Stevens, G. C. & Kaufman, D. M. (1996) The geographic range: size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics*, **27**, 597-623.
- Chamberlin, T. (1965) The method of multiple working hypotheses. *Science*, **148**, 754-759. doi:10.1126/science.148.3671.754.
- Chapman, A. D. (2005a) *Principles and methods of data cleaning*. GBIF.
- Chapman, A. D. (2005b) *Principles of data quality*. GBIF.
- Commonwealth Institute of Entomology (1976) Distribution Maps of Pests, Series A (Agricultural), Map No. 9 (revised). Commonwealth Agricultural Bureaux, London.
- Davis, A. J., Lawton, J. H., Shorrocks, B. & Jenkinson, L. S. (1998) Individualistic species responses invalidate simple physiological models of community dynamics under global environmental change. *Journal of Animal Ecology*, **67**, 600-612.
- de Villiers, M., Hattingh, V. & Kriticos, D. J. (2012) Combining field phenological observations with distribution data to model the potential range distribution of the fruit fly *Ceratitis rosa* Karsch (Diptera: Tephritidae). *Bulletin of Entomological Research*, **103**, 60-73. doi:DOI: 10.1017/S0007485312000454.

- Dhileepan, K., Senaratne, K. A. D. W. & Raghu, S. (2006) A systematic approach to biological control agent exploration and prioritisation for prickly acacia (*Acacia nilotica* ssp. *indica*). *Australian Journal of Entomology*, **45**, 303-307. doi:10.1111/j.1440-6055.2006.00555.x.
- Duboz, R., Versmisse, D., Travers, M., Ramat, E. & Shin, Y. J. (2010) Application of an evolutionary algorithm to the inverse parameter estimation of an individual-based model. *Ecological Modelling*, **221**, 840-849. doi:10.1016/j.ecolmodel.2009.11.023.
- Fielding, A. H. & Bell, J. F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38-49.
- Fitzpatrick, E. A. (1963) Estimates of pan evaporation from mean maximum temperature and vapour pressure. *Journal of Applied Meteorology*, **2**, 780-792.
- Fitzpatrick, E. A. & Nix, H. A. (1969) A model for simulating soil water regime in alternating fallow-crop systems. *Agricultural Meteorology*, **6**, 303-319. doi:[http://dx.doi.org/10.1016/0002-1571\(69\)90023-5](http://dx.doi.org/10.1016/0002-1571(69)90023-5).
- Geier, P. (1963) The life history of the codling moth, *Cydia pomonella* (L.) (Lepidoptera, Tortricidae) in the Australian Capital Territory. *Australian Journal of Zoology*, **11**, 323-367.
- Guichard, S., Guis, H., Tran, A., Garros, C., Balenghien, T. & Kriticos, D. J. (2014) Worldwide niche and future potential distribution of *Culicoides imicola*, a major vector of bluetongue and African horse sickness viruses. *PLoS One*, **9**, e112491.
- Hall, M. & Wall, R. (1995) Myiasis of humans and domestic animals. *Advances in Parasitology*, Vol 35, **35**, 257-334.
- Helton, J. C., Davis, F. J. & Johnson, J. D. (2005) A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling. *Reliability Engineering & System Safety*, **89**, 305-330. doi:10.1016/j.res.2004.09.006.
- Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, **199**, 142-152.
- Jarošík, V., Honěk, A. & Dixon, A. F. (2002) Developmental rate isomorphy in insects and mites. *The American Naturalist*, **160**, 497-510.
- Jarosik, V., Honek, A., Magarey, R. D. & Skuhrovec, J. (2011) Developmental database for phenology models: Related insect and mite species have similar thermal requirements. *Journal of Economic Entomology*, **104**, 1870-1876. doi:10.1603/ec11247.
- Jarošík, V., Kratochvíl, L., Honěk, A. & Dixon, A. F. (2004) A general rule for the dependence of developmental rate on temperature in ectothermic animals. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **271**, S219-S221.
- Johnson, C. J., N., S. E., Merrill, E. H., McDonald, T. L. & Boyce, M. S. (2006) Resource selection functions based on use availability data: Theoretical

Parameter Reference

- motivation and evaluation methods. *The Journal of Wildlife Management*, **70**, 347-357.
- Keane, R. M. & Crawley, M. J. (2002) Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution*, **17**, 164-170.
- Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334-350. doi:10.1111/j.1461-0248.2008.01277.x.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B. & Rubel, F. (2006) World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, **15**, 259-263.
- Krebs, C. J. (1978) *Ecology. The experimental analysis of distribution and abundance*. Harper and Row, New York.
- Kriticos, D. J. (2012) Regional climate-matching to estimate current and future biosecurity threats. *Biological Invasions*, **14**, 1533-1544. doi:10.1007/s10530-011-0033-8.
- Kriticos, D. J., Lamoureaux, S., Bourdôt, G. W. & Pettit, W. (2004) Nassella tussock: current and potential distributions in New Zealand. *New Zealand Plant Protection*, **57**, 81-88.
- Kriticos, D. J. & Leriche, A. (2009) The effects of spatial data precision on fitting and projecting species niche models. *Ecography*, **33**, 115-127. doi:DOI: 10.1111/j.1600-0587.2009.06042.x.
- Kriticos, D. J., Leriche, A., Palmer, D., Cook, D. C., Brockerhoff, E. G., Stephens, A. E. A. & Watt, M. S. (2013) Linking climate suitability, spread rates and host-impact when estimating the potential costs of invasive pests. *PLoS One*, **8**, e54861. doi:10.1371/journal.pone.0054861.
- Kriticos, D. J., Ota, N., Hutchison, W. D., Beddow, J., Walsh, T., Tay, W. T., Borchert, D. M., Paula-Moreas, S. V., Czapak, C. & Zalucki, M. P. (2015) The Potential distribution of invading *Helicoverpa armigera* in North America: Is it just a matter of time? *PLoS One*, **10**, e0119618. doi:10.1371/journal.pone.0119618.
- Kriticos, D. J., Potter, K. J., Alexander, N., Gibb, A. R. & Suckling, D. M. (2007) Using a pheromone lure survey to establish the native and potential distribution of an invasive lepidopteran, *Uraba lugens*. *Journal of Applied Ecology*, **44**, 853-863.
- Kriticos, D. J., Sutherst, R. W., Brown, J. R., Adkins, S. A. & Maywald, G. F. (2003a) Climate change and biotic invasions: A case history of a tropical woody vine. *Biological Invasions*, **5**, 145-165.
- Kriticos, D. J., Sutherst, R. W., Brown, J. R., Adkins, S. A. & Maywald, G. F. (2003b) Climate change and the potential distribution of an invasive alien plant: *Acacia nilotica* ssp. *indica* in Australia. *Journal of Applied Ecology*, **40**, 111-124.
- Kriticos, D. J., Yonow, T. & McFadyen, R. E. (2005) The potential distribution of *Chromolaena odorata* (Siam weed) in relation to climate. *Weed Research*, **45**, 246-254.

- Lobo, J. M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145-151.
- Makowski, D. (2013) Uncertainty and sensitivity analysis in quantitative pest risk assessments: Practical rules for risk assessors. *Neobiota*.
- Miller, D. R. (1974) Sensitivity analysis and validation of simulation models. *Journal of Theoretical Biology*, **48**, 345-360.
- New, M. & McSweeney, C. (2010) Climate and Invasive Species: The Limits to Climate Information. *Bioinvasions and Globalization: Ecology, Economics, Management and Policy* (eds C. Perrings, H. A. Mooney & M. Williamson), pp. 30-41. Oxford University Press, Oxford.
- Pardey, P. G., Beddow, J. M., Hurley, T. M., Kriticos, D. J., Park, R. F., Duveiller, E., Sutherst, R. W., Burdon, J. & Hodson, D. (2013) Right-sizing stem rust research. *Science*, **340**, 147-148.
- Pimentel, D. (2002) *Biological Invasions: Economic and environmental costs of alien plant, animal and microbe species*. CRC Press, New York, USA.
- Pinkard, E. A., Kriticos, D. J., Wardlaw, T. J. & Carnegie, A. J. (2010) Estimating the spatio-temporal risk of disease epidemics using a bioclimatic niche model. *Ecological Modelling*, **221**, 2828-2838. doi:10.1016/j.ecolmodel.2010.08.017.
- Pitcairn, M. J., Pickel, C., Falcon, L. A. & Zalom, F. G. (1991) Development and survivorship of *Cydia pomonella* (L.) (Lepidoptera: Tortricidae) at ten constant temperatures. *Pan-Pacific Entomologist*, **67**, 189-194.
- Pitt, J. P. W., Kriticos, D. J. & Dodd, M. B. (2011) Limits to simulating the future spread of invasive species: *Buddleja davidii* in New Zealand. *Ecological Modelling*, **222**, 1880-1887.
- Portmann, F. T., Siebert, S. & Doll, P. (2010) MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, **24**, Gb1011. doi:10.1029/2008gb003435.
- Ramankutty, N. & Foley, J. A. (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, **13**, 997-1027.
- Renner, G. & Ekart, A. (2003) Genetic algorithms in computer aided design. *Computer-Aided Design*, **35**, 709-726. doi:10.1016/s0010-4485(03)00003-4.
- Robertson, M. P., Kriticos, D. J. & Zachariades, C. (2008) Climate matching techniques to narrow the search for biological control agents. *Biological Control*, **46**, 442-452.
- Sacks, W. J., Deryng, D., Foley, J. A. & Ramankutty, N. (2010) Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, **19**, 607-620. doi:10.1111/j.1466-8238.2010.00551.x.
- Sands, P. & Hughes, R. D. (1976) A simulation model of seasonal changes in the value of cattle dung as a food resource for an insect. *Agricultural Meteorology*, **17**, 161-183.

Parameter Reference

- Shel'deshova, G. G. (1967) Ecological factors determining the distribution of the codling moth *Laspeyresia pomonella* L. (Lepidoptera, Tortricidae) in the northern and southern hemispheres. *Entomological review*, **46**, 349-361.
- Shelford, V. E. (1963) *The ecology of North America*. University of Illinois Press, Urbana, IL., USA.
- Siebert, S., Doll, P., Hoogeveen, J., Faures, J. M., Frenken, K. & Feick, S. (2005) Development and validation of the global map of irrigation areas. *Hydrology and Earth System Sciences*, **9**, 535-547.
- Stinner, R. E., Gutierrez, A. P. & Butler, G. D. (1974) An algorithm for temperature-dependent growth rate calculation. *The Canadian Entomologist*, **106**, 519-523.
- Sutherst, R. W. (2003) Prediction of species geographical ranges. *Journal of Biogeography*, **30**, 805-816.
- Sutherst, R. W. & Bourne, A. S. (2009) Modelling non-equilibrium distributions of invasive species: a tale of two modelling paradigms. *Biological Invasions*, **11**, 1231-1237.
- Sutherst, R. W., Ingram, J. S. I. & Scherm, H. (1998) Global change and vector-borne diseases. *Parasitology Today*, **14**, 297-299.
- Sutherst, R. W. & Maywald, G. (2005) A climate model of the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera : Formicidae): Implications for invasion of new regions, particularly Oceania. *Environmental Entomology*, **34**, 317-335.
- Sutherst, R. W. & Maywald, G. F. (1985) A computerised system for matching climates in ecology. *Agriculture, Ecosystems and Environment*, **13**, 281-299.
- Sutherst, R. W., Maywald, G. F. & Skarratt, D. B. (1995) Predicting insect distributions in a changed climate. *Insects in a Changing Environment* (eds R. Harrington & N. E. Stork), pp. 59-91. Academic Press, London.
- Taylor, S. & Kumar, L. (2012) Sensitivity analysis of CLIMEX parameters in modelling potential distribution of *Lantana camara* L. *PLoS One*, **7**, 1-16. doi:10.1371/journal.pone.0040969.
- Tetens, O. (1930) Uber einige meteorologische Begriffe. *Zeitschrift fur Geophysik*, **6**, 297.
- Trudgill, D. L. (1995) Why do tropical poikilothermic organisms tend to have higher threshold temperatures for development than temperate ones. *Functional Ecology*, **9**, 136-137.
- van der Ploeg, R. R., Böhm, W. & Kirkham, M. B. (1999) On the origin of the theory of mineral nutrition of plants and the law of the minimum. *Soil Science Society of America Journal*, **63**, 1055-1062.
- Venette, R. C., Kriticos, D. J., Magarey, R., Koch, F., Baker, R. H. A., Worner, S., Gómez, N. N., McKenney, D., Dobesberger, E., Yemshanov, D., De Barro, P., Hutchison, W. D., Fowler, G., Kalaris, T. & Pedlar, J. (2010) Pest risk maps for invasive alien species: a roadmap for improvement. *Bioscience*, **80**, 349-362.
- Vera, M. T., Rodriguez, R., Segura, D. F., Cladera, J. L. & Sutherst, R. W. (2002) Potential geographical distribution of the Mediterranean fruit fly, *Ceratitidis*

- capitata* (Diptera : Tephritidae), with emphasis on Argentina and Australia. *Environmental Entomology*, **31**, 1009-1022.
- Weedon, G. P., Gomes, S., Viterbo, P., Österle, H., Adam, J. C., Bellouin, N., Boucher, O. & Best, M. (2010) The WATCH Forcing Data 1958-2001: a meteorological forcing dataset for land surface and hydrological models. *WATCH Technical Report*, pp. 41pp.
- Wharton, T. N. & Kriticos, D. J. (2004) The fundamental and realised niche of the Monterey pine aphid, *Essigella californica* (Essig) (Hemiptera: Aphididae): implications for managing softwood plantations in Australia. *Diversity and Distributions*, **10**, 253-262.
- Woodward, F. I. (1987) *Climate and plant distribution*. Cambridge University Press.
- Yemshanov, D., Koch, F. H., Ducey, M. J., Haack, R. A., Siltanen, M. & Wilson, K. (2013) Quantifying uncertainty in pest risk maps and assessments: Adopting a risk-averse decision maker's perspective. *Neobiota*.
- Zalucki, M. P., Rochester, W. A., Oberhauser, K. & Solensky, M. (2004) Spatial and temporal population dynamics of monarchs down-under: lessons for North America. *Monarch butterfly biology and conservation* (eds, pp. 219-228. Cornell University Press, Ithaca, USA.

Parameter Reference

Appendix 1. Parameter Mnemonic Reference.

Growth parameters

Temperature Index

DV0 Limiting low temperature. No population growth takes place at or below this average weekly temperature.

DV1 Lower optimal temperature. The lower limit of the range of ideal temperatures for population growth. Population growth rate is reduced if the average weekly minimum temperature is below this value.

DV2 Upper optimal temperature. The upper limit of the range of ideal temperatures for population growth. Population growth rate is reduced if the average weekly maximum temperature exceeds this value.

DV3 Limiting high temperature. No population growth takes place when the average weekly maximum temperature equals or exceeds this value.

PDD Minimum degree-days above DV0 necessary to complete a generation. For univoltine species, a minimum number of degree-days must be accumulated in a year in order to complete a generation and so survive at a given location. If the value of PDD is not reached at a particular location, the EI for that location will be 0. The default value of PDD is zero.

Moisture Index

SM0 Limiting low moisture. No population growth takes place at or below this average weekly soil moisture.

SM1 Lower optimal moisture. The lower limit of the range of ideal soil moisture for population growth. Population growth rate is reduced if the average weekly soil moisture is below this value.

SM2 Upper optimal moisture. The upper limit of the range of ideal soil moisture for population growth. Population growth rate is reduced if the average weekly soil moisture is above this value.

SM3 Limiting high moisture. No population growth takes place at or above this average weekly soil moisture.

The following are used only in the Species Interaction application:

LGMC Lower Competitor Growth Rate. The growth rate of the competing species below which it does not affect the available soil moisture. With a growth rate of the competing species above this value, less moisture is available for the current species.

MIF Moisture Inhibitor Floor. The lowest value that the soil moisture output can reach (when the competing species has a Growth Index of 1).

Diapause Index

DPDO Diapause Induction Day length. This is the average daylength that induces diapause. The default value is 0, indicating no diapause.

DPT0 Diapause Induction Temperature. This is the average weekly minimum (for winter diapause) or maximum temperature (for summer diapause) that induces diapause.

DPT1 Diapause Termination Temperature. This is the average weekly minimum (for winter diapause) or maximum (for summer diapause) temperature that terminates diapause.

DPD Diapause Development Days. This is the minimum number of days below (winter diapause) or above (summer diapause) DPT0 required for diapause development to be complete. If the required number of days in diapause (DPD) are not achieved, the species is assumed not to have survived diapause and the annual GI becomes zero. The default value for DPD is 0, which corresponds to a facultative diapause.

DPSW Indicator for Summer or Winter Diapause. The value of this parameter determines whether the species undergoes summer or winter diapause. A value of 0 indicates winter diapause, a value of 1 indicates summer diapause.

Light Index

LT0 is the daylength (in hours) above which growth is at a maximum.

LT1 is the daylength below which growth is 0.

Radiation Index

RV0 Limiting low radiation. No population growth takes place at or below this value of average weekly radiation.

RV1 Lower optimal radiation. The lower limit of the range of ideal radiation for population growth. Population growth rate is reduced if the average weekly radiation falls below this value.

RV2 Upper optimal radiation. The upper limit of the range of ideal radiation for population growth. Population growth rate is reduced if the average weekly radiation is above this value.

RV3 Limiting high radiation. No population growth takes place at or above this value average weekly radiation.

The following are used only in the Species Interaction application:

LGRC Lower Competitor Growth Rate. The growth rate of the competing species below which it does not affect the available radiation. With a growth rate of the competing species above this value, less radiation is available for the current species.

RIF Radiation Inhibitor Floor. The lowest value that the radiation output can reach (when the competing species has a Growth Index of 1).

Parameter Reference

Physical Substrate Index

SVPO Lower Physical Substrate Variable Threshold. No population growth takes place at or below this value of the average weekly physical substrate variable.

SVPI Lower optimal Physical Substrate Variable Value. The lower limit of the range of physical substrate variable values ideal for population growth. Population growth rate is reduced if the average weekly value of the physical substrate variable falls below this value.

SVP2 Upper optimal Physical Substrate Variable Value. The upper limit of the range of physical substrate variable values ideal for population growth. Population growth rate is reduced if the average weekly value of the physical substrate variable is above this value.

SVP3 Upper Physical Substrate Variable Threshold. No population growth takes place at or above this value of the average weekly physical substrate variable.

Biotic Substrate Index

SVBO Lower Biotic Substrate Variable Threshold. No population growth takes place at or below this value of the average weekly biotic substrate variable.

SVBI Lower optimal Biotic Substrate Variable Value. The lower limit of the range of biotic substrate variable values ideal for population growth. Population growth rate is reduced if the average weekly value of the biotic substrate variable falls below this value.

SVB2 Upper optimal Biotic Substrate Variable Value. The upper limit of the range of biotic substrate variable values ideal for population growth. Population growth rate is reduced if the average weekly value of the biotic substrate variable is above this value.

SVB3 Upper Biotic Substrate Variable Threshold. No population growth takes place at or above this value of the average weekly biotic substrate variable.

The following is used only in the Species Interaction application:

Biotic Index

SIP0 Species Interaction Parameter. This proportion of the Growth Index of the "other" species is added to (or subtracted from, if the parameter is negative) the raw weekly Growth Index to obtain the final weekly Growth Index.

Stress Parameters

As stresses are experienced when climatic extremes outside the ranges suitable for growth occur, the threshold values must always be outside the ranges for growth (for example, TTCS < DV0, TTHS > DV3, etc).

Cold Stress

Cold stress can occur in three different ways. In the degree-day method, stress occurs because the days are not warm enough to maintain metabolism. In the lethal temperature methods, stress occurs in response to excessively low temperatures (either minima or averages). Multiple stress types may occur for the one species.

TTCS Cold Stress Temperature Threshold. This parameter represents the average weekly minimum temperature below which Cold Stress accumulates.

THCS Cold Stress Temperature Rate. This is the rate at which Cold Stress accumulates once temperatures drop below the threshold value of TTCS.

DVCS Cold Stress Day-degree Temperature Threshold. The temperature threshold that is used for the cold stress day-degree calculations.

DTCS Cold Stress Day-degree Threshold. Cold stress will begin to accumulate when this threshold number of degree-days above DVCS is not achieved.

DHCS Cold Stress Degree-day Rate. This is the rate at which Cold Stress accumulates once the threshold number of degree-days (DTCS) above DTCS is not achieved.

TTCSA Cold Stress Temperature Threshold (Average). This parameter represents the mean weekly average temperature below which Cold Stress accumulates.

THCSA Cold Stress Temperature Rate (Average). This is the rate at which Cold Stress accumulates once average temperatures drop below the threshold value of TTCS.

Heat Stress

Heat stress can occur in two different ways. In the degree-day method, stress occurs because the days are too hot to allow normal metabolism. In the lethal temperature method, stress occurs because of excessively high maximum temperatures.

TTHS Heat Stress Temperature Threshold. This parameter represents the average weekly maximum temperature above which Heat Stress accumulates.

THHS Heat Stress Temperature Rate. This is the rate at which Heat Stress accumulates once temperatures exceed the threshold value of TTHS.

DTHS Heat Stress Threshold. Heat stress will begin to accumulate when this threshold number of degree-days above DV3 is exceeded.

Parameter Reference

DHHS Heat Stress Degree-day Rate. This is the rate at which Heat Stress accumulates once the threshold number of degree-days above DV3 (DTHS) is exceeded.

Dry Stress

SMDS Dry Stress Threshold. Dry stress is accumulated when the average weekly soil moisture level drops below this value.

HDS Dry Stress Rate. This is the rate at which Dry Stress accumulates once the soil moisture drops below the Dry Stress threshold of SMDS.

Wet Stress

SMWS Wet Stress Threshold. Wet stress is accumulated when the average weekly soil moisture levels exceed this value.

HWS Wet Stress Rate. This is the rate at which Wet Stress accumulates once the soil moisture exceeds the Wet Stress threshold of SMWS.

Temperature/Moisture Stress Interactions

Interactive effects occur when combinations of stress values in their own right are not lethal, but are when they occur simultaneously (eg. Cold and wet or hot and wet).

Cold-Dry Stress

DTCD Cold-Dry Degree-day Threshold. Cold-Dry Stress will begin to accumulate when this threshold number of degree-days above DV0 is not achieved.

MTC Cold-Dry Moisture Threshold. This is the average weekly soil moisture level below which Cold-Dry Stress accumulates.

PCD Cold-Dry Stress Rate. This is the rate at which Cold-Dry Stress accumulates once the threshold number of degree-days > DV0 is not reached **and** the average soil moisture level drops below the threshold value.

Cold-Wet Stress

DTCW Cold-Wet Temperature Threshold. Cold-Wet Stress will begin to accumulate when this threshold number of degree-days above DV0 is not achieved.

MTCW Cold-Wet Moisture Threshold. This is the average weekly soil moisture level above which Cold-Wet Stress accumulates.

PCW Cold-Wet Stress Rate. This is the rate at which Cold-Wet Stress accumulates once the threshold number of degree-days > DV0 is not reached **and** the average soil moisture level exceeds the threshold value of MTCW.

Hot-Dry Stress

TTHD Hot-Dry Temperature Threshold. This is the average weekly maximum temperature above which Hot-Dry Stress accumulates.

MTHD Hot-Dry Moisture Threshold. This is the average weekly soil moisture level below which Hot-Dry Stress accumulates.

PHD Hot-Dry Stress Rate. This is the rate at which Hot-Dry Stress accumulates once the average maximum temperature exceeds the threshold of TTHD and the average soil moisture level drops below the threshold value of MTHD.

Hot-Wet Stress

TTHW Hot-Wet Temperature Threshold. This is the average weekly maximum temperature above which Hot-Wet Stress accumulates.

MTHW Hot-Wet Moisture Threshold. This is the average weekly soil moisture level above which Hot-Wet Stress accumulates.

PHW Hot-Wet Stress Rate. This is the rate at which Hot-Wet Stress accumulates once the average maximum temperature and the average soil moisture level exceed the threshold values of TTHW and MTHW.

14 Appendix 2. The LOC/MET file format.

The two files hold the location data (.LOC files) and the meteorological data (.MET files) respectively. Each of these files can be in either a fixed-field or comma-delimited format. The MetManager automatically detects whether a file is in fixed-field or comma-delimited format.

(a) Fixed field format

An example of the *.loc file (ie bolivia.loc) is given below and could be used to add two new locations (Chulumani and Patacamaya) in Bolivia.

```
*1.00 SOUTH AMERICA
*1.01 Bolivia
#*   nostates
0   Chulumani      16.4 S  67.5 W 1580
0   Patacamaya     17.2 S  67.9 W 3789
```

The file starts with a header that includes the continent, country, and optionally, a state. Note that these items appear on separate lines of which the first two are prefixed by a * and four numeric characters followed by a space and the name of the continent or country. *1.00 denotes continent and 1.xx (where xx is any non zero integer) denotes country. The optional line for state is prefixed by # * and four spaces followed by the name of the state or 'nostates', where state definitions do not exist (the state line may be omitted altogether if you do not intend to use states for the entry of new locations). The name of the continent **MUST** be one of the continents in the CLIMEX database, or the import will fail. The name of the country and state entered should correspond to one already being used in the CLIMEX database. If an unknown name is used for a country or state, a new country or state with that name will be created and the new locations will not be added to that respective continent and/or country and/or state. Be particularly careful with spelling of these items to avoid errors.

Following the header on a new line appears the **Location Data** for each new location. This includes information about the importance level, the name of the location, its coordinates and elevation (in metres. The Location Data is of a fixed format with data in columns with each location line specified as below (for columns not referred to, insert a normal space).

```
1      Importance level code (0-5; blank=0)
7-25   Location name
26-29  Latitude
31     North/South indicator (N or S)
34-38  Longitude
40     East/West indicator (E or W)
41-45  Elevation (in metres)
```

The **Meteorological Data** file (example below) holds all the climate data for each location. For each month (1 - 12) the file holds average maximum daily temperature, average minimum daily temperature, average monthly rainfall, average daily relative humidity at 9 am, average daily relative humidity at 3 pm and the 10-character location code defined in the Location Data file above. Note in the example below that the 3pm Relative Humidity values are all zero. CLIMEX interprets these as being absent, and calculates them using the 9am value (9am RH x 0.85). The Meteorological Data file is given a **.met** file extension (e.g., bolivia.met).

```

1 26.9 15.9 171 12.9 .0 CHULUMANI
2 26.8 15.8 223 12.5 .0 CHULUMANI
3 26.9 15.5 141 12.1 .0 CHULUMANI
4 26.9 14.7 56 11.6 .0 CHULUMANI
5 25.4 14.6 38 11.2 .0 CHULUMANI
6 24.5 11.5 17 11.0 .0 CHULUMANI
7 24.7 11.3 20 11.1 .0 CHULUMANI
8 26.3 12.3 58 11.5 .0 CHULUMANI
9 27.0 13.4 97 11.9 .0 CHULUMANI
10 27.7 14.9 99 12.3 .0 CHULUMANI
11 28.5 15.1 117 12.7 .0 CHULUMANI
12 27.5 15.7 188 13.0 .0 CHULUMANI
1 22.7 5.8 88 12.9 .0 PATACAMAYA
2 21.9 6.0 69 12.5 .0 PATACAMAYA
3 21.1 4.7 47 12.1 .0 PATACAMAYA
4 19.3 2.1 15 11.6 .0 PATACAMAYA
5 16.5 -1.6 8 11.2 .0 PATACAMAYA
6 14.6 -4.8 2 11.0 .0 PATACAMAYA
7 14.9 -4.7 1 11.1 .0 PATACAMAYA
8 16.8 -2.6 9 11.4 .0 PATACAMAYA
9 19.1 1.5 37 11.9 .0 PATACAMAYA
10 22.6 3.0 18 12.4 .0 PATACAMAYA
11 24.3 4.5 33 12.8 .0 PATACAMAYA
12 22.6 5.7 76 13.0 .0 PATACAMAYA

```

The met file data must contain the same number of locations and be in the same order as the loc file entries. The format of the file is as follows (columns not used are spaces):

```

1-3      Month
5-9      Average Daily Maximum Temperature
11-15    Average Daily Minimum Temperature
17-21    Average Monthly Rainfall (mm)
23-27    Average Daily 9am Relative Humidity (%)
29-33    Average Daily 3pm Relative Humidity (%)

```

(b) Comma-delimited format

The Continent, Country and (optional) State lines start with a # symbol in this format. The lines that correspond to a location must contain the following items, separated by commas:

Latitude, Longitude, Elevation, Location Name, Importance level

Note the "Location Name" may not contain a comma. An example of a comma-delimited .loc file is shown below:

```

#Continent, Africa
#Country, Algeria
#State, nostates
32.8,-0.6,1072,Ain-Sefra,0
36.8,3,50,Algiers,0

```

The .MET file contains the following data items, each separated by a comma from the next item:

LOG/MET Files

Month, Maximum Temperature, Minimum Temperature, Rainfall, 9am Relative Humidity, 3pm Relative Humidity

An example of a comma-delimited .MET file containing the data for two locations is shown below:

```
1,12.2,-0.6,10,79,52
2,15.6,1.1,10,76,50
3,17.2,3.9,15,68,36
4,23.3,7.8,10,58,29
5,27.2,11.1,15,59,29
6,32.8,15.6,28,46,26
7,37.8,19.4,8,39,18
8,35.6,18.9,8,46,21
9,32.2,15.6,15,51,29
10,24.4,9.4,28,68,34
11,16.1,5.28,77,46
12,12.8,1.1,18,78,48
1,15.9.4,112,75,66
2,16.1,9.4,109,72,60
3,17.2,11.1,74,71,59
4,20,12.8,41,67,57
5,22.8,15,46,72,60
6,25.6,18.3,15,72,60
7,28.3,21.1,3,73,60
8,29.4,21.7,5,70,60
9,27.2,20.6,41,74,62
10,23.3,17.2,79,72,60
11,18.9,13.3,130,73,63
12,15.6,10.6,137,72,64
```

15 Appendix 3. The Species File format.

CLIMEX species files must have the extension '.exp' to be recognised as such. They consist of a header (everything up to, but not including, the line "Begin Values") and a values block (the rest of the file). The header block contains the species notes/comments (between the "Begin Comment" and "End Comment" lines). The "Protected" keyword, if present, prohibits the user from editing this file within CLIMEX. In the values block, parameters are hierarchically nested within their processes. Each parameter is specified in 2 lines, as follows:

```
"Begin Parameter" "parameter name" "value"
```

```
"End Parameter"
```

Be careful when changing any parameter value manually to **change only the value of the parameter**. The name must never be changed – it must correspond to the name that CLIMEX uses for that parameter – or CLIMEX will not be able to find that parameter in the file. It is also important to retain the quotes around each value.

```
"DYMEX Parameter Version" "2.1"
```

```
"Begin Parameter Block" "ClimexModel"
```

```
"Begin Description"
```

```
"Name" "_Template - Savannah"
```

```
"Protected"
```

```
"Date Modified" "Jan 02 2001, 12:18"
```

```
"Begin Comment"
```

```
Accommodates tropical savannah-type species
```

```
"End Comment"
```

```
"End Description"
```

```
"Begin Values"
```

```
"Begin Process" "Temperature Index"
```

```
"Begin Parameter" "Limiting low temperature" "18"
```

```
"End Parameter"
```

```
"Begin Parameter" "Lower optimal temperature" "28"
```

```
"End Parameter"
```

```
"Begin Action" "Maximum Temperature"
```

```
"Begin Parameter" "Upper optimal temperature" "32"
```

```
"End Parameter"
```

```
"Begin Parameter" "Limiting high temperature" "40"
```

```
"End Parameter"
```

```
"End Action"
```

```
"End Process"
```

```
"Begin Process" "Moisture Index"
```

```
"Begin Action" "Soil Moisture"
```

```
"Begin Parameter" "Limiting low moisture" "0.3"
```

```
"End Parameter"
```

```
"Begin Parameter" "Lower optimal moisture" "0.6"
```

```
"End Parameter"
```

```
"Begin Parameter" "Upper optimal moisture" "0.7"
```

```
"End Parameter"
```

```
"Begin Parameter" "Limiting high moisture" "0.8"
```

```
"End Parameter"
```

```
"End Action"
```

```
"End Process"
```

"Begin Process" "Light Index" "Unused"
 "Begin Action" "Daylength"
 "Begin Parameter" "Daylength at no growth" "0"
 "End Parameter"
 "Begin Parameter" "Daylength at maximum growth" "0.01"
 "End Parameter"
 "End Action"
"End Process"
"Begin Process" "Diapause Index" "Unused"
 "Begin Parameter" "Diapause Induction Daylength" "0"
 "End Parameter"
 "Begin Parameter" "Diapause Induction Temperature" "0"
 "End Parameter"
 "Begin Parameter" "Diapause Termination Temperature" "0"
 "End Parameter"
 "Begin Parameter" "Diapause Development Days" "0"
 "End Parameter"
 "Begin Parameter" "Summer/Winter Diapause Indicator" "0"
 "End Parameter"
"End Process"
"Begin Process" "Cold Stress" "Unused"
 "Begin Action" "Minimum Temperature"
 "Begin Parameter" "Cold Stress Temperature Threshold" "2"
 "End Parameter"
 "Begin Parameter" "Cold Stress Temperature Rate" "0"
 "End Parameter"
 "End Action"
 "Begin Action" "Degree-days above Threshold"
 "Begin Parameter" "Cold Stress Degree-day Threshold" "15"
 "End Parameter"
 "Begin Parameter" "Cold Stress Degree-day Rate" "0"
 "End Parameter"
 "End Action"
 "Begin Action" "Average Temperature"
 "Begin Parameter" "Cold Stress Temperature Threshold (Average)" "0"
 "End Parameter"
 "Begin Parameter" "Cold Stress Temperature Rate (Average)" "0"
 "End Parameter"
 "End Action"
"End Process"
"Begin Process" "Heat Stress"
 "Begin Action" "Maximum Temperature"
 "Begin Parameter" "Heat Stress Temperature Threshold" "36"
 "End Parameter"
 "Begin Parameter" "Heat Stress Temperature Rate" "0.0002"
 "End Parameter"
 "End Action"
 "Begin Action" "Degree-days above Threshold"
 "Begin Parameter" "Heat Stress Degree-day Threshold" "0"
 "End Parameter"
 "Begin Parameter" "Heat Stress Degree-day Rate" "0"
 "End Parameter"
 "End Action"
"End Process"
"Begin Process" "Dry Stress"

"Begin Action" "Soil Moisture"
 "Begin Parameter" "Dry Stress Threshold" "0.2"
 "End Parameter"
 "Begin Parameter" "Dry Stress Rate" "-0.005"
 "End Parameter"
 "End Action"
 "End Process"
 "Begin Process" "Wet Stress"
 "Begin Action" "Soil Moisture"
 "Begin Parameter" "Wet Stress Threshold" "1.3"
 "End Parameter"
 "Begin Parameter" "Wet Stress Rate" "0.005"
 "End Parameter"
 "End Action"
 "End Process"
 "Begin Process" "Cold-Dry Stress" "Unused"
 "Begin Action" "Degree-days above Threshold"
 "Begin Parameter" "Cold-Dry Degree-day Threshold" "0"
 "End Parameter"
 "End Action"
 "Begin Action" "Soil Moisture"
 "Begin Parameter" "Cold-Dry Moisture Threshold" "0"
 "End Parameter"
 "End Action"
 "Begin Parameter" "Cold-Dry Stress Rate" "0"
 "End Parameter"
 "End Process"
 "Begin Process" "Cold-Wet Stress" "Unused"
 "Begin Action" "Degree-days above Threshold"
 "Begin Parameter" "Cold-Wet Degree-day Threshold" "0"
 "End Parameter"
 "End Action"
 "Begin Action" "Soil Moisture"
 "Begin Parameter" "Cold-Wet Moisture Threshold" "0"
 "End Parameter"
 "End Action"
 "Begin Parameter" "Cold-Wet Stress Rate" "0"
 "End Parameter"
 "End Process"
 "Begin Process" "Hot-Dry Stress" "Unused"
 "Begin Action" "Maximum Temperature"
 "Begin Parameter" "Hot-Dry Temperature Threshold" "0"
 "End Parameter"
 "End Action"
 "Begin Action" "Soil Moisture"
 "Begin Parameter" "Hot-Dry Moisture Threshold" "0"
 "End Parameter"
 "End Action"
 "Begin Parameter" "Hot-Dry Stress Rate" "0"
 "End Parameter"
 "End Process"
 "Begin Process" "Hot-Wet Stress" "Unused"
 "Begin Action" "Maximum Temperature"
 "Begin Parameter" "Hot-Wet Temperature Threshold" "0"
 "End Parameter"

LOG/MET Files

"End Action"
"Begin Action" "Soil Moisture"
 "Begin Parameter" "Hot-Wet Moisture Threshold" "0"
 "End Parameter"
"End Action"
"Begin Parameter" "Hot-Wet Stress Rate" "0"
"End Parameter"
"End Process"
"Begin Action" "Day-degree accumulation above DV0"
"End Action"
"Begin Action" "Day-degree accumulation above DV3"
 "Begin Parameter" "Dummy Parameter 1" "100"
 "End Parameter"
"End Action"
"Begin Action" "Day-degree accumulation above DVCS"
 "Begin Parameter" "Cold Stress DD Threshold Temperature" "18"
 "End Parameter"
"End Action"
"Begin Parameter" "Degree-days per Generation" "650"
"End Parameter"
"End Values"
"End Parameter Block"

16 Version history

16.1 Major Changes from Version 2

Two species can now be fitted with interactions between them (either competition or synergy) specified via parameters. The Biotic Index that describes this interaction is described in Section 5.3.8.

Radiation (as **Radiation Index**, section 5.3.4) is available as an additional component to the Growth Index.

Two non-specific components (definable by the user) can be added to the Growth Index. These are the **Physical Substrate Index** and the **Biotic Substrate Index** (Section 5.3.5). The variables determining these indices can either be specified as a single value for all locations or they can be location specific and read from the MetManager.

Automatic fitting of the parameter values that determine the Stress indices is available via a genetic algorithm based fitting routine (Section 6.2).

“Regional matching” (i.e., the use of a set of locations for the ‘Home’ location in the **Match Climates** function) is available (see Section 0.2).

The MetManager application has been extended to allow the importation of up to 5 user-defined location constants as well as up to 5 user-defined variables.

16.2 Major changes from Version 1.1 to 2

CLIMEX has been considerably enhanced from Version 1.1. For users familiar with Version 1.1, the following list details the major differences. Those changes that may cause models to behave differently when moved to Version 2 are shown in italics.

The various CLIMEX applications are now DYMEX models. This has meant a radical change to the user interface and operation of the program which will require some relearning by users of CLIMEX V1.1. However, it has also allowed CLIMEX to use the very flexible display facilities provided by DYMEX. In particular, map capabilities are now much more powerful, with zoom and pan support. In addition, maps are constructed from standard shape files, many of which are available from various internet sites. A new MapManager allows map region files to be defined visually. Many map projections have been provided. Map printing has been enhanced and a Print Preview function provided.

The same flexibility is also now available for Tables and Charts, which can be configured in a multitude of ways, and the configurations can be saved and re-used in subsequent runs. Results can be saved as Microsoft Access® database tables. Charts and maps can be exported in raster (Windows bitmap, bmp or jpg) or vector (Windows enhanced metafile, emf) form.

The new MetManager program is a Microsoft Access application and overcomes the data importing problems that users commonly had with the previous version. It is easy to create multiple databases of meteorological data and Location Selections from them. Data can still be imported with the .loc/.met format, as well as from .bmt files

or Access files. A 0.5° world grid of meteorological data is included for those tasks where the standard database is not sufficient. Acknowledgement should be given to the Climatic Research Unit (CRU), Norwich (<http://www.cru.uea.ac.uk/cru/data/hrg.htm>) when the data are used. A reformatted 0.1° grid is available for non-commercial applications by courtesy of the CRU Norwich (<http://www.cru.uea.ac.uk/cru/data/hrg.htm>) and users are required to agree to comply with CRU restrictions when they download the database from the CSIRO/Hearne website.

Soil Moisture and Average Temperature have been added as variables that can be included in the Match Index for Match Climates.

Species Parameters are now stored in a separate file for each species (with the extension .exp). The comments related to a species are stored in its parameter file. A conversion utility (cxdxconvert.exe) is provided to convert species and comment files from the Version 1 format to Version 2.

The degree-day dependent Cold Stress uses degree-days based on a new threshold parameter (DVCS) rather than DV0. Note, however, that the default value of DVCS is the same as DV0. The new flexibility allows the users to use thresholds other than DV0.

A third component has been added to the Cold Stress. This component acts in the same way as the component based on minimum temperature, but uses average temperature instead of the minimum. It is useful when comparing responses to climates with different diurnal ranges of temperature.

The Interaction Stresses now include the time since the stress began as part of the calculation (as is done for all other Stress Indices). When such parameter files are converted using the conversion utility, the new parameter file will contain an indicator that directs CLIMEX to use the old Interaction Stress method. These species should preferably be refitted to use the new method.

The differences in the Ecoclimatic Index (EI) can be plotted on a map after two sequential runs using the same locations. Thus users can quickly appreciate the changes that would result from, for example, the application of a Climate Change Scenario or a change in a parameter value.

17 Index

- abundance, 27, 56, 63
- Access database, 143, 182
- annual plants, 44
- arid, 57
- Average Temperature, 47, 63, 104, 170, 182
- Away location, 24, 25, 103, 106
- bimodal rainfall, 44
- biological control, 10, 24
- Biotic Index, 31, **43**, 54, 169, 181
- Biotic Substrate, 12
- Biotic Substrate Index, 12, 13, 40, 149, 169, 181
- BMT files, 122, 182
- census data, 56
- Climate Change, **116–19**
- climate database, 29
- climate matching, 24
- climatic preferences, 29, 57
- CLIMEX Application, 8
 - Compare Locations (1 species), 11
 - Compare Locations (1 species, extended), 12
 - Compare Locations (2 species), 11
 - Compare Locations (Genetic Algorithm), 13, 93
 - Compare Locations (Species Interaction), 13
 - Compare Years, 14
 - Match Climates, 24
 - Match Climates (Regional), 26
- CLIMEX Model, 3, 8, 9, 27
- CMI, 103, 105
- Cold Stress, **46–48**, 62, 169
- Cold-Dry Stress, 51, 171
- Cold-Wet Stress, 51, 172
- community, 31
- Compare Locations, **9**, 11
- Compare Years, **14**
- competition, 13, 43, 181
- competitive exclusion, 54
- Composite Match Index, 103
- conversion utility, 182
- core distribution, 59, 62, 82, 87
- Core Distribution**, 55
- crossover rate, 87, 94
- cumulative stress, 45, 48, 49
- data averaging, 15
- data sets, 15
- daylength, 41, 42, 52, 63, 66
- degree-days, 33, 46, 48
- density-dependence, 32
- DHCS, 46, 170
- DHHS, 48, 171
- diapause, 30, 31, 40–42, 52, 63, 73
 - facultative, 167
 - Genetic algorithm, 94
 - obligate, 4, 30, **52–53**, 66, 73
 - summer, 42, 73
 - winter, 41, 66, 73
- Diapause
 - obligate, 40
 - Summer or Winter indicator, 41
- diapause development, 41, 63, 74
- Diapause Index, 31, **40–42**, 42, 53, 167
- differences, 183
- distribution, 27
- distribution data, 59
- distribution limits, 59, 62
- distribution maps, 56
- DLD0, 63
- DLT0, 63
- DLT1, 63
- dormancy, 31

- DPD, 41, 42, 67, 167
DPD0, 41, 67, 167
DPSW, 41, 67, 167
DPT0, 41, 67, 167
DPT1, 41, 67, 167
Dry Stress, **49**, 70, 171
DTCD, 171
DTCS, 46, 170
DTCW, 172
DTHS, 48, 171
DV0, 32, 48, 52, 63, 67, 166, 182
DV1, 32, 63, 166
DV2, 32, 63, 166
DV3, 32, 48, 63, 166
DVCS, 46, 170, 182
DYMEX, 2, 122, 181
Ecoclimatic Index, 4, 29, 30, 33, 41, 53–**54**, 59
El. see Ecoclimatic Index
equatorial zone, 118
Evaporation, 144
evapotranspiration, 35
extreme seasons, 9
extreme temperatures, 47, 48
extreme values, 15
extremes, 44
generation, 32, 33, 52, 69
generations, 94
Genetic Algorithm, **82–96**
geographical distribution, 9, 44, 56
geography, 59
gradients, 59
grid data, 122, 182
Growth Index, 56
 Annual, 3, **31**, 41, 53
 Weekly, **31**, 41, 43, 53, 74
growth season, 4
HDS, 49, 171
Heat Stress, **48–49**, 69, 170
Home location, 25, 103
Home Region, 106
host availability, 27
Hot-Dry Stress, 51, 172
Hot-Wet Stress, 51, 172
HWS, 50, 171
hypothesis, 59, 62
import data, 122
Import formats, 173
Index
 Cold Stress, 182
 Ecoclimatic, 4
 Growth, 3
 Interaction Stress, 182
 Stress, 3
inference, 4, 56
installation, 5
Interaction Stress, **50–52**, 171
irrigation, 11, 70, **119**
iteration, 60
iterative, 56
known distribution, 61, 86
laboratory data, 4, 56, 59, 60
Law of the Minimum, 3
Law of Tolerance, 3
LGMC, 37
LGRC, 39, 168
Light Index, 31, **42**, 63, 167
limiting conditions, 9
Location Selection, 61, 120, 144, **146–48**, 182
Location Sequence, 60, 61, 120, 144
LT0, 42, 167
LT1, 42, 167
Map, **122–26**

- Legend, 126
- Projection, 125, 181
- Region Definition, **122**, 126
- Region File, 181
- Zoom, 126
- Map Manager, **122–26**, 181
- masking, 103
- Match Climates, 3, **103–5**
 - Regional, **105–9**
- Match Index, 103, 105
 - Humidity, 105
 - Rainfall Pattern, 104
 - Soil Moisture, 105
 - Temperature, 104
 - Total Rainfall, 104
- Match Index Weights, 103
- Match Level, 103
- Metadata
 - Selection, 144
 - Selection Criteria, 144
- Metdatabase
 - Creation, 144
- MetManager format, 173–75
- MetManager program, 120, **144**, 173, 182
- MIF, 37
- Moisture Index, 31, **35–36**, 166
- Moisture Parameters, 64
- Moisture Stress, 70
- MTCD, 171
- MTCW, 172
- MTHD, 172
- MTHW, 172
- multivoltine, 32
- mutation rate, 87, 94
- obligate diapause. *see* diapause, obligate
- parameter, 9, 111
- Parameter File
 - comments, 115
- parameter fitting, 4, **56–77**
 - automatic, **82–96**, 181
- Parameter Grid window, 112
- PCD, 171
- PCW, 172
- PDD, 30, **33**, **52**, 53, 62, 68, 166
- pest risk analysis, 10, 24
- Phanaerophytes, 79
- PHD, 172
- phenology, 4, 56, 63
- PHW, 172
- Physical Substrate Index, 168, 181
- population, 3
- population growth, 9, 30, 31
- population growth rate, 63
- population size, 94
- potential distribution, 9
- quarantine, 24
- Radiation Index, **37–39**, 167, 181
- rate of increase, 32
- Reference File, 86, **88**, 89, **91**
- Regional Matching. *see* Match Climates, Regional
- RIF, 168
- RIF_i, 39
- RV0, 37, 167
- RV1, 37, 168
- RV2, 37, 168
- RV3, 37, 168
- Scenario Files, 117
- seasonal growth, 31
- seasonal growth potential, 10
- sensitivity, 56, 59
- Sensitivity Analysis, 96
- Sequence. *see* Location Sequence
- Shape File, 123, 181
- Simulator (DYMEX), 110

- SIP0, 169
- SM0, 36, 64, 166
- SM1, 36, 64, 166
- SM2, 36, 64, 166, 167
- SM3, 36, 64, 166, 167
- SMDS, 49, 64, 171
- SMWS, 50, 64, 171
- Soil Moisture, 64, 79, 103, 182
- soil moisture model, 35
- soil moisture storage, 35
- Soil Moisture Utility, 64
- soil type, 27, 35, 56
- species comments, 60
- Species File, 60, 111, 112, 182
- creation, 112
 - deletion, 112
 - folder, 111
 - printing, 112, 116
 - protected, 115
- species interaction, 13, 37, 38, 43, 54, 82
- species name, 116
- species template, 57, 59, 65
- stress, 9
- stress accumulation, 44, 45, 51, 56
- stress boundaries, 61
- stress gradient, 62
- Stress Index, 44–46, 56, 59, 169
- Stress Indices, 30
- Substrate Index, 39–40
- sub-tropical, 57
- summer diapause. *see* diapause, summer
- SVB0, 169
- SVB1, 169
- SVB2, 169
- SVB3, 169
- SVP0, 168
- SVP1, 168
- SVP2, 168
- SVP3, 168
- synergy, 13, 43, 181
- temperate, 57, 65
- Temperature Index, 31, 32–35, 34, 166
- Annual, 35
- Temperature parameters, 63
- Template (Species Analysis), 28
- THCS, 47, 170
- THCSA, 48, 170
- THHS, 49, 170
- tropical, 57
- TTCS, 47, 170
- TTCSA, 47, 170
- TTHD, 172
- TTHS, 49, 170
- TTHW, 172
- validation, 60
- Validation
- cross-validation, 4
- variation, 9
- vernalisation, 40
- Wet Stress, 50, 171
- winter diapause. *see* diapause, winter